Is Philosophy Exceptional? A Corpus-Based, Quantitative Study

Moti Mizrahi
School of Arts and Communication, Florida Institute of Technology, Melbourne FL, USA
mmizrahi@fit.edu

Michael A, Dickinson
University of Illinois Urbana-Champaign, Urbana IL, USA

Abstract: Drawing on the epistemology of logic literature on anti-exceptionalism about logic, we set out to investigate the following metaphilosophical questions empirically: Is philosophy special? Are its methods (dis)continuous with science? More specifically, we test the following metaphilosophical hypotheses empirically: philosophical deductivism, philosophical inductivism, and philosophical abductivism. Using indicator words to classify arguments by type (namely, deductive, inductive, and abductive arguments), we searched through a large corpus of philosophical texts mined from the JSTOR database (n = 435,703) to find patterns of argumentation. The results of our quantitative, corpus-based study suggest that deductive arguments are significantly more common than abductive arguments and inductive arguments in philosophical texts overall, but they are gradually and steadily giving way to non-deductive (i.e., inductive and abductive) arguments in academic philosophy.

Keywords: abductive argument, deductive argument, exceptionalism, indicator words, inductive argument, metaphilosophy, philosophical abductivism, philosophical deductivism, philosophical inductivism, text analysis

1. Introduction

Anti-exceptionalism about logic (AEL) is the view that logic is not special. More explicitly, AEL consists of the following theses:

> Logic isn't special. Its theories are continuous with science; its method continuous with scientific method. Logic isn't a priori, nor are its truths analytic truths. Logical theories are revisable, and if they are revised, they are revised on the same grounds as scientific theories (Hjortland 2017, p. 632).

In terms of methodology in particular, which is the focus of the present study, AEL is the view that "Theory choice within logic is similar in important respects to that of the recognised sciences" (Martin and Hjortland 2021, p. 286). According to Martin and Hjortland (2021, p. 286), anti-exceptionalists about logic reject the claim that "logic is *epistemically foundational*, and thus that logical propositions are known immediately through non-inferential means," whereas exceptionalists embrace it (emphasis in original). Accordingly, foundationalism about logic is the view that "At least some logical propositions are known by non-inferential means" (Martin 2021, p. 9071).

Now, substituting logic for philosophy, specifically academic philosophy as it is practiced in colleges and universities across the Anglophone world, one could ask similar

questions about academic philosophy. Is academic philosophy special? Are its theories and methods continuous with science? Are philosophical theories revised on the same grounds as scientific theories? In his *The Philosophy of Philosophy* (2007), Williamson advances a sort of anti-exceptionalism about philosophy.[1] Williamson (2007, p. 3) acknowledges that "there are real methodological differences between philosophy and other sciences, as actually practiced," but he insists that "they are less deep than is often supposed." More specifically, he argues that the intuitions that academic philosophers appeal to are no different in kind from, and are continuous with, the sort of judgments that scientists ordinarily make. Any methodological difference between philosophical and non-philosophical inquiry, then, is merely a difference in degree of systematicity, not a difference in kind. As Williamson puts it, "the methodology of much past and present philosophy consists in just the unusually systematic and unrelenting applications of ways of thinking required over a vast range of non-philosophical inquiry."

Now, some philosophers have begun to address these metaphilosophical questions empirically. In one empirical study, Knobe (2015) compared two samples of published papers on the philosophical study of mind: one sample of papers from 1960 to 1999 and another sample of papers from 2009 to 2013. Knobe (2015) found that 62% of the papers from the 1960-1999 sample used purely *a priori* methods, whereas only 12% of the papers from the 2009-2013 sample used purely *a priori* methods. This evidence leads Knobe (2015, p. 38) to conclude that there has been "a strong shift [in method] toward the use of systematic empirical data, including original experiments conducted by philosophers."[2]

In another empirical study, Ashton and Mizrahi (2018a) test the view that philosophy is essentially an *a priori* discipline empirically. According to Ashton and Mizrahi (2018a, p. 62), "if philosophy is indeed *a priori*, and in the business of discovering necessary truths from the armchair, we would expect philosophers to advance mostly deductive, not inductive, arguments." Consistent with the view that philosophy is an *a priori* discipline, Ashton and Mizrahi (2018a) find that the proportion of philosophy articles in which deductive arguments are made is higher than that of philosophy articles in which inductive arguments are made. However, contrary to the view that philosophy is an *a priori* discipline, Ashton and Mizrahi (2018a) also find that the proportions of philosophy articles in which deductive arguments are made and those in which inductive arguments are made are converging over time and that the difference between the ratios of inductive arguments and deductive arguments is declining over time. As Ashton and Mizrahi (2018a, pp. 68-69) put it, their results suggest that "deductive arguments are gradually losing their status as the dominant form of argumentation in philosophy."[3]

---

[1] Thanks to an anonymous reviewer for this point.

[2] It is important to keep in mind the distinction between methods of theory-choice or types of argument, on the one hand, and sources of evidence, on the other hand. For one could accept the empirical evidence pointing to a shift from the traditional methods of academic philosophy toward empirical methods, and yet insist that the sources of evidence that academic philosophers use are still the traditional sources of intuition, introspection, and the like. Similarly, as both Martin (2021) and Martin and Hjortland (2021) observe, one can reject the foundational status of logical knowledge, and thereby embrace an abductivist or a predictivist picture of theory-choice in logic, while at the same time insist that the "data" logicians use are still *a priori*. This issue is beyond the scope of this paper. For present purposes, our focus is on methods of theory-choice or styles of argument rather than sources of evidence in academic philosophy. But again we are grateful to an anonymous referee for raising this important point.

[3] See also Fletcher et al. (2021) for empirical evidence pointing to a turn away from formal, logical methods toward probabilistic methods in academic philosophy.

In this paper, we aim to contribute to this growing body of empirical work on whether academic philosophy is special and/or methodologically (dis)continuous with science. We adopted the methodology used by Ashton and Mizrahi (2018a), but we scaled it up significantly to encompass more data from philosophical texts. In addition, we also scaled out the methodology used by Ashton and Mizrahi (2018a) by including in our empirical study a type of argument that was left out of Ashton and Mizrahi's (2018a) empirical study, namely, abductive arguments. For some anti-exceptionalists about logic also subscribe to the view that "theories of logic, not unlike scientific theories in general, are chosen on the basis of *abductive arguments*, that is, inference to the best explanation" (Hjortland 2017, p. 632; emphasis added). This view is known as logical abductivism (cf. Martin and Hjortland 2021, p. 286). Applied to academic philosophy, then, *philosophical abductivism* is the view that theories of academic philosophy, much like scientific theories, are chosen on the basis of abductive arguments.

A related view can be dubbed *philosophical inductivism*. This is the view that theories of academic philosophy are chosen on the basis of inductive arguments. Now, if induction is as prevalent in science as abduction is taken to be,[4] then philosophical inductivism and philosophical abductivism are both consistent with anti-exceptionalism about philosophy.[5] For, as we have seen, one of the main tenets of AEL is that logical theories and methods are continuous with science. By the same token, according to anti-exceptionalism about academic philosophy, philosophical theories and methods are continuous with science. So, if both abduction and induction are integral parts of scientific method(s), which is what philosophers of science generally take to be the case,[6] then both philosophical inductivism and philosophical abductivism are consistent with anti-exceptionalism about academic philosophy.[7]

Contrary to philosophical abductivism and philosophical inductivism, the view according to which academic philosophy is special and discontinuous with scientific method(s) can be dubbed *philosophical deductivism*. This view about philosophy is the counterpart of logical rationalism in the epistemology of logic. As Martin (2021, p. 9077) puts it:

> According to logical rationalism, we gain justification for our logical beliefs directly from intuitions regarding a particular proposition. We simply *see* that the proposition *p* is true or false. Consequently, given logical rationalism, we would expect logical arguments to be full of appeals to intuitions, especially when it comes to certain fundamental propositions on which the remainder of one's logical theory rests (emphasis in original).[8]

---

[4] According to McCain and Poston (2017, p. 1), "Not only are rigorous inferences to the best explanation (IBE) used pervasively in the sciences, explanatory reasoning is virtually ubiquitous in everyday life."

[5] According to Okasha (2016, p. 19), "Most philosophers think it's obvious that science relies heavily on induction, indeed so obvious that it hardly needs arguing for."

[6] According to Chakravartty (2017), abduction "seems ubiquitous in science." See also Douven (2017) on the ubiquity of abduction. On the ubiquity of induction, see Henderson (2020): "We appear to rely on inductive inference ubiquitously in daily life, and it is also generally thought that it is at the very foundation of the scientific method."

[7] Cf. Martin and Hjortland (2021) on logical predictivism.

[8] In another empirical study using the methods of data mining and text analysis, Ashton and Mizrahi (2018b) test the "received wisdom" about intuition talk and appeals to intuition in academic philosophy. The results of their empirical study show that, contrary to the "received wisdom," both intuition talk (as indicated by words such as

The sort of inferences one can then make from intuitive or self-evident logical truths are supposed to be deductive (valid) inferences. Likewise, the sort of inferences one can make from intuitive or self-evident philosophical truths are supposed to be deductive (valid) inferences as well. Indeed, according to Chudnoff (2007, p. 29), who is a leading proponent of the evidential use of intuitions in academic philosophy, i.e., appeals to intuition, "Philosophers prize deductive arguments over all others."

We set out to test these metaphilosophical hypotheses, namely, philosophical abductivism, philosophical inductivism, and philosophical deductivism, empirically. According to Martin (2021, p. 9078), "if logical abductivism were correct, we would expect logical arguments to be full of appeals to well-recognised data which it is commonly accepted within the community that theories ought to be able to accommodate." Likewise, if philosophical abductivism were true, we would expect to find significantly more abductive arguments than deductive arguments or inductive arguments in philosophical texts. If philosophical inductivism were true, we would expect to find significantly more inductive arguments than deductive arguments or abductive arguments in philosophical texts. And if philosophical deductivism were true, we would expect to find significantly more deductive arguments than inductive arguments or abductive arguments in philosophical texts. Using data mining and text analysis methods, we study a large corpus of philosophical texts mined from the JSTOR database (n = 435,703) in order to test these hypotheses empirically.[9] For the purposes of this study, any article published in an academic journal of philosophy counts as a philosophical text. Using indicator words to classify arguments by type (namely, deductive, inductive, and abductive arguments), we searched through our corpus to find patterns of argumentation.[10]

Before we report the results of our quantitative, corpus-based study in Section 3, we describe our methodology in more detail in Section 2. (See also Appendix I for a detailed discussion of our text mining methods in R.) In Section 4, we will discuss the results of our empirical study. The results of our quantitative, corpus-based study suggest that deductive arguments are significantly more common than inductive arguments and abductive arguments in philosophical texts overall, but they are gradually and steadily giving way to non-deductive (i.e., inductive and abductive) arguments in academic philosophy.

2. Methods

Introductory textbooks to logic and argumentation typically contain a brief discussion of indicator words. There are premise indicators—words such as 'because' and phrases such as 'infer from' and the like—which indicate a premise of an argument, and there are conclusion indicators—words such as 'therefore' and phrases such as 'it follows that' and the like—which

'intuit' and 'intuitive') and appeals to intuition (as indicated by phrases such 'it seems that' and 'it appears that') go as far back as the 1800s.

[9] Ashton and Mizrahi (2018a) use data mined from JSTOR as well. Other philosophers have used JSTOR Data for Research to model topics in philosophy journals. See, for example, Weatherson' *A History of Philosophy Journals: Volume 1 Evidence from Topic Modeling, 1876-2013* (http://www-personal.umich.edu/~weath/lda/).

[10] For an example of the application of corpus-based methods to philosophy of logic, see Mizrahi (2019). Mizrahi (2019, p. 203) uses corpus-based methods "to test empirically how the idea that 'logic is obvious' is reflected in logical and philosophical practice."

indicate a conclusion of an argument. For example, Morrow and Weston (2011, p. 5) tell students to look for indicator words in order to distinguish between premises and conclusions. According to Morrow and Weston (2011, p. 5):

> Some words or phrases are *conclusion indicators*. These are words or phrases that tell you that you're about to read or hear the conclusion of an argument. Other words or phrases are *premise indicators*. These tell you that you're about to read or hear a premise (emphasis in original).

Morrow and Weston (2011, p. 5) then provide a list of premise indicators, which includes words like 'because' and 'since', and a list of conclusion indicators, which includes words like 'therefore' and 'hence'. Likewise, according to Govier (2013, p. 4), "Indicator words suggest the presence of argument and help to indicate its structure. Some indicator words, like *therefore*, come before the conclusion in an argument. Other indicator words, like *since* and *because*, come before premises." Govier's (2013, pp. 4-5) list of premise indicators include the following: 'since', 'because', 'for', 'as indicated by', 'follows from', 'may be inferred from', 'may be derived from', 'on the grounds that', 'for the reason that', 'as shown by', 'given that', and 'may be deduced from'. And her list of conclusion indicators includes the following: 'therefore', 'thus', 'so', 'consequently', 'hence', 'then', 'it follows that', 'it can be inferred that', 'in conclusion', 'accordingly', 'for this reason (or for all these reasons) we can see that', 'on these grounds it is clear that', 'proves that', 'shows that', 'indicates that', 'we can conclude that', 'we can infer that', and 'demonstrates that' (Govier 2013, pp. 5-6).

In addition to helping students identify the premises and conclusions of arguments, indicators also help students distinguish between different types of arguments, such as deductive arguments and inductive arguments. For example, according to Baronett (2016, p. 23):

> To help identify arguments as either deductive or inductive, one thing we can do is look for key words or phrases. For example, the words "necessarily," 'certainly," "definitely," and "absolutely" suggest a deductive argument. [...] On the other hand, the words "probably," "likely," "unlikely," "improbable," "plausible," and "implausible" suggest inductive arguments.

Similarly, according to Hurley and Watson (2018, p. 34), "In deciding whether an argument is inductive or deductive, we look to certain objective features of the argument" (2018, pp. 34-35). One of those objective features is "the occurrence of special indicator words" (Hurley and Watson 2018, pp. 34-35). According to Hurley and Watson (2018, p. 35), "inductive indicators" include words and phrases such as 'probably', 'improbable', 'plausible', 'implausible', 'likely', 'unlikely', and 'reasonable to conclude', whereas "deductive indicators" include words and phrases such as 'it necessarily follows that', 'certainly', 'absolutely', and 'definitely'.[11]

[11] According to Salmon (2013), "Expressions such as *must, it must be the case that, necessarily, inevitably, certainly,* and *it can be deduced that* frequently indicate that an argument is deductive," (p. 86), whereas expressions such as "*probably, usually, tends to support, likely, very likely,* and *almost always*" typically indicate that an argument is inductive (p. 94).

We can use these deductive indicators and inductive indicators, then, to look for deductive arguments and inductive arguments in philosophical texts in much the same way that students of logic and philosophy use them to look for arguments in any text. In that respect, we are following Ashton and Mizrahi's (2018a) methodology, but with a novel addition. That is, to the aforementioned deductive and inductive indicator words, we have added indicator words for abductive arguments, i.e., arguments in which the conclusion is supposed to be the best explanation for some phenomenon (Govier 2013, pp. 298-302). Abductive indicators include words and phrases such as 'account for', 'best explain', 'make sense of', and 'best explanation for' (Overton 2013). Accordingly, the types of arguments we searched for in this empirical study and their associated indicators are listed in Table 1.

Table 1. Types of arguments and their indicator words with examples from philosophical texts

Argument Types	**Indicators**	**Examples**
Abductive	account for, best explain, makes sense of, best explanation for	"We infer that middle-sized objects exist, because their existence *provides the best explanation for* the patterns in our sense experience" (Trout 1998, p. 97).
Deductive	absolutely, certainly, definitely, necessarily	"if, as he says, such an infinite series really is impossible then it does *absolutely follow that* if anything exists in time at all, there must have been a moment, before which nothing existed" (Moore 1954, p. 175).
Inductive	likely, unlikely, probably, improbable	"Whatever may be the case for lesser breeds without the law, the nature of 'open' and—*sotto voce*—Western societies is such that conspiracy theories involving Western governments are unlikely to be true, and *hence unlikely* to be justified" (Pigden 2017, p. 123).

Of course, we have to keep in mind that these abductive, deductive, and inductive indicators are just that--*indicators*. They are not sure signs of the presence (or absence) of arguments in texts. In other words, "the mere occurrence of an indicator word by no means guarantees the presence of an argument" (Hurley and Watson 2018, p. 16). Nevertheless, they are still useful and reliable indicators for the presence of arguments in text, which is why students of logic and philosophy are instructed to look for them. As Lepore and Cumming (2013, p. 6) put it, "Although there are no sure signs of whether an argument is present, fairly reliable indicators exist." Lepore and Cumming (2013, p. 6) proceed to list some of the aforementioned indicator words as those listed in Table 1. In addition, since our aim is to study arguments made by academic philosophers, which are published in academic journals of philosophy, and academic "philosophers are careful folk, trained in the ways of argument" (Currie 2016, p. 200), we can be quite confident that, as professionals, academic philosophers rarely misuse indicators in an effort to make non-arguments appear as arguments (see also Ashton and Mizrahi 2018a, p. 62).

The quantitative, corpus-based methods we use in this empirical study allow us to overcome the limitations of relying on selective quotation. After all, one can easily find instances of the aforementioned indicator words in philosophical texts (see Table 1). However, selected quotations may or may not be representative of academic philosophy as a whole. By using text mining and analysis methods, we can study a large corpus of philosophical texts, and thus obtain a broader view of the argumentative landscape in academic philosophy. Of course, empirical methodologies have limitations of their own. As far as our corpus-based methods are concerned, there are two major limitations. First, we can only study and analyze what is explicitly mentioned in the corpus. For the purpose of this empirical study, then, our corpus of philosophical texts must contain explicit mentions of the indicator words listed in Table 1, so that we could analyze ratios, means, and patterns of usage. It is reasonable to assume that there would be such explicit mentions of the indicator words listed in Table 1 in philosophical texts if academic philosophers are indeed professional arguers; that is, "trained in the ways of argument" (Currie 2016, p. 200).

Second, as with any empirical methodology, there may be some false positives and/or false negatives. When it comes to the corpus-based methods used in this study, false negatives could occur when we search for a specific word *w* in a corpus, but do not find it, even though the corpus contains a synonym of *w*. For example, although unlikely, it is possible that our corpus of philosophical texts contains no instances of 'probably', and so a search for 'probably' would return zero results, because academic philosophers use 'likely' instead of 'probably' in all the philosophical texts that make up our corpus. On the other hand, false positives could occur when we find instances of a word *w* in our corpus, but those instances contain irrelevant uses of *w*. For the purpose of this empirical study, then, the corpus of philosophical texts must contain not only explicit mentions of the abductive, deductive, and inductive indicators listed in Table 1, but also explicit mentions of those indicators in the context of argumentation. For example, instances of 'certainly' that occur outside of any argumentative context would be considered false positives for the purposes of this study.

Now, there are a couple of things we can do to overcome the limitations of our quantitative, corpus-based approach. First, we can refine our searches by expanding our search terms to include as many indicator words as we can. For each argument type, we have four indicator words (see Table 1). This search methodology is designed to minimize the number of false negatives, i.e., occurrences of abductive, deductive, and inductive arguments in philosophical texts that are indicated by words other than the standard ones, such as 'best explain', 'necessarily', and 'probably', by using synonymous indicator words and phrases, such as 'account for', 'certainly', and 'likely'.[12]

In that respect, it is important to note another possibility for false positives in this study. Take the indicator word 'likely' again. Following the logic textbooks cited above, we take the word 'likely' to be a reliable indicator for inductive arguments. However, as an anonymous reviewer rightly points out, the word 'likely' could also occur in the context of abductive argument. The same can be said about the word 'probably'. Indeed, broadly speaking, abductive

[12] In Hylnad's (2005) taxonomy of metadiscourse signals, 'probably' and 'likely' are classified as hedges, whereas in Salagar-Meyer's (1994) taxonomy they are classified as shields. For a critical assessment of these taxonomies, see Thabet (2018).

arguments may be considered inductive arguments insofar as the premises of an abductive argument are intended to make its conclusion *probably*, but not *necessarily*, true. Accordingly, if "[a]n *inductive argument* is one in which it is claimed that the premises make the conclusion *probable*" (Baronett 2016, p. 23; emphasis in original), and the premises of abductive arguments are intended to provide *probable* support for their conclusions, then abductive arguments can be considered a type of inductive arguments. Nevertheless, some philosophers and logicians treat abductive arguments as a distinct type of argumentation. Indeed, Baronett (2016) himself discusses abduction and Inference to the Best Explanation (IBE) in a chapter titled "Causality and Scientific Arguments," which is separate from the chapters on deduction and induction in his logic textbook. According to Baronett (2016, p. 652), "In *inference to the best explanation*, we reason from the premise that a hypothesis would explain certain facts to the conclusion that the hypothesis is the best explanation for those facts" (emphasis in original). For this reason, we follow these logic textbooks in treating abductive arguments as a distinct type of argument, which is different from inductive arguments, and is identified by indicator words and phrases such as 'account for', 'best explain', 'make sense of', and 'best explanation for' (Overton 2013).

Second, we can further refine our searches by pairing the argument type indicators with indicator words for arguments, such as 'therefore' and 'hence'. Since our aim is to test the aforementioned metaphilosophical hypotheses, namely, philosophical abductivism, philosophical inductivism, and philosophical deductivism, empirically, we need to find out what types of arguments academic philosophers actually make in philosophical publications. To this end, we need to search for the abductive, deductive, and inductive indicators listed in Table 1 in argumentative contexts by pairing the abductive, deductive, and inductive indicators listed in Table 1 with indicator words for arguments, such as 'therefore' and 'hence'. Although they are frequently mentioned as premise indicators, we chose not to use the words 'since' and 'because'. For, as Copi et al. (2011, p. 18) point out, "those words are used both in explanations and in arguments." Instead, words like 'therefore' and 'hence' tend to indicate arguments rather than explanations more reliably. By anchoring the abductive, deductive, and inductive indicators listed in Table 1 to argument indicators, such as 'therefore' and 'hence', we can be quite confident that our indicators for argument types (see Table 1) actually indicate arguments in the corpus, given that an argument must have a conclusion, and thus that the number of false positives will be minimized. This procedure results in the argument indicator pairs listed in Table 2, which are the indicator words most commonly mentioned in logic textbooks.

Table 2. Indicator pairs for deductive, inductive, and abductive arguments

Deductive indicator pairs	**Inductive indicator pairs**	**Abductive indicator pairs**
therefore necessarily	therefore probably	therefore account for
therefore certainly	therefore likely	therefore best explain
therefore definitely	therefore unlikely	therefore make sense of
therefore absolutely	therefore improbable	therefore best explanation for
hence necessarily	hence probably	hence account for
hence certainly	hence likely	hence best explain
hence definitely	hence unlikely	hence make sense of
hence absolutely	hence improbable	hence best explanation for
so necessarily	so probably	so account for

so certainly	so likely	so best explain
so definitely	so unlikely	so make sense of
so absolutely	so improbable	so best explanation for
consequently necessarily	consequently probably	consequently account for
consequently certainly	consequently likely	consequently best explain
consequently definitely	consequently unlikely	consequently make sense of
consequently absolutely	consequently improbable	consequently best explanation for
proves necessarily	proves probably	proves account for
proves certainly	proves likely	proves best explain
proves definitely	proves unlikely	proves make sense of
proves absolutely	proves improbable	proves best explanation for
thus necessarily	thus probably	thus account for
thus certainly	thus likely	thus best explain
thus definitely	thus unlikely	thus make sense of
thus absolutely	thus improbable	thus best explanation for
follows necessarily	follows probably	follows account for
follows certainly	follows likely	follows best explain
follows definitely	follows unlikely	follows make sense of
follows absolutely	follows improbable	follows best explanation for
accordingly necessarily	accordingly probably	accordingly account for
accordingly certainly	accordingly likely	accordingly best explain
accordingly definitely	accordingly unlikely	accordingly make sense of
accordingly absolutely	accordingly improbable	accordingly best explanation for
infer necessarily	infer probably	infer account for
infer certainly	infer likely	infer best explain
infer definitely	infer unlikely	infer make sense of
infer absolutely	infer improbable	infer best explanation for

These indicator words were selected because they are the ones that appear most frequently in various logic textbooks, as we discussed above.[13]

By searching for these argument indicator pairs (as listed in Table 2) in our corpus, we can find out what types of arguments academic philosophers make in their published works and with what frequency. For each of the pairs listed in Table 2, we ran three kinds of searches: (a) a search allowing for up to three words between argument type indicator, e.g., ‘necessarily’, and argument indicator, e.g., ‘therefore’, (b) a search allowing for up to six words between argument type indicator, e.g., ‘probably’, and argument indicator, e.g., ‘hence’, and (c) a search allowing for up to ten words between argument type indicator, e.g., ‘account for’, and argument indicator, e.g., ‘so’. This search methodology allows us to test the aforementioned metaphilosophical hypotheses while minimizing the number of false positives and false negatives as follows:

[13] Thanks to an anonymous reviewer for pressing on this point. As Flake and Fried observe, “When information about the measures in a study is lacking, the information needed to evaluate the validity of the study is also lacking” (2020, p. 459).

1. Hypothesis 1: *Philosophical abductivism*: theories of academic philosophy are chosen on the basis of abductive arguments for the most part (i.e., more abductive than deductive or inductive arguments).
    1.1. Prediction 1.1: Philosophical texts contain significantly more abductive arguments than deductive arguments or inductive arguments.
2. Hypothesis 2: *Philosophical inductivism*: theories of academic philosophy are chosen on the basis of inductive arguments for the most part (i.e., more inductive than deductive or abductive arguments).
    2.1. Prediction 2.1: Philosophical texts contain significantly more inductive arguments than deductive arguments or abductive arguments.
3. Hypothesis 3: *Philosophical deductivism*: theories of academic philosophy are chosen on the basis of deductive arguments for the most part (i.e., more deductive than inductive or abductive arguments).
    3.1. Prediction 3.1: Philosophical texts contain significantly more deductive arguments than inductive arguments or abductive arguments.

Testing these hypotheses empirically would in turn get us a little closer to answering our research questions: Is academic philosophy special? Are its theories and methods continuous with science? Are philosophical theories revised on the same grounds as scientific theories?

It is important to emphasize, as an anonymous reviewer urged us to do, that these hypotheses are not to be read as universal generalizations. That is, Hypothesis 1 is not to be construed as the claim that all theories of academic philosophy, without exception, are chosen on the basis of abductive arguments. Similarly, Hypothesis 2 is not to be construed as the claim that all theories of academic philosophy, without exception, are chosen on the basis of inductive arguments. Likewise, Hypothesis 3 is not to be construed as the claim that all theories of academic philosophy, without exception, are chosen on the basis of deductive arguments. Instead, these hypotheses should be read as statistical generalizations that can be subjected to empirical and statistical testing. Accordingly, on Hypothesis 1, theories of academic philosophy are chosen on the basis of abductive arguments more often than not. In other words, if we were to pick a theory of academic philosophy at random, that theory would be more likely than not to have been chosen on the basis of abductive arguments. Similarly, on Hypothesis 2, theories of academic philosophy are chosen on the basis of inductive arguments more often than not. In other words, if we were to pick a theory of academic philosophy at random, that theory would be more likely than not to have been chosen on the basis of inductive arguments. Finally, on Hypothesis 3, theories of academic philosophy are chosen on the basis of deductive arguments more often than not. In other words, if we were to pick a theory of academic philosophy at random, that theory would be more likely than not to have been chosen on the basis of deductive arguments.

It would be ideal to have data from academic disciplines other than academic philosophy that we could then compare to our data from academic philosophy. That way, we could see whether argumentation in academic philosophy is different from argumentation in other disciplines. Since our corpus is made up of philosophical texts only, however, we cannot do that. So, we have to leave this work to future studies. For the purposes of this study, we are not testing any comparative hypotheses about academic philosophy in relation to other disciplines. Recall

that "the central exceptionalist claim is that the justification of logical theories is a priori" (Hjortland 2017, p. 633), whereas anti-exceptionalism about logic (AEL) is the view that "Logic isn't a priori, nor are its truths analytic truths" (Hjortland 2017, p. 632). Analogously, philosophical exceptionalism is the view that the justification of philosophical theories is *a priori*, whereas anti-exceptionalism about philosophy is the view that philosophy is not *a priori*. These claims are not comparative, and so no comparison to other disciplines is required, although future studies with comparisons to other disciplines would be a welcome addition to the growing body of empirical work on whether academic philosophy is special, *a priori*, and (dis)continuous with science. Furthermore, the hypotheses we set out to test in this study, namely, philosophical abductivism (Hypothesis 1), philosophical inductivism (Hypothesis 2), and philosophical deductivism (Hypothesis 3), are hypotheses about argumentation *within* philosophy.

It is also important to emphasize again that our search methodology is not totally immune from counting false negatives and/or false positives, as we discussed above. One reason to think that there might be some false negatives in our datasets is that academic philosophers could be omitting indicator words from their academic publications deliberately because they are writing for a professional audience of academic philosophers. Presumably, being academic philosophers themselves, readers of philosophy journals do not need indicator words to identify arguments in text. This is possible, of course, although omitting indicator words might seem to run counter to academic philosophers' professed commitment to rigor and clarity in philosophical writing. For omitting indicator words would make it less clear to any reader, academic philosopher or not, where the argument in the text is, what type of argument is being made, and what the premises and the conclusion of the argument are. But academic philosophers, particularly those working in the analytic tradition, "pride themselves on skill in argumentation" (Rorty 2006, p. 70) and clarity of thought. As Lackey (2005, p. 277) puts it, "Analytic philosophers pride themselves on being logical, rigorous, and clear."[14]

3. Results

When working with data for research from JSTOR, it is standard practice not to include data from the most recent years because the JSTOR database may not have a complete archive of the most recent publications. For this reason, our datasets include publications from the earliest year in the JSTOR database, i.e., 1867, up to the year 2014, after which JSTOR does not seem to have a complete archive of philosophical publications. Now, in searches permitting three words between argument indicator root and anchor, the mean ratio of deductive arguments is higher than the mean ratio of inductive arguments and the mean ratio of abductive arguments. However, the ratios appear to be converging over time, with more recent years showing slightly higher ratios of inductive arguments than deductive arguments in philosophical texts.

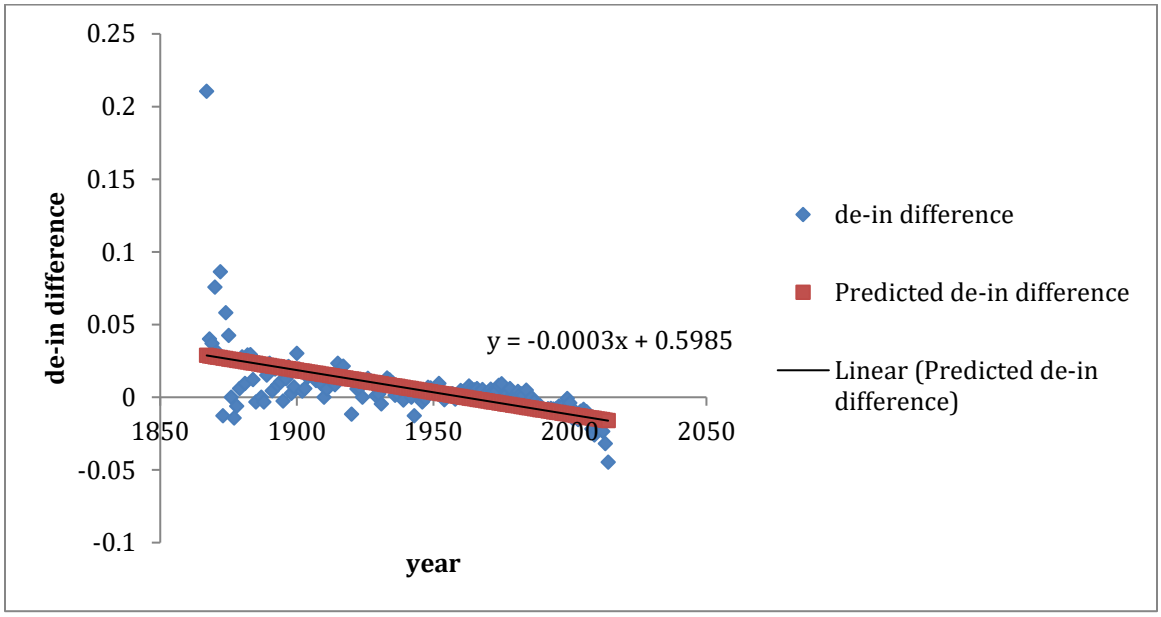
Welch's *t*-tests were conducted to compare the ratios of argument types from the results for searches allowing a three-word maximum range. First, there was a statistically significant difference between deductive arguments ($M = 0.0361$, $SD = 0.0208$, $N = 148$) and abductive arguments ($M = 0.0051$, $SD = 0.0046$, $N = 148$), $t(162) = -17.6104$, $p < .001$. Second, there was a statistically significant difference between inductive arguments ($M = 0.0298$, $SD = 0.0162$, $N =$

[14] See Appendix I for details on the text mining methods we used in R.

148) and abductive arguments (M = 0.0051, SD = 0.0046, N = 148), $t(171) = -17.7228$, $p < .001$. Finally, there was a statistically significant difference between deductive arguments (M = 0.0361, SD = 0.0208, N = 148) and inductive arguments (M = 0.0298, SD = 0.0162, N = 148), $t(277) = 2.9081$, $p = .003$, two-tailed. These results suggest that philosophical texts contain significantly more deductive arguments than either inductive arguments or abductive arguments.

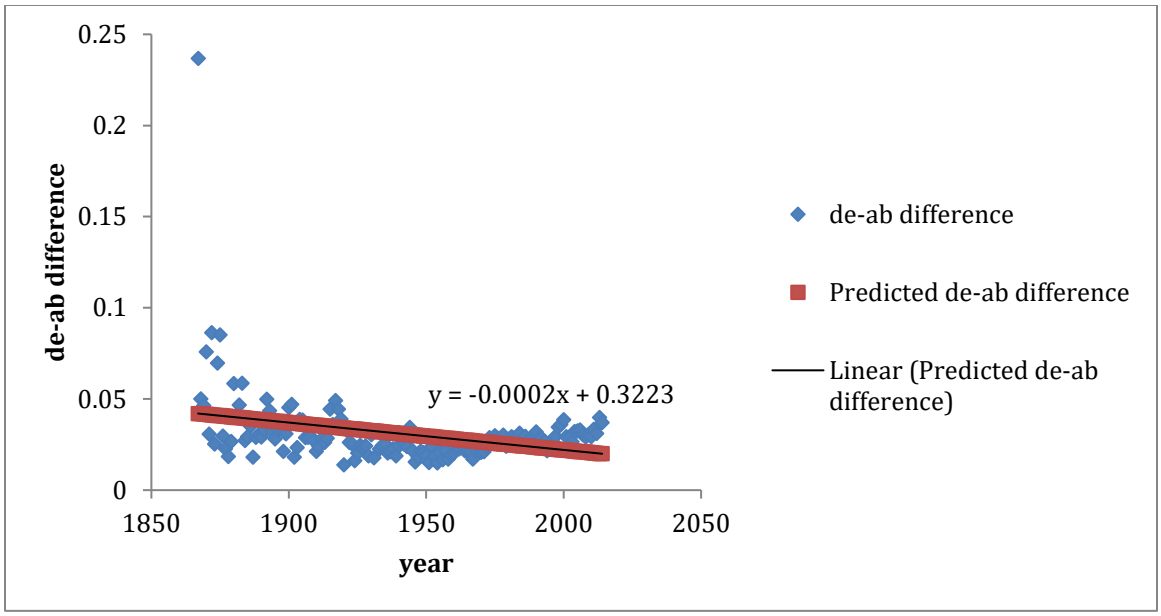
Following Ashton and Mizrahi (2018a), we also modeled the differences in ratios between the ratios of deductive arguments and the ratios of inductive arguments in philosophical texts over time. When we look at the ratios of deductive arguments and the ratios of inductive arguments over the years (1867-2014) in our three-word dataset, and run a regression analysis, the result is a linear model with the equation y = -0.0003x + 0.5985. Since the slope of the line is negative, the difference between the ratios of deductive arguments and inductive arguments is declining over the years, $R^2 = 0.31$, $F(1, 146) = 65.9686$, $p < .001$. (See Figure 1.)

Figure 1. Predicted differences in ratios between deductive and inductive arguments in the three-word dataset



Additionally, building up on Ashton and Mizrahi (2018a), we also modeled the differences in ratios between deductive arguments and abductive arguments in philosophical texts over time. When we look at the ratios of deductive arguments and the ratios of abductive arguments over the years (1867-2014), the equation for the line is y = -0.0002x + 0.3223. Since the slope of the line is negative, the difference between the ratios of deductive arguments and abductive arguments is also declining over the years, $R^2 = 0.10$, $F(1, 146) = 15.5432$, $p < .001$. (See Figure 2.)

Figure 2. Predicted differences in ratios between deductive and abductive arguments in the three-word dataset

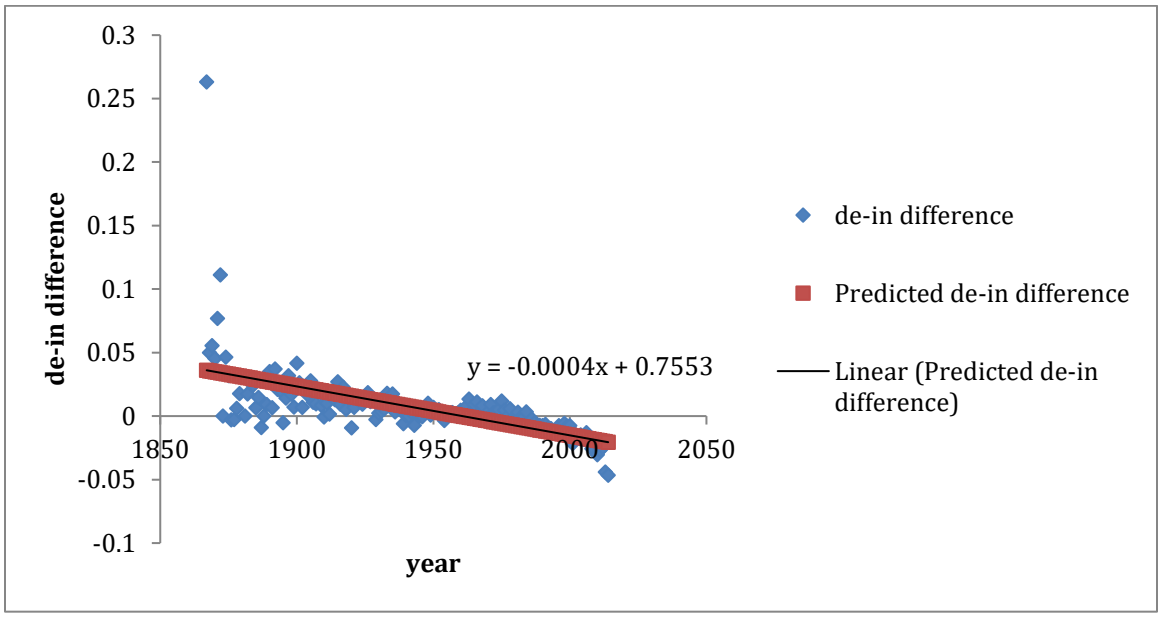


In searches permitting six words between argument indicator root and anchor, the mean ratio of deductive arguments is higher than the mean ratio of inductive arguments and the mean ratio of abductive arguments. However, the ratios appear to be converging over time, with more recent years showing slightly higher ratios of inductive arguments than deductive arguments in philosophical texts. This pattern is similar to the one exhibited by the data from our three-word searches.

Welch's *t*-tests were conducted to compare the ratios of argument types from the results for searches allowing a six-word maximum range. First, there was a statistically significant difference between deductive arguments (M = 0.0531, SD = 0.0252, N = 148) and abductive arguments (M = 0.0082, SD = 0.0068, N = 148), $t(169) = 20.8582$, $p < .001$. Second, there was a statistically significant difference between inductive arguments (M = 0.0454, SD = 0.0227, N = 148) and abductive arguments (M = 0.0082, SD = 0.0068, N = 148), $t(173) = 18.9993$, $p < .001$. Finally, there was a statistically significant difference between deductive arguments (M = 0.0531, SD = 0.0252, N = 148) and inductive arguments (M = 0.0454, SD = 0.0227, N = 148), $t(291) = 2.7732$, $p = .005$, two-tailed. Like the results from our three-word dataset, these results suggest that philosophical texts contain significantly more deductive arguments than either inductive arguments or abductive arguments.

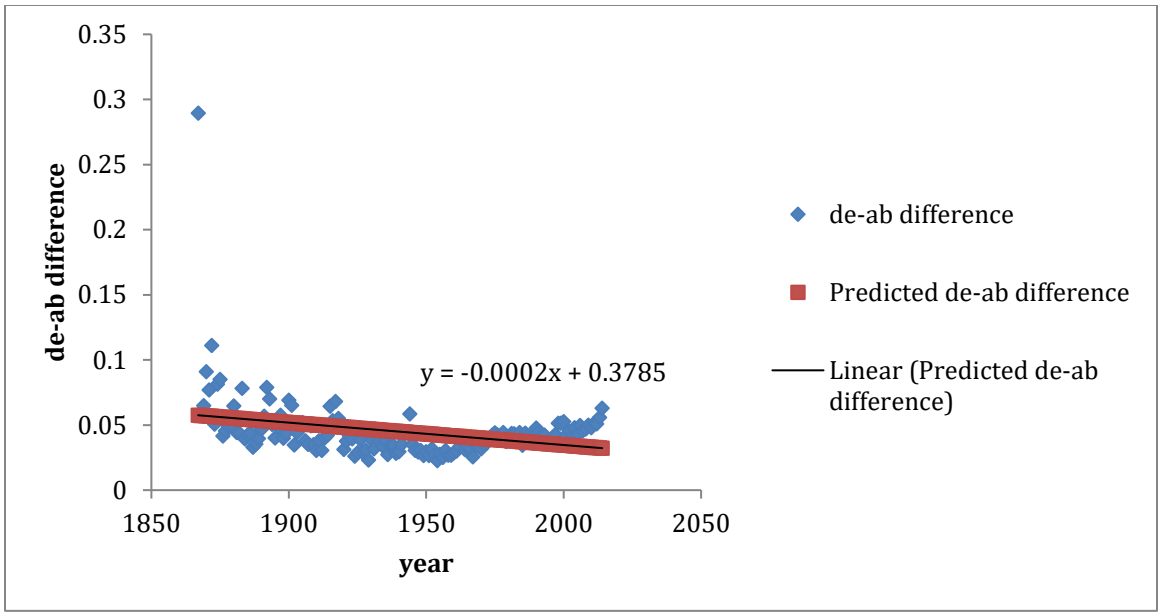
As we did for our three-word searches, and following Ashton and Mizrahi (2018a), we also modeled the differences in ratios between deductive arguments and inductive arguments in philosophical texts over time. When we look at the ratios of deductive arguments and the ratios of inductive arguments over the years (1867-2014) in our six-word dataset, and run a regression analysis, the result is a linear model with the equation y = -0.0004x + 0.7553. Since the slope of the line is negative, the difference between the ratios of deductive arguments and inductive arguments is declining over the years, $R^2 = 0.33$, $F(1, 146) = 73.6216$, $p < .001$. (See Figure 3.)

Figure 3. Predicted differences in ratios between deductive and inductive arguments in the six-word dataset



Additionally, building up on Ashton and Mizrahi (2018a), we also modeled the differences in ratios between deductive arguments and abductive arguments in philosophical texts over time, as we did for our three-word dataset. When we look at the ratios of deductive arguments and the ratios of abductive arguments over the years (1867-2014), the equation for the line is y = -0.0002x + 0.3785. Since the slope of the line is negative, the difference between the ratios of deductive arguments and abductive arguments is also declining over the years, R^2 = 0.08, $F(1, 146) = 14.3294$, $p < .001$. (See Figure 4.) These results are similar to the ones we have observed in our three-word dataset as well.

Figure 4. Predicted differences in ratios between deductive and abductive arguments in the six-word dataset

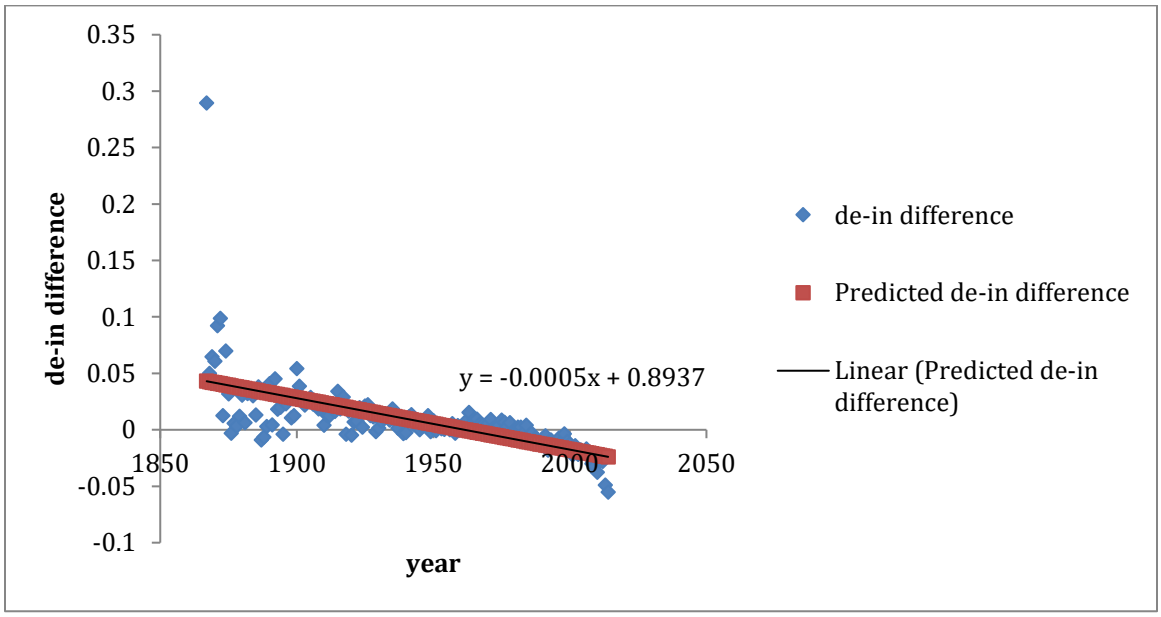


Finally, in searches permitting ten words between argument indicator root and anchor, the mean ratio of deductive arguments is higher than the mean ratio of inductive arguments and the mean ratio of abductive arguments. However, the ratios appear to be converging over time, with more recent years showing slightly higher ratios of inductive arguments than deductive arguments in philosophical texts. Again, this pattern is similar to the ones exhibited by the data from our three-word and six-word searches.

Welch's *t*-tests were conducted to compare the ratios of argument types from the results for searches allowing a ten-word maximum range. First, there was a statistically significant difference between deductive arguments (M = 0.0698, SD = 0.0279, N = 148) and abductive arguments (M = 0.0117, SD = 0.0093, N = 148), $t(179) = 23.9772$, $p < .001$. Second, there was a statistically significant difference between inductive arguments (M = 0.0603, SD = 0.0287, N = 148) and abductive arguments (M = 0.0117, SD = 0.0093, N = 148), $t(178) = 19.5432$, $p < .001$. Finally, there was a statistically significant difference between deductive arguments (M = 0.0698, SD = 0.0279, N = 148) and inductive arguments (M = 0.0603, SD = 0.0287, N = 148), $t(294) = 2.8927$, $p = .004$, two-tailed. Like the results from our three-word and six-word datasets, these results suggest that philosophical texts contain significantly more deductive arguments than either inductive arguments or abductive arguments.

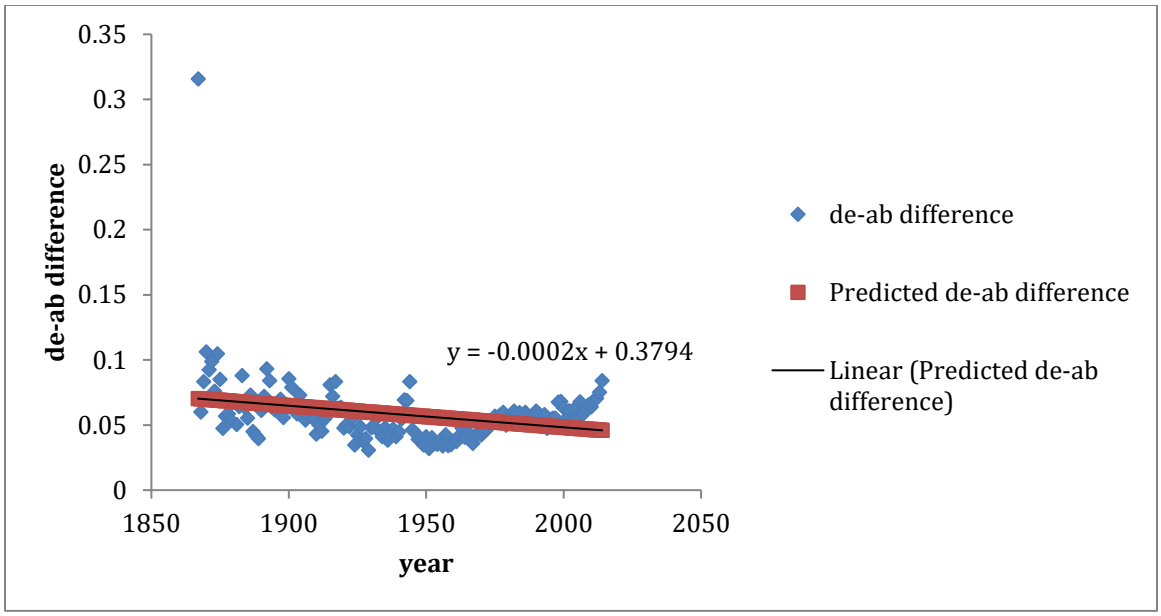
As we did for our three-word and six-word searches, and following Ashton and Mizrahi (2018a), we also modeled the differences in ratios between deductive arguments and inductive arguments in philosophical texts over time. When we look at the ratios of deductive arguments and the ratios of inductive arguments over the years (1867-2014) in our ten-word dataset, and run a regression analysis, the result is a linear model with the equation y = -0.0005x + 0.8937. Since the slope of the line is negative, the difference between the ratios of deductive arguments and inductive arguments is declining over the years, $R^2 = 0.37$, $F(1, 146) = 88.0541$, $p < .001$. (See Figure 5.)

Figure 5. Predicted differences in ratios between deductive and inductive arguments in the ten-word dataset



Additionally, building up on Ashton and Mizrahi (2018a), we also modeled the differences in ratios between deductive arguments and abductive arguments in philosophical texts over time, as we did for our three-word and six-word datasets. When we look at the ratios of deductive arguments and the ratios of abductive arguments over the years (1867-2014), the equation for the line is y = -0.0002x + 0.3794. Since the slope of the line is negative, the difference between the ratios of deductive arguments and abductive arguments is also declining over the years, $R^2 = 0.07$, $F(1, 146) = 11.5516$, $p < .001$. (See Figure 6.) These results are similar to the ones we have observed in our three-word and six-word datasets as well.

Figure 6. Predicted differences in ratios between deductive and abductive arguments in the ten-word dataset



4. Discussion

As we discussed in Section 1, our quantitative, corpus-based study was designed to test metaphilosophical hypotheses about argumentation in philosophical texts. According to philosophical abductivism (Hypothesis 1), philosophical theories are chosen mostly on the basis of abductive arguments. According to philosophical inductivism (Hypothesis 2), philosophical theories are chosen mostly on the basis of inductive arguments. And according to philosophical deductivism (Hypothesis 3), philosophical theories are chosen mostly on the basis of deductive arguments. Now, if philosophical abductivism were true, we would expect to find significantly more abductive arguments than deductive arguments or inductive arguments in philosophical texts. If philosophical inductivism were true, we would expect to find significantly more inductive arguments than deductive arguments or abductive arguments in philosophical texts. And if philosophical deductivism were true, we would expect to find significantly more deductive arguments than inductive arguments or abductive arguments in philosophical texts.

Our results suggest that philosophical publications contain all three types of arguments, namely, abductive, deductive, and inductive arguments, but in different proportions. The results of *t*-tests suggest that, on average, academic philosophers make significantly more deductive arguments than either inductive arguments or abductive arguments in their published works overall. Since we have observed these patterns in our three-word, six-word, and ten-word datasets, we can be quite confident that these results are robust.

While deductive arguments are significantly more common than inductive arguments and abductive arguments in philosophical texts overall, our results also suggest that deductive arguments are gradually and steadily giving way to inductive arguments and abductive arguments in academic philosophy. These findings are supported by linear models that show that the difference in deductive arguments in proportion to inductive arguments, and the difference in deductive arguments in proportion to abductive arguments, are gradually declining over the

years. Again, since we have observed these patterns in our three-word, six-word, and ten-word datasets, we can be quite confident that these results are robust.

Overall, then, the results of our study could be construed as providing some empirical support to philosophical deductivism (Hypothesis 3) over philosophical inductivism (Hypothesis 2) and philosophical abductivism (Hypothesis 1). For, if philosophical deductivism were true, we would expect to find significantly more deductive arguments than inductive arguments or abductive arguments in philosophical texts. Indeed, we have found that the ratio of deductive arguments is significantly higher than that of inductive arguments and that of abductive arguments in philosophical publications overall. It may be premature, however, to say that these results confirm philosophical deductivism (Hypothesis 3) insofar as there may be other alternative hypotheses (other than philosophical inductivism and philosophical abductivism) that have not been ruled out in this empirical study. We briefly discuss one such alternative hypothesis below.

Moreover, our results also suggest that we should expect inductive arguments and abductive arguments to become more prevalent in academic philosophy at the expense of deductive arguments. If these trends do continue as our linear models predict, then academic philosophy would become increasingly less exceptional in its use of deductive arguments, and more continuous with science in its growing use of non-deductive arguments, namely, inductive arguments and abductive arguments. For "inductive inference [...] is at the very foundation of the scientific method" (Henderson 2020)[15] and abductive inference is "ubiquitous in scientific practice" (Chakravartty 2017).[16] In other words, as inductive arguments and abductive arguments become more prevalent in academic philosophy, as our linear models predict, academic philosophy will become more continuous with science and, in turn, exceptionalism about philosophy will become less likely as a metaphilosophical view of academic philosophy, since both inductive arguments and abductive arguments are non-deductive types of arguments that are considered integral parts of scientific method(s).[17]

In that respect, our findings are in line with the results of other empirical studies on the status of philosophy as an *a priori* or empirical field of inquiry, particularly those conducted by Knobe (2015) and Ashton and Mizrahi (2018a), which are discussed in the introduction above (Section 1).[18] The interesting thing about our results, which differentiates them from the results obtained by Knobe (2015) and those obtained by Ashton and Mizrahi (2018a), is that they suggest that abductive arguments account for the steady and gradual move away from deductive arguments almost as much as inductive arguments do. In other words, deductive arguments seem to be giving way to not only inductive arguments but also abductive arguments in philosophical publications. These patterns were observed in our three-word, six-word, and ten-word datasets, which is why we can be quite confident that these results are robust. Moreover, since our results were obtained from a survey of a large corpus of philosophical texts mined from the JSTOR

[15] According to Okasha (2016, p. 19), "science relies heavily on induction."

[16] According to McMullin (1992), abduction is "the inference that makes science."

[17] For a corpus-based, empirical study on whether scientists use mostly deductive terms or inductive terms when they talk about testing hypotheses in scientific publications, see Mizrahi (2020).

[18] See also Fletcher et al. (2021) for empirical evidence for a turn away from formal, logical methods toward probabilistic methods in academic philosophy.

database (n = 435,703), we can be quite confident that they are representative of academic philosophy in general.

Nevertheless, it is important to note that the steady and gradual move away from deductive arguments to non-deductive (i.e., inductive and abductive) arguments in academic philosophy may simply be a reflection, not of a methodological change in the types of arguments academic philosophers make, but of a terminological change. That is, changes in parlance may occur over long periods of time, and so the changes in indicator words over the years may simply be a reflection of such a change in the terminological preferences of philosophers from different periods. This explanation of our results is plausible, of course, and further studies are needed in order to find out whether it is the correct explanation. The quantitative corpus-based methods we have used in this study do not seem to be suitable for deciding between alternative explanations of the results of our study, namely, whether the changes in indicator words over time are a reflection of methodological changes or terminological changes in academic philosophy. It may be that a more qualitative approach is needed in order to address this question.[19] For this reason, we leave this question to future work.

It is also important to note that the results of our quantitative, corpus-based study apply to academic philosophy as it is practiced in the English language only. This is because our data was mined from philosophy journals that publish articles written in English. It is possible that philosophical texts written in languages other than English might exhibit patterns of argumentation that are different from those we have found in our study. Accordingly, we think it would be interesting to conduct quantitative, corpus-based studies similar to the one we have conducted but with data mined from philosophical texts written in languages other than English. Likewise, as we mentioned in Section 2, for the purposes of this empirical study, we were only concerned with argument types, not arguers, in philosophical publications. In future studies, then, it would be interesting to find out if there are any significant differences between philosophical arguers and the types of arguments they make in philosophical texts.

5. Conclusion

In this paper, we have reported the results of our quantitative, corpus-based study of argument types in philosophical publications. Using indicator words to classify arguments by type (namely, deductive, inductive, and abductive arguments), we searched through a large corpus of philosophical texts mined from the JSTOR database (n = 435,703) in order to test the following metaphilosophical hypotheses: philosophical abductivism, philosophical deductivism, and philosophical inductivism. Our findings suggest that, while deductive arguments are significantly more common than inductive arguments and abductive arguments in philosophical texts overall, they are gradually and steadily giving way to inductive arguments and abductive arguments in academic philosophy. Our linear models show that inductive arguments and abductive arguments are expected to become more prevalent in academic philosophy at the expense of deductive arguments. Insofar as our results point to a steady and gradual change from deductive to non-deductive (i.e., inductive and abductive) types of arguments in academic philosophy, they suggest that what the future has in store for academic philosophy is less exceptionalism and more continuity with science.

[19] For an example of a more qualitative approach to corpus analysis, see Sytsma et al. (2019).

Acknowledgments

We are grateful to two anonymous reviewers of *Social Epistemology* for invaluable comments on earlier drafts of this paper.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Moti Mizrahi is Associate Professor of Philosophy at the Florida Institute of Technology. He is the author of *The Relativity of Theory: Key Positions and Arguments in the Contemporary Scientific Realism/Antirealism Debate* (Springer, 2020), the editor of *The Kuhnian Image of Science: Time for a Decisive Transformation?* (Rowman & Littlefield, 2018), and the editor of *For and Against Scientism: Science, Methodology, and the Future of Philosophy* (Rowman & Littlefield, 2022). He has published extensively on the philosophy of science, the scientific realism/anti-realism debate, the epistemology of philosophy, and argumentation. His work has appeared in journals such as *Argumentation*, *British Journal for the Philosophy of Science*, *Erkenntnis*, *Philosophical Studies*, *Studies in History and Philosophy of Science*, and *Synthese*. A major theme in his work has been the application of digital, statistical, and data-driven methods to problems concerning moral, philosophical, and scientific reasoning.

Mike Dickinson is the Planning, Landscape Architecture, & Agriculture Librarian at the University of Illinois. His librarianship has centered on data management. Mike's research interests include data education, large-scale library reference analysis, and corpus-based textual analysis of philosophical literature.

Data availability statement

Raw data sets for this study can be found on OSF at https://osf.io/6acpy/?view_only=6e0ce2ef55f546f48175f372464b156f.

References

Ashton, Z. and Mizrahi, M. (2018a). Show Me the Argument: Empirically Testing the Armchair Philosophy Picture. *Metaphilosophy* 49 (1-2): 58-70.

Ashton, Z. and Mizrahi, M. (2018b). Intuition Talk is not Methodologically Cheap: Empirically Testing the "Received Wisdom" about Armchair Philosophy. *Erkenntnis* 83 (3): 595-612.

Baronett, S. (2016). *Logic*. New York: Oxford University Press.

Chakravartty, A. (2017). Scientific Realism. In E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2017 Edition).

https://plato.stanford.edu/archives/sum2017/entries/scientific-realism/.

Chudnoff, E. (2007). *A Guide to Philosophical Writing*. President and Fellows of Harvard College. https://hwpi.harvard.edu/files/hwp/files/philosophical_writing.pdf.

Copi, I. M., Cohen, C., and McMahon, K. (2011). *Introduction to Logic*. Fourteenth Edition. Upper Saddle River, NJ: Prentice Hall.

Currie, G. (2016). Does Great Literature Make Us Better? In P. Catapano and S. Critchley (eds.), *The Stone Reader: Modern Philosophy in 133 Arguments* (pp. 198-202). New York: W. W. Norton & Co.

Douven, I. (2017). Abduction. In E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2017 Edition). https://plato.stanford.edu/archives/sum2017/entries/abduction/.

Flake, J. K., and Fried, E. I. (2020). Measurement Schmeasurement: Questionable Measurement Practices and How to Avoid Them. *Advances in Methods and Practices in Psychological Science* 3 (4): 456-465.

Fletcher, S. C., Knobe, J., Wheeler, G., and Woodcock, B. A. (2021). Changing Use of Formal Methods in Philosophy: Late 2000s vs. Late 2010s. *Synthese*. https://doi.org/10.1007/s11229-021-03433-9.

Govier, T. (2013). *A Practical Study of Argument*. Enhanced Seventh Edition. Belmont, CA: Wadsworth, Cengage Learning.

Henderson, L. (2020). The Problem of Induction. In E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2020 Edition). https://plato.stanford.edu/archives/spr2020/entries/induction-problem/.

Hjortland, O. T. (2017). Anti-exceptionalism about Logic. *Philosophical Studies* 174 (3): 631-658.

Hurley, P. J. and Watson, L. (2018). *A Concise Introduction to Logic*. Thirteenth Edition. Boston, MA: Cengage Learning.

Hyland, K. (2005). *Metadiscourse*. New York: Continuum.

Knobe, J. (2015). Philosophers are Doing Something Different Now: Quantitative Data. *Cognition* 135: 36-38.

Knobe, J., and Nichols, S. (2008). An Experimental Philosophy Manifesto.” In J. Knobe and S. Nichols (eds.), *Experimental Philosophy* (pp. 3-14). New York: Oxford University Press.

Lackey, M. (2005). Review of *A House Divided: Comparing Analytic and Continental Philosophy*. *The Journal of Speculative Philosophy* 19 (3): 176-280.

Lepore, E., and Cumming, S. (2013). *Meaning and Argument: An Introduction to Logic through Language*. Malden, MA: Blackwell.

Martin, B. (2021). Identifying Logical Evidence. *Synthese* 198 (10): 9069-9095.

Martin, B. and Hjortland, O. (2021). Logical Predictivism. *Journal of Philosophical Logic* 50 (2): 285-318.

McCain, K. and Poston, T. (2017). Best Explanations: An Introduction. In K. McCain and T. Poston (eds.), *Best Explanations: New Essays on Inference to the Best Explanation* (pp. 1-4). New York: Oxford University Press.

McMullin, E. (1992). *The Inference that Makes Science*. Milwaukee: Marquette University Press.

Mizrahi, M. (2019). What Isn't Obvious about 'Obvious': A Data-Driven Approach to Philosophy of Logic. In A. Aberdein and M. Inglis (eds.), *Advances in Experimental Philosophy of Logic and Mathematics* (pp. 201-224). London: Bloomsbury.

Mizrahi, M. (2020). Hypothesis Testing in Scientific Practice: An Empirical Study. *International Studies in the Philosophy of Science* 33 (1): 1-21.

Mizrahi, M., and Dickinson, M. (2022). Philosophical Reasoning about Science: A Quantitative, Digital Study. *Synthese* 200. https://doi.org/10.1007/s11229-022-03670-6.

Moore, G. E. (1954). *Some Main Problems of Philosophy.* London: Routledge.

Morrow, D. R. and Weston, A. (2011). *A Workbook for Arguments: A Complete Course in Critical Thinking*. Indianapolis: Hackett Publishing Co.

Okasha, S. (2016). *Philosophy of Science: A Very Short Introduction*. 2nd Ed. Oxford: Oxford University Press.

Overton, J. A. (2013). "Explain" in Scientific Discourse. *Synthese* 190 (8): 1383-1405.

Pigden, C. R. (2017). Are Conspiracy Theories Epistemically Vicious? In K. Lippert-Rasmussen, K. Brownlee, and D. Coady (eds.), *A Companion to Applied Philosophy* (pp. 120-132). Malden, MA: Wiley Blackwell.

Rorty, R. (2006). *Take Care of Freedom and Truth Will Take Care of Itself: Interviews with Richard Rorty*. Edited by Eduardo Mendieta. Stanford, CA: Stanford University Press.

Salager-Meyer, F. (1994). Hedges and Textual Communicative Function in Medical English Written Discourse. *English for Specific Purposes* 13 (2): 149-170.

Salmon, M. H. (2013). *Introduction to Logic and Critical Thinking*. Sixth Edition. Boston, MA: Wadsworth.

Sytsma, J., Bluhm, R, Willemsen, P., and Reuter, K. (2019). Causal Attributions and Corpus Analysis. In E. Fischer and M. Curtis (eds.), *Methodological Advances in Experimental Philosophy* (pp. 209-238). London: Bloomsbury.

Thabet, R. A. (2018). A Cross-Cultural Corpus Study of the Use of Hedging Markers and Dogmatism in Postgraduate Writing of Native and Non-native Speakers of English. In K. Shaalan, A. E. Hassanien, and F. Tolba (eds.), *Intelligent Natural Language Processing: Trends and Applications* (pp. 677-710). Gewerbestrasse: Springer.

Trout, J. D. (1998). *Measuring the Intentional World.* New York: Oxford University Press.

Weatherson, B. (2020). *A History of Philosophy Journals: Volume 1 Evidence from Topic Modeling, 1876-2013*. Bookdown 0.21: http://www-personal.umich.edu/~weath/lda/ .

Williamson, T. (2007). *The Philosophy of Philosophy*. Malden, MA: Blackwell Publishing.

Appendix I: Text Mining Methods in R

A combination of several text-mining packages in R language were used to manipulate the corpus of philosophical texts throughout this study. RStudio was used as an interactive-development environment to process the data. The corpus of documents included a .txt file containing the full-text of the philosophical work, and a corresponding .xml file to the full-text file, composed of the metadata information about each full-text file.

The *readtext* package was utilized to load the text files into the RStudio environment. The *readtext* function takes a folder path as an input parameter (i.e., > readtext("filepath"). The *readtext()* function will then load all files in the target folder into RStudio as a dataframe comprised of two columns.

The first column is titled "doc_id," which lists the file names as individual elements within a string vector. The second column is titled "text" and includes the full-text from each of the individual text files as a single character string. Thus, the result of the second column is a vector of character strings, with each string containing the full-text of an input text file. The .xml files were converted to .txt files from the Windows Command Prompt. The .txt metadata files were then also read into the R environment using the *readtext()* function.

To search for indicator pairs within the full-text documents, the *string_detect()* function from the *stringr* package was used in combination with a regular expression as a pattern search parameter. The argument indicator root and anchor were included within the regular expression to search for specific words.

The regular expression pattern allows for the root of the argument indicator pairs to both precede and follow the anchor word(s) within a certain range of words, exclusively. The function

was applied to the corpus across three word-ranges. The ranges selected permitted 3, 6, or 10 words between the argument indicator root and the anchor word(s). For example, to search for pattern matches across a range of 3 words, the regular expression returns a positive match in the following cases:

Root $word_1$ $word_2$ $word_3$ *Anchor* | OR | *Anchor* $word_1$ $word_2$ $word_3$ *Root*

Any pattern in which the argument indicator roots and anchors are separated by less than the maximum range (i.e., 3, 6, or 10) is also considered a positive match. For example, as applied within a 3-word maximum range, the following case would be considered a positive match:

Anchor $word_1$ $word_2$ *Root*

Applied in this manner, the *string_detect()* function will return a list of TRUE or FALSE logical values, where TRUE indicates the presence of the argument indicator and the anchor at least one time within each full-text string. FALSE indicates no pattern match within the text. The logical values were then converted to numeric data, with 1 replacing TRUE and 0 replacing FALSE. This detection process was repeated for each indicator pair of the deductive, inductive, and abductive lists and across all three word-ranges. The resulting lists were then summed, and the number of positive matches were recorded to a separate .csv file.

Separate .csv files containing matched full-text documents across each of the word-ranges were also generated from these lists. Publication years were extracted from the metadata text files using the *str_extract()* function from the *stringr* package. As with the indicator pair matches, a regular expression was used to isolate XML tags containing the publication years from the corresponding metadata files for each full-text item in the corpus. Ratios were then calculated from the number of documents containing positively matched indicator pairs for each year in proportion to the total number of documents across the entire corpus for the same year. Before these ratios could be calculated, duplicate rows were removed from the aggregated data. This was done to ensure the total number of matched documents for each word-range and argument type would not exceed the total number of publications in a given year, but instead list all the publications that matched the argument-type only once. A regression analysis was then performed on the resulting ratios.

It should be noted that the algorithm searches the entire corpus for each indicator pair, but can only match each indicator pair one time within a single article. For example, if article *x*, published in 1950, contains 2 indicator pairs, ‘therefore necessarily’ and ‘hence certainly’, the algorithm would return a count of 2. This would be the case even if one of the root-anchor indicator pairs repeated more than once throughout article *x*. So, if ‘therefore necessarily’ were to repeat 5 times and ‘hence certainly’ were to repeat 3 times in article *x*, the returned count would still be 2. With that said, a separate article published in the same year could contain the same indicator pairs. In that case, those pairs would be counted again. If article *y* was also published in 1950 and contains the same indicator pairs as article *x*, they would also be counted, and the algorithm would return a count of 4 for 1950.

It is also important to emphasize again that this search algorithm is not totally immune from counting false negatives and/or false positives, as we discussed in Section 2. As anonymous reviewers pointed out, one reason to think that there might be some false negative results in our datasets is that academic philosophers could be omitting indicator words from their academic publications deliberately because they are writing for a professional audience of academic philosophers. Presumably, being academic philosophers themselves, such an audience does not need indicator words to identify arguments in text. This is possible, of course, although omitting indicator words might seem to run counter to academic philosophers' professed commitment to rigor and clarity in philosophical writing. For omitting indicators words would make it less clear to any reader, academic philosopher or not, where the argument in the text is, what type of argument is being made, and what the premises and the conclusion of the argument are. And academic philosophers, particularly those working in the analytic tradition, "pride themselves on skill in argumentation" (Rorty 2006, p. 70). As Lackey (2005, p. 277) puts it, "Analytic philosophers pride themselves on being logical, rigorous, and clear."

The size of our corpus (n = 435,703) makes it rather difficult to extract a random sample of articles for human coders to look through, and identify arguments in those articles, in order to then compare what the coders find to what the pattern detection algorithm finds. We have conducted a related study, however, in which we use a similar pattern detection algorithm and run it on a much smaller corpus. We then asked three coders to identify arguments in a small sample of articles extracted from that corpus. We then compared the results from the three coders with the results from the pattern detection algorithm and checked for interrater reliability. We have found that there is substantial agreement between the three coders and the algorithm (see Mizrahi and Dickinson 2022).

Nevertheless, we did test the algorithm on small amounts of textual data. Once the algorithm could count the correct number of exact matches to the provided root-anchor pairings on small amounts of text, it was scaled up to run on larger sections of the corpus, and eventually on the entire corpus at once. Another limitation is that processing times for each indicator pair could range from 10-30 minutes within the RStudio application. RStudio could also potentially run out of RAM while loading in digital objects. However, repeated tests did result in the same number of matches per indicator-pair and across each word-range.

Finally, as far as our research questions are concerned, we need not worry about *arguers*, i.e., who is making the argument, only about the type of the argument made. In that respect, the algorithm would positively match arguments in text whether or not they are the arguments of the author of that text. As long as there is an argument present in a philosophical text, no matter who the arguer is, we would like our algorithm to count it, given that our research questions are about *types of arguments* made in philosophical publications.