



SERRC
Social Epistemology
Review & Reply Collective

<http://social-epistemology.com>
ISSN: 2471-9560

Can We Tell Whether Philosophy is Special?

Chad Gonnerman, University of Southern Indiana, cgonnerman@usi.edu; Stephen Crowley,
Boise State University, stephencrowley@boisestate.edu

Gonnerman, Chad and Stephen Crowley. 2022. "Can We Tell Whether Philosophy is Special?" *Social Epistemology Review and Reply Collective* 11 (11): 1-11. <https://wp.me/p1Bfg0-7jm>.

In “Is Philosophy Exceptional? A Corpus-Based, Quantitative Study” (2022), Moti Mizrahi and Michael Adam Dickinson use corpus methods to determine the kinds of arguments that turn up in philosophical writing. They use the results to contribute to debates on philosophy’s “specialness” or “exceptionality”. To what extent is philosophy interestingly unlike other knowledge-making disciplines? Specifically, does it deploy different forms of argument than the sciences or other disciplines?

These questions are interesting, and Mizrahi and Dickinson’s methodological approach is impressive. Nonetheless, we have concerns about the approach. In this short reply, we argue there are reasons to worry about the accuracy of the results that the authors report and reasons to wonder how their results, even if accurate, shed light on questions about philosophy’s specialness. We end with some thoughts about the strengths and weaknesses of the project and some suggestions for next steps for those interested in moving this project (and others of its ilk) forward.

1. A Reconstruction

The methodological approach taken by Mizrahi and Dickinson is sufficiently novel (but cf. Ashton and Mizrahi 2018) that we want to begin with a brief description of their method, as best as we could reconstruct it from the body of the paper and the appendix. That way, the reader will have a better sense of what it is that we are talking about in our reply. It will also provide the authors with an opportunity to correct any errors or omissions in our reconstruction.

One of Mizrahi and Dickinson’s main goals was to go some way towards capturing the sorts of arguments that philosophers give. Of course, philosophers give lots of arguments across many contexts. Conferences, conversations, colloquia, and classrooms are just a few. But since records of such exchanges are mostly non-existent, the authors rightly focus on philosophical publications. Their data-collection strategy on this front went through two main stages. The first involved settling on a collection of *indicator pairs*. Each pair consisted of a conclusion indicator (e.g., ‘therefore’) and a word or phrase deemed to be an indicator of the type of argument at play (e.g., ‘necessarily’ for deductive, ‘likely’ for inductive, and ‘best explains’ for abductive; for the full list, see Table 2, page 7). The second main stage of data collection involved carrying out a series of searches of JSTOR’s corpus of philosophical texts for the years 1867 to 2014. For each indicator pair, the authors searched the corpus for strings of varying lengths that included both members of the pair. For example, in their *ten-word searches*, a hit was returned for the pair ‘therefore ... necessarily’ if both terms occurred within ten words of each other, allowing for cases where ‘therefore’ appeared before ‘necessarily’ or after.

Scoring for the different argument types (deductive, inductive, or abductive) seemed to work as follows. If publication P published in year Y used the deductive indicator pair being searched for (e.g., ‘therefore ... necessarily’) within the search’s word-range (e.g., within ten words), then P was categorized among the *deductive publications* for Y. Matters appeared to

work similarly for the inductive and abductive indicator pairs. What remains a bit obscure to us, however, is how the scoring worked when a publication contained two or more indicator pairs. What if P from 1950 had used both ‘therefore ... necessarily’ and ‘hence ... certainly’? Evidently, the search algorithm would have delivered two deductive hits for P (see second paragraph on p. 18). But it appears that one of the hits would have been removed so that P counts only once for the category of deductive publications in 1950 (see first paragraph on page 18). Or, to take another kind of case, what if P published in 1962 had used a deductive indicator pair (e.g., ‘therefore ... necessarily’) and an inductive indicator pair (e.g., ‘follows ... unlikely’)? As best as we can tell, Mizrahi and Dickinson don’t directly comment on this possibility. Our best guess is that P would have been counted as a deductive *and* an inductive publication for 1962.

Setting aside complications in scoring the articles, it is reasonably clear that publications were grouped by year and a *deductive ratio* was calculated for each year as follows: the number of papers “testing positive” for deductive publications in that year were compared to the total number of papers published in the year. This resulted in a total of 148 deductive ratios, one for each year between 1867 and 2014. Similar procedures were used to compute annual *inductive* and *abductive ratios*. Armed with these values, Mizrahi and Dickinson could then compare the average ratios. They consistently found that the mean deductive ratio was larger than the mean inductive ratio, which in turn was consistently larger than the mean abductive ratio (9-11). In addition, our authors computed a series of differences in annual ratios (deductive-inductive, deductive-abductive). Regression analyses indicate that these differences are getting smaller over the years. It would thus appear that while philosophers are, on the whole, more likely to advance deductive arguments than inductive or abductive arguments, this tendency is weakening through the years (9-12).

2. Puzzles over Specialness

One might think philosophy is “special” for a variety of reasons, but roughly “specialness” will show up in either product or process. That is, either philosophical products or philosophical processes will be interestingly different from the products and processes of other forms of knowledge-making. Our authors focus on *process specialness*. For example, in the abstract, immediately after raising the question of whether philosophy is special, they ask, “Are its methods (dis)continuous with science?” We will follow the authors in their focus.

For Mizrahi and Dickinson, process specialness amounts to something about the combination of deductive, inductive, and abductive arguments deployed by a discipline. But it wasn’t always clear to us what about this combination would make a discipline like philosophy special, or exceptional, were the argumentative details to pattern out in the right ways. It appears that our authors toggle between two possibilities:

1. Philosophy is special because most of its arguments are deductive in contrast to science where we have some reason to think that most arguments are non-deductive (e.g., inductive, abductive) in nature (e.g., 13).

2. Philosophy is special because its form(s) of argument are interestingly different from those of other disciplines (e.g., 1). (On this version, our authors point out that their work can only be preliminary as we currently lack data on the form(s) of argument used in other disciplines.)

In the case of the first option, Mizrahi and Dickenson suggest (12-13) that philosophy *is* special since, in view of their empirical results, it appears that, historically, philosophy has mostly used deductive arguments *but* that this may change before too long since philosophy's argumentative tendencies are drifting towards non-deductive modes of inference. But is this the conclusion suggested by their data? The differences in deductive and inductive ratios graphed in Figures 1, 3, and 5 (9, 10, 12) appear to have swung fully in favor of inductive ratios at least by the year 2000.¹ And if we take abductive arguments into account, the swing towards non-deductive inferences probably happened even earlier. So, we are unclear why the authors don't conclude that philosophy was special but is no longer. There must be something going on with 'specialness' here that we don't fully appreciate.

In any case, our major worry lies with the second option. We'll offer our worry as a dilemma: Can all (or at least the vast majority) of arguments be classified as either deductive, inductive or abductive?

- If the answer is 'no', then it seems that paying attention to just how many deductive, inductive, or abductive arguments are used in a discipline won't settle the issue of how process special it is since much of what goes on in a discipline does not boil down to how deductive, inductive, and abductive it is.
- If the answer is 'yes', then it seems that paying attention to just how many deductive, inductive, or abductive arguments are used in a discipline won't settle the issue of how process special it is since all argument-giving disciplines, including mathematics (Baker 2020), engage in some combination of deductive, inductive, and abductive argumentation and so we won't have the sort of information that helps us to set the discipline apart and identify it as special. Instead, philosophy will be just one point among many smeared across a multi-dimensional space of argumentative proclivities.

¹ Our estimation of the point where the inductive ratios overtake the deductive ratios is based on an eyeball inspection of the figures. Another possibility is to lean on the regression equations that the authors report in each of the figures. But when we do, we get some implausible results. For example, for the *six-word searches* where up to six words could appear between conclusion and argument-type indicators, Mizrahi and Dickenson report the following equation: $y = -0.0004x + 0.7553$. Evidently, y is the difference between deductive and inductive ratios, and x is the year of publication. If so, then we can calculate the year at which the regression analysis predicts that the difference between the two ratios will be zero. We then discover that $y = 0$ when $x = 1888$. So, the prediction is that the inductive ratios will overtake the deductive ratios by 1889? (The regression equation for the three-word searches delivers a more plausible value of 1995 for when $y = 0$; the equation for the ten-word searches produces a wildly implausible value of 1787.)

Consider the following example. Field ecologists use statistical methods to identify patterns in the information they collect. So do ecological modelers. In the one case the data are generated by fieldwork and in the other by simulation. Whether we say these two areas of study use the same processes or not will depend on how you make sense of ‘process’: do data generating practices count as part of the process or not? Furthermore, even if you treat the two processes as the same (use of statistical inferences) does it make sense to say the areas are (or aren’t) interestingly different? That will depend on what makes a difference interesting and that is exactly what we want to know more about.

In sum, before we can work out the implications of Mizrahi and Dickinson’s empirical work for metaphilosophical debates on the specialness of philosophy, we believe some degree of clarification is in order. What counts as special? And how do their measures help to reveal philosophy’s specialness, if there is any to be disclosed?

3. Philosophy’s Drift: The Interesting and Less Interesting

One worry raised against digital humanities (DH) is that it rarely delivers anything interesting. In an acerbic and sometimes hyperbolic critique of DH, Kirsch (2014) repeatedly turns to this worry. About the Google NGram Viewer, which allows users to graph the occurrence of words or phrases in Google books over time, he writes, “it turns out once again that the digital analysis of literature tells what we already know...” This is a point that Kirsch reinforces in connection with Moretti’s (2009) finding that the title of novels published in Britain between 1740 and 1850 got shorter; Kirsch states, “you hardly need a computer to tell you that: the bulky eighteenth-century title is commonplace and a target of jokes even today.”

When charges of humdrum digital results are leveled against patterns produced for the express purpose of figuring out whether the underlying tools perform as expected, as sometimes occurs in humanistic criticisms of DH, the charges land quite wide of the mark (Hammond 2017). To borrow an example from the metaphilosophical literature on intuitions (Cummins 1998), it would be a bit like objecting to Galileo’s telescope on the grounds that it doesn’t tell us anything new when we point it at a well-explored mountain in the distance. “That’s the whole point,” Galileo might note, “when calibrating your instrument.” But perhaps something like the oft-heard critique of DH applies to one aspect of Mizrahi and Dickinson’s results. Maybe a case could be made that there is something a tad uneventful about their finding that philosophy is now more non-deductive than it once was.

The leveraging of systematically gathered empirical data towards philosophical ends has a long history (e.g., Sytsma et al. forthcoming). Still, there is something different going on nowadays. There is just so much more empirical engagement happening in philosophy compared to the previous century. For example, as our authors note (2), when Knobe (2015) compared the most highly cited philosophical articles on the mind published between 1960-1999 to those from 2009-2013, he found a significant increase in works that rely on empirical results. Also, surveys of philosophers of science point to a widespread interest and experience in philosophical work that is deeply engaged with the sciences (Plaisance et al.

2021). Further, an examination of PhilPapers indicates that there are on average around 1.16 times more publications falling in the experimental philosophy category *per year* for the years 2000 to 2020 (Sytsma et al. forthcoming). Given such indications of increased attention to and leveraging of empirical or scientific data in philosophy, along with the reasonable suggestion that such data lend themselves to non-deductive modes of inference, it isn't surprising that Mizrahi and Dickinson found an arc towards the non-deductive in philosophy. If anything, if they hadn't, we might wonder about the reliability of their approach.

Importantly, none of this is to say that there is nothing interesting about the empirical work reported by our authors. Very far from it. For starters, the case we just sketched for expecting a trend towards non-deductive arguments in philosophy is probabilistic at best, so some degree of confirmation can only help. Moreover, out of their corpus-analytic work, Mizrahi and Dickinson don't merely extract an indication that there has been a drift towards the non-deductive in philosophy; they also work out some estimates of the magnitude of the trend. For example, when it comes to their ten-word searches, their regression analysis of differences in deductive and inductive ratios delivered a regression coefficient of -0.0005 . This indicates that we can expect the difference between these two ratios to decrease each year by around 0.0005 . But how trustworthy are their estimates? We turn to this question next.

4. Sensitivity and Specificity

Good estimation requires good measurement. At the heart of Mizrahi and Dickinson's measurements are their *classification schemes*: the set of rules they used for mapping philosophical publications into givers of different argument types. Thus, any reasonably complete appraisal of their measurements of argument types in philosophy has to include an assessment of their classification schemes. How do they fare?

When summarizing the performance of a classification scheme designed for empirical research purposes, Fox et al. (2021, 143) note that researchers typically deploy a handful of measures. Here, we focus on two of these measures and associated proportions. First, there is sensitivity. In epidemiological research, the *sensitivity* of a classification scheme is the probability that a subject who was exposed to the factor of interest (e.g., maternal smoking during pregnancy) is correctly categorized as having been exposed to the factor. This value gives rise to the scheme's *false-negative proportion*: the proportion of subjects who were exposed to the factor but who were incorrectly categorized among the unexposed. The other measure of interest for us is the scheme's *specificity*. In epidemiology, this is the probability that a subject who was not exposed to the factor being studied is correctly categorized as being unexposed. The *false-positive proportion*, then, is the proportion of subjects who were unexposed but were categorized among the exposed.

Applied to Mizrahi and Dickinson's work, the foregoing gives rise to a large set of numbers. For example, there is the sensitivity of their scheme for categorizing philosophical publications among abductive publications by searching for abductive indicator pairs

allowing up to three words between pair members. The false-negative proportion of this scheme would be $1 - \text{this sensitivity value}$. And there is the specificity of their scheme for finding inductive publications that relies on inductive indicator pairs with no more than six words between each. Its false-positive proportion would be $1 - \text{its specificity}$. All told, since in effect the authors deploy three classification schemes for each of three argument types, a complete summary would consist of nine sensitivities, nine specificities, and 18 proportions. Fortunately, much of what we have to say about the authors' classification schemes over the next couple sections can be said without all this numerical detail.

5. Sensitivity and False Negatives

Mizrahi and Dickinson note that their study of philosophical arguments covered 435,703 publications (1). That is a large number. No single philosopher could read that many papers. It seems to be the kind of task well-suited for *distant reading*, where, to borrow from Pence and Ramsey (2018), we let a computer “do the ‘reading’ for us, extracting large-scale patterns that would be invisible or impractical to find otherwise” (931).

Also quite large, however, is the number of publications that, in effect, didn't make their way into the comparisons or statistical models reported by Mizrahi and Dickinson. This large number is a direct consequence of the low sensitivity of their classification schemes. Exactly how low they are is hard to say without more data. Ideally, what is wanted here are the results of a well-designed validation study (see Fox et al. 2021, 153-154). But, even in the absence of such results, we can go some way towards substantiating the claim that the sensitivities at play in the paper's classification schemes are probably quite low, and thus that their false-negative proportions are fairly high.

Every philosophical publication has at least one argument. Also, every argument is either deductive, inductive, or abductive. One upshot is that a fully sensitive classification scheme for detecting deductive, inductive, and abductive publications will capture *every* article in the set of philosophical publications to which the scheme was applied. But, turning to the Mizrahi and Dickinson's classification schemes, if we add up the average ratios of publications that they classified as giving a deductive, inductive, or abductive argument, we fall way short of capturing every article in the corpus. For example, when it comes to their ten-word searches, we get to a ratio of 0.1418 (deductive) + 0.0698 (inductive) + 0.0603 (inductive) + 0.0177 (abductive). So, when the three ratios ought to sum to one (or probably more, since a single article can give more than one type of argument), Mizrahi and Dickinson's most liberal classification schemes get us only about 14% of the way there. Approximately 86% of JSTOR's philosophical publications are skirting past their most permissive schemes.

With that said, it may be that the argument we just gave is a bit too strong. One of its central premises is that every philosophical publication gives at least one argument. Maybe there are exceptions. Also, our argument is premised on the claim that every argument is either deductive, inductive, or abductive. Perhaps there are others. We are happy to concede as much. But notice that, when we do, we aren't forfeiting much ground. The number of argument-less philosophical publications is probably tiny. The number that advance

arguments outside the deductive- inductive-abductive triad is likely small as well. So, these are mostly quibbles at the margins of our argument. The point remains that, collectively, the false-negative proportions for our author's most liberal classification schemes hover somewhere north of 0.80. This translates to around 350,000 papers that are, in effect, getting incorrectly categorized as not giving either a deductive, inductive, or abductive argument.

Of course, there is nothing inherently wrong in drawing conclusions about the relative sizes of various categories in a population, whether at a time or across a time period, using a classification scheme that captures only a subset of the population. Mizrahi and Dickinson aren't carrying out a philosophical census, after all. But when drawing conclusions about the relative sizes of the targeted categories it is desirable to have some reason to think that the scheme being deployed results in sample categories whose relative sizes adequately reflect the relative sizes of population categories. Thus, it is unwise to rest one's predictions of who is going to win an unfolding election entirely based on early-vote tallies, at least in the U.S. And this is all the more true when trying to predict the *magnitude* of the victory.

In response, the reader might suggest that the most reasonable default assumption is that what goes for philosophical publications categorized according to the classification schemes of Mizrahi and Dickinson's paper likely goes for the overall population of philosophical publications. Thus, a trend towards a greater use of non-deductive indicator pairs in philosophical publications probably approximates the size of a similar trend towards non-deductive arguments in general. On this front, we are sympathetic to the thought. But we hasten to add that default assumptions make for a thin basis to rest empirical research on. When doing empirical, especially academic, research, we rightfully ask for more, when we can get it. Let us be clear: our goal is *not* a volley in a game of burden tennis with our authors. Rather, we suggest the burden of doing better lies on all of us, or at least all of us interested in understanding patterns and trends in philosophical argumentation.

6. Specificity and False Positives

Sensitivity values and false-positive proportions are just one dimension along which we can assess the performance of a classification scheme. Another dimension focuses on the scheme's specificity and false-positive proportion. In this section, we raise the worry that the specificity of the classification scheme used by Mizrahi and Dickinson for detecting inductive arguments may be getting worse over time.

Recall that the sensitivity of a classification scheme relates to the probability that it correctly categorizes the non-members of the target category outside the category. Thus, a classification scheme for detecting inductive publications will have a low specificity score if it slurps up many publications that don't give any inductive arguments. "Treat any philosophical publication as an inductive publication" would have a low specificity. And if one's rule started with a fixed percentage of randomly selected publications getting placed in the category, with this percentage increasing by, say, a percentage-point a year, the specificity of the scheme would eventually get worse each year until a ceiling is hit. Since the error introduced by this second scheme tends to increase in size across years, it would give rise to

what Viswanathan (2005) calls *systematic correlational error*. It is our contention that the induction-detecting scheme used by Mizrahi and Dickinson may be susceptible to a similar form of error.

Fletcher et al. (2021) argue that philosophers are increasingly using probabilistic methods in their research. The basis of their argument is an empirical study of articles published in *Philosophical Studies* across two time periods (2005-2009 vs. 2015-2019). They found that while the proportion of papers using methods in the logic family stayed similar across the two periods, the proportion using methods in the probability family underwent a threefold increase. One driver of this increase is a doubling in the number of articles coded as using probability theory. This matters for current purposes because an increase in philosophical deployments of probability theory introduces a complication for Mizrahi and Dickinson's inductive scheme. Specifically, since probability theory is centrally concerned with, well, probabilities, deployments of it often use induction-like language, language which will automatically get categorized by the authors' scheme as giving an inductive argument when it is not.

To illustrate, consider Weisberg (2021)'s SEP entry on formal epistemology. In his discussion of the O'Connor-Weatherall model of scientific polarization (e.g., O'Connor and Weatherall 2018), we find the following sentence: "And the blues have become so confident, they're unlikely to ever move close enough to any of the reds to get to their ear." (Here, the "blues" are doctors who have credence above 0.5 in a new medical treatment; the "reds" are doctors with zero credence in the treatment.) Notice that this sentence includes the words 'so' and 'unlikely', and they appear within five words of each other. Thus, for at least a couple of their searches, Mizrahi and Dickinson's classification scheme would categorize this encyclopedic entry as giving an inductive argument. But is there an inductive argument here? Setting aside the fact that 'so' is not functioning as an argument indicator in the sentence, we are talking about a conclusion (purported to be) entailed by the relevant premises and mathematics.² Isn't that a deductive argument that merely looks like an inductive argument?

It is important to note that our point here isn't that the classification scheme is subject to false positives. That is a trifle, at best. Our point is that, given the increased use of probability theory (and other probability methods) in philosophy, the number of deductive inferences masquerading as inductive arguments has probably increased as well, and this introduces an artificial inflation of increasing size to Mizrahi and Dickinson's annual ratios of inductive publications. It just may be that the trend towards non-deductive arguments that Mizrahi and Dickinson report is not quite as strong as their estimates suggest.

² While Mizrahi and Dickinson are probably right that "academic philosophers rarely misuse indicators in an effort to make non-arguments appear as arguments" (5), this sentence helps to illustrate that it is also probably true that philosophers often correctly use indicators in their non-indicator senses.

7. Conclusion

Mizrahi and Dickinson have put together a thought-provoking piece. Their methodological approach is impressive. But it seems to us that there are reasons to think that more work needs to be done before we can be confident that the pattern of results reported by the authors sheds much light on questions about the specialness of philosophy. That said, we hope it is clear that our concerns focus primarily on implementation rather than the project itself. Corpus analysis strikes us as a powerful tool for metaphilosophical work, and the investigation of argumentative practices represents a potentially fruitful way to attack metaphilosophical issues as well as being interesting in its own right. Even though we have raised worries about the way the project has been implemented, we acknowledge that by making the choices necessary for implementation (and then actually implementing it) the authors have provided the rest of us with exactly the sort of information needed to push projects like this one forward. We hope our concerns will present themselves in this light—as suggestions for next steps—rather than a dismissal of this work.

To that end we'd like wrap up by emphasizing what we take to be the most important next steps:

1. The false negative problem. It seems that philosophers use conclusion indicators more reliably than argument-type indicators. Is the failure to mark kind even handed or more prevalent regarding certain types of argument? That is, do philosophers mark the same percentage of each argument type (e.g., 3% of all deductive arguments are marked at deductive arguments)? If so, then marked arguments are a representative sample of argument types as a whole. If not, then we need to supplement our study of marked arguments with more information. Might a qualitative sample help answer the question of how representative the marked argument sample is?
2. The false positive problem. Our grounds for concern here rely on a finding about the increased use of probabilistic methods in philosophy combined with a plausible speculation about how that shift will affect Mizrahi and Dickinson's schemes for detecting inductive publications. What to do about this if we are right is unclear to us at this point. But the first step is to check if we are right. Can we run Fletcher et al.'s data set through Mizrahi and Dickinson's classification scheme to see how those papers are scored and then qualitatively assess a random sample of the paper to see whether their scheme produces more false positives when evaluating articles that use probabilistic methods?
3. The "what is special" problem. This seems like the easiest challenge to address. We are mostly asking for a clarification of what specialness amounts to vis-a-vis philosophy's argumentative patterns. If it is a claim mostly about the prevalence of deductive arguments, then clarity about the implications of the results reported by Mizrahi and Dickinson would be most appreciated. If

the claim is about where philosophy sets in a multi-dimensional space of argumentative tendencies, then some principled way of carving out special places in this space seems needed. Either way, corpus analyses have a role to play.

To conclude, the authors' project strikes us as exciting but its implementation potentially limited. That said, those limitations can and should serve as inspirations for the rest of the community. We are excited to see what comes next.

References

- Ashton, Zoe and Moti Mizrahi. 2018. "Show Me the Argument: Empirically Testing the Armchair Philosophy Picture." *Metaphilosophy* 49 (1-2): 58-70.
- Baker, Alan. 2020. "Non-Deductive Methods in Mathematics." In *Stanford Encyclopedia of Philosophy* edited by Edward N. Zalta. Summer 2020. <https://plato.stanford.edu/archives/sum2020/entries/mathematics-nondeductive/>.
- Cummins, Robert C. 1998. "Reflection on Reflective Equilibrium." In *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry* edited by Michael DePaul and William Ramsey, 113-128. Lanham, MD: Rowman & Littlefield.
- Fletcher, Samuel C., Joshua Knobe, Gregory Wheeler, Brian A. Woodcock 2021. "Changing Use of Formal Methods in Philosophy: Late 2000s vs. Late 2010s." *Synthese* 199: 14555-14576.
- Fox, Matthew P., Richard F. MacLehose, Timothy L. Lash. 2021. *Applying Quantitative Bias Analysis to Epidemiologic Data*, 2nd edition. Dordrecht: Springer.
- Hammond, Adam. 2017. "The Double Bind of Validation: Distant Reading and the Digital Humanities' 'Trough Of Disillusionment'." *Literature Compass* 14: e12402.
- Kirsch, Adam. 2014. "Technology is Taking Over English Departments: The False Promise of the Digital Humanities." *New Republic*. 2 May. <https://newrepublic.com/article/117428/limits-digital-humanities-adam-kirsch>.
- Knobe, Joshua. 2015. "Philosophers are Doing Something Different Now: Quantitative Data." *Cognition* 135: 36-38.
- Mizrahi, Moti and Michael Adam Dickinson. 2022. "Is Philosophy Exceptional? A Corpus-Based, Quantitative Study." *Social Epistemology*. <https://doi.org/10.1080/02691728.2022.2109529>.
- Moretti, Franco. 2009. "Style, Inc. Reflections on Seven Thousand Titles (British Novels, 1740–1850)." *Critical Inquiry* 36 (1):134–158.
- O'Connor, Cailin and James Owen Weatherall. 2018. "Scientific polarization." *European Journal for Philosophy of Science* 8: 855-875.
- Pence, Charles H. and George Ramsey. 2018. "How to do Digital Philosophy of Science." *Philosophy of Science* 85 (5): 930-941.
- Plaisance, Kathryn, Alexander V. Graham, John McLevey, Jay Michaud. 2021. "Show Me the Numbers: A Quantitative Portrait of the Attitudes, Experiences, and Values of Philosophers of Science Regarding Broadly Engaged Work." *Synthese* 198 (5): 4603–4633.

- Sytsma, Justin Joseph Ulatowski, Chad Gonnerman. forthcoming. "All in the Family: The History and Philosophy of Experimental Philosophy." In *A Compact Compendium of Experimental Philosophy* edited by Alexander Max Bauer and Stephan Kornmesser. Berlin: De Gruyter.
- Viswanathan, Madhu. 2005. *Measurement Error and Research Design*. Thousand Oaks, CA: SAGE.
- Weisberg, Jonathan. 2021. "Formal Epistemology." In *Stanford Encyclopedia of Philosophy* edited by Edward N. Zalta. Spring 2021 edition. <https://plato.stanford.edu/archives/spr2021/entries/formal-epistemology/>.