

Show Me the Argument: Empirically Testing the "Armchair Philosophy" Picture

Zoe Ashton, Simon Fraser University

Moti Mizrahi, Florida Institute of Technology

Forthcoming in *Metaphilosophy*

Abstract: Many philosophers subscribe to the view that philosophy is a priori and in the business of discovering necessary truths from the armchair. In this paper, we set out to empirically test this picture of philosophy as “armchair philosophy.” If philosophy were indeed a priori, and in the business of discovering necessary truths from the armchair, then we would expect to see that reflected in philosophical practice. In particular, we would expect philosophers to advance mostly deductive, rather than inductive, arguments, which are the sort of arguments whose premises, if true, necessitate the truth of their conclusions. Our results show that the percentage of philosophy articles advancing deductive arguments is higher than the percentage of articles advancing inductive arguments, which is what we would expect from the vantage point of the “armchair philosophy” picture. However, our results also show that the percentages of articles advancing deductive arguments and those advancing inductive arguments are converging over time and that the difference between inductive and deductive ratios is declining over time. This trend suggests that deductive arguments are gradually losing their status as the dominant form of argumentation in philosophy.

Keywords: a priori; armchair philosophy; deductive argument; indicator words; inductive argument; necessary truth; philosophical methodology

1. Introduction

According to Restall (2009, p. 37), “Philosophers love a priori knowledge: we delight in truths that can be known from the comfort of our armchairs, without the need to venture out in the world for confirmation.” Many philosophers think that this is what philosophy is; that is, that “philosophy is a priori” (Williamson 2007, p. 1). According to this conception of philosophy, the “traditional methods of philosophy are armchair ones [hence, “armchair philosophy”]: they consist of thinking, without any special interaction with the world beyond the chair, such as measurement, observation or experiment would typically involve” (Williamson 2007, p. 1). In other words, philosophical inquiry “seeks to arrive at conceptually grounded necessary truths that are general in scope” (Henderson and Horgan 2013, p. 115).¹ Some go so far as to claim that philosophy not only *is* but also “should be regarded as [...] a priori” (Deutsch 2015, p. xx). Those who subscribe to this picture of philosophy as “armchair philosophy” draw a sharp distinction between philosophical methodology and the methodologies of the sciences (see, e.g., Bealer 1998, pp. 201-240). And some are even critical of empirical methodologies, such as experimental philosophy, and the like (see, e.g., Williamson 2016, pp. 22-36).²

In this paper, we set out to empirically test this picture of philosophy as “armchair philosophy” against actual philosophical practice.³ If philosophy is indeed a priori, where “a priori” means “derived from mere thinking, independent of sense experience,” to be

¹ See also Melnyk (2008, p. 267) on conceptual analysis as “a priori discovery of necessary truths” from the armchair.

² See also the essays collected in Horvath and Grundmann (2012).

³ Of course, there are different conceptions of what philosophy is or what its aim is. For a recent discussion, see Kriegel (2016).

distinguished from “a posteriori,” i.e., “derived from sense experience” (Gutting 2009, p. 11), and it is in the business of discovering necessary truths from the armchair, then we would expect to see that reflected in philosophical practice.

In the next section, we describe how we think the “armchair philosophy” picture, if it were an accurate picture of what philosophy is, would be reflected in philosophical practice. Then, in Section 3, we will report the results of our experimental study and discuss their implications in Section 4.

2. Methods

Most logic textbooks instruct students to look for indicator words when trying to identify arguments in context. For example, according to Copi et al. (2011, p. 11):

One useful method [for recognizing arguments] depends on the appearance of certain common indicators, certain words or phrases that typically serve to signal the appearance of an argument’s conclusion or of its premises.⁴

Copi et al. (2011, pp. 11-12) go on to provide the following list of conclusion indicator words and premise indicator words (Table 1).

⁴ See also Salmon (2013), p. 15.

Table 1. A list of conclusion indicators and premise indicators

Conclusion indicators	Premise indicators
therefore	since
hence	because
so	for
accordingly	follows from
consequently	as shown by
proves that	the reason is that
thus	may be derived from
it follows that	may be deduced from
we may infer	in view of the fact that

For instance, Moore’s use of the phrase ‘prove that’ in the following passage signals to the reader what the conclusion of the argument that Moore puts forth is (emphasis added):

I can *prove* now, for instance, that two human hands exist. How? By holding up my two hands, and saying, as I make a certain gesture with the right hand, ‘Here is one hand’, and adding, as I make a certain gesture with the left, ‘and here is another’. And if, by doing this, I have *proved ipso facto* the existence of external things, you will all see that I can also do it now in numbers of other ways: there is no need to multiply examples (Moore 1959, pp. 145-146).

The conclusion of Moore’s anti-skeptical argument, then, is that there is an external world, i.e., there are things that exist outside the mind, such as hands, papers, shoes, etc., as indicated by his use of the indicator word ‘prove’.

Indicator words are also supposed to help students distinguish between 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 (2016, p. 31), “inductive indicators” include terms and phrases such as “probably,” “improbable,” “plausible,” “implausible,” “likely,” “unlikely,” and “reasonable to conclude,” whereas “deductive indicators” include terms and phrases such as “it necessarily follows that,” “certainly,” “absolutely,” and “definitely.”

Table 2. A list of deductive and inductive indicator words

Deductive indicators	Inductive indicators
necessarily	(im)probably
certainly	(un)likely
definitely	(im)plausible
absolutely	(un)reasonable

For example, the presence of the conclusion indicator ‘so’ and the inductive indicator ‘plausible’ in the following passage suggests that the argument is supposed to be inductive (emphasis added):

It would be a miracle, a coincidence on a near cosmic scale, if a theory made as many correct empirical predictions as, say, the general theory of relativity or the photon theory of light without what that theory says about the fundamental structure of the universe being correct or “essentially” or “basically” correct. But we shouldn’t accept miracles, not at any rate if there is a non-miraculous alternative. If what these theories say is going on “behind” the phenomena is indeed true or “approximately true” then it is no wonder that they get the phenomena right. *So it is plausible to conclude that* presently accepted theories are indeed “essentially” correct (Worrall 1989, p. 101).

Indeed, this argument, which is known as the “no-miracles argument” for scientific realism, is commonly taken by philosophers of science to be an inference to the best explanation (Sankey 2016, pp. 22-23). Of course, whether the no-miracles argument is strong or cogent is another matter (see, e.g., Mizrahi 2012).⁵

Likewise, the presence of the premise indicator ‘since’ and the deductive indicator ‘necessarily’ in the following passage suggests that the argument is supposed to be deductive (emphasis added):

if human nature is so constituted as to desire nothing which is not either a part of happiness or a means of happiness, we can have no other proof, and we require no other, that these are the only things desirable. If so, happiness is the sole end of human action, and the promotion of it is the test by which to judge all human conduct; *from whence it necessarily follows that* it must be the criterion of morality, *since* a part is included in the whole (Mill 1871, p. 58).

Again, whether Mill’s proof is valid or sound is another matter (see, e.g., Atkinson 1957).

Accordingly, if philosophy is indeed a priori, and in the business of discovering necessary truths from the armchair, then the sort of arguments we would expect philosophers to make are deductive arguments, namely, the sort of arguments whose premises, if true, “provide *conclusive* support for their conclusions,” “*necessitate* the truth of the conclusion,” and “are *indefeasible*” (Fogelin and Sinnott-Armstrong 2005, p. 251; emphasis added). This means that we would expect philosophers to frame their arguments using mostly deductive indicators, such as ‘necessarily’, ‘certainly’, and ‘conclusively’, rather than inductive indicators, such as

⁵ In our experimental study, we do not distinguish between inductive arguments and abductive arguments (or inferences to the best explanation). Insofar as they are the sort of arguments whose premises, if true, make the conclusion more likely to be true or probable, rather than necessitate or guarantee its truth, abductive arguments can be considered a type of inductive arguments. It would be interesting, however, to use our empirical methodology to study how abductive arguments are used in philosophical practice. We leave this question for another occasion.

‘probably’, ‘plausibly’, and ‘likely’. If we find that philosophers frame their arguments using mostly inductive indicators, then that would suggest that philosophers advance mostly inductive, not deductive, arguments. But this would be contrary to what we would expect if philosophy were a priori, and in the business of discovering necessary truths from the armchair. Since, unlike deductive arguments, which are the sort of arguments whose premises (if true) necessitate the truth of their conclusions, inductive arguments are the sort of arguments whose premises (if true) make their conclusions probably (or more likely to be true). In other words, deductive inference is monotonic, whereas inductive inference is ampliative and non-monotonic (Levi 2005, pp. 177-192).

To be clear, we are not suggesting that, according to the “armchair philosophy” picture, all philosophical arguments must be deductive arguments. As Gutting (2009, p. 11) points out:

The days are long gone when adequate philosophical argument had to be valid deduction from self-evident premises. We allow that a good philosophical argument may be inductive or based on premises expressing widely shared common-sense judgments.

However, if the “armchair philosophy” picture of philosophical methodology is accurate, i.e., if philosophy is indeed a priori, and in the business of discovering necessary truths from the armchair, then we would expect philosophers to advance mostly deductive, not inductive, arguments. In addition, both inductive arguments and deductive arguments may work together. For example, inductive arguments may be given as the evidence for the premises of a deductive argument. But such reliance on induction, even in support of premises, would still not be in accordance with the “armchair philosophy” picture painted above.

It is also important to note that indicator words are not a *foolproof* method for detecting arguments in context. As they instruct students to look for indicator words when trying to

identify arguments in context, most logic textbooks are also careful to point out that the presence or absence of indicator words does not *guarantee* the presence or absence of an argument. For instance, according to Hurley (2016, p. 15), “the mere occurrence of an indicator word by no means guarantees the presence of an argument.” Likewise, the mere absence of an indicator word does not guarantee the absence of an argument. Nevertheless, indicator words are still pretty good *indicators* of the presence and/or absence of arguments, which is why logic textbooks instruct students to look for them. Furthermore, since the arguments we have looked at for the purposes of our empirical study are arguments made by professional philosophers, we can be pretty confident that professional philosophers, who are trained in logic and argumentation, seldom misuse indicator words in an effort to dress up arguments artificially.

In order to get an idea of how many philosophy articles employ deductive and inductive arguments, we searched the JSTOR database using the deductive and inductive indicator words outlined in Table 2. We wanted to ensure that the indicators were actually attached to an argument, though, so we combined argument indicators from Table 1 with deductive indicators from Table 2, as well as argument indicators from Table 1 with inductive indicators from Table 2, to create indicator pairs. The argument indicators were distributed resulting in 32 deductive indicator pairs and 32 inductive indicator pairs.

Table 3. A list of deductive and inductive indicator pairs.

Deductive pairs	Inductive pairs
therefore necessarily	therefore probably
therefore certainly	therefore likely
therefore definitely	therefore unlikely
therefore absolutely	therefore improbable
hence necessarily	hence probably
hence certainly	hence likely
hence definitely	hence unlikely
hence absolutely	hence improbable
so necessarily	so probably
so certainly	so likely
so definitely	so unlikely
so absolutely	so improbable
accordingly necessarily	accordingly probably
accordingly certainly	accordingly likely
accordingly definitely	accordingly unlikely
accordingly absolutely	accordingly improbable
consequently necessarily	consequently probably
consequently certainly	consequently likely
consequently definitely	consequently unlikely
consequently absolutely	consequently improbable
proves necessarily	proves probably
proves certainly	proves likely
proves definitely	proves unlikely
proves absolutely	proves improbable
thus necessarily	thus probably
thus certainly	thus likely
thus definitely	thus unlikely
thus absolutely	thus improbable
follows necessarily	follows probably
follows certainly	follows likely
follows definitely	follows unlikely
follows absolutely	follows improbable

We searched simultaneously for any match of the 32 indicator pairs where the indicator terms were within 3 words of each other. In addition to the original 3-word search, we also conducted searches allowing 6 and 10 words between indicator words. Some examples of the search results are listed here (emphasis added):

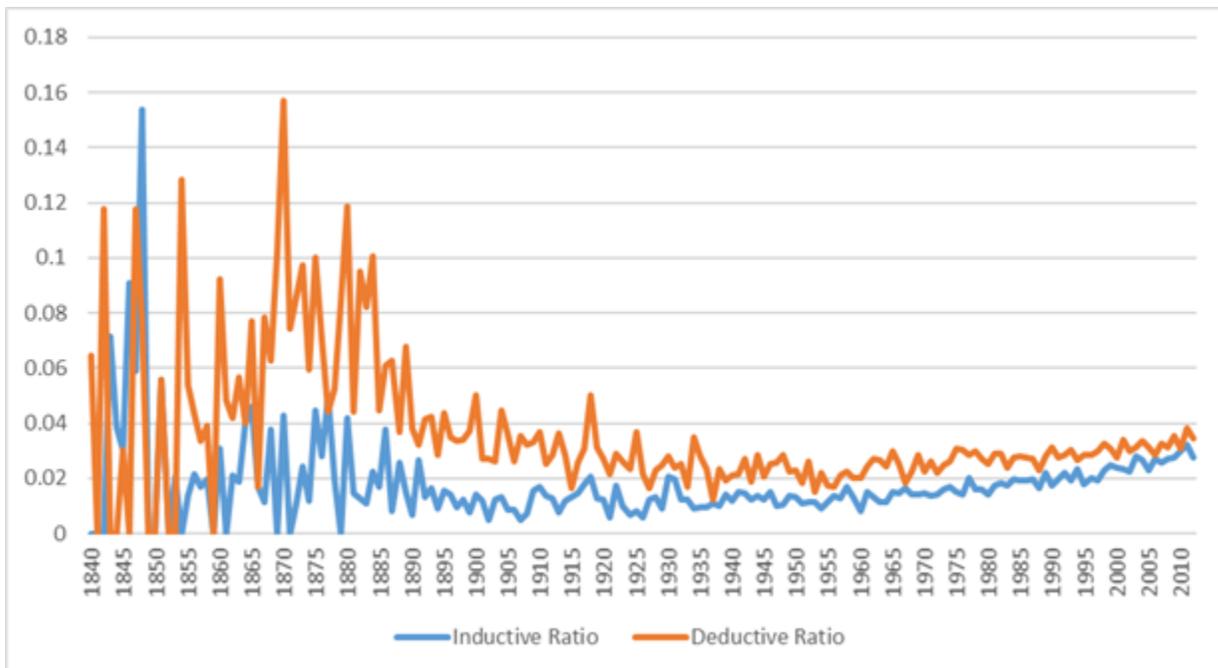
1. “The method for creating better theory and better practice (always at the heart of the pragmatists’ vision) is *therefore necessarily* collaborative, with action tested by many in a variety of circumstances” (Ronald and Roskelly 2001, p. 619).
2. “*Thus* there *certainly* seems to be categories whose basic members are verbs that combine with an adverb somewhere along the route to becoming heads of complete VP’s” (McConnell-Ginet 1982, p. 165).
3. “*Hence*, it is very *likely* that folk concepts of the pragmatic dimension of intentional talk are more richly understood than the core notions of the cognitive machinery that underlies intentional action” (Adams and Steadman 2004, p. 174).
4. “It *therefore* seems *unlikely* that the difference between people’s responses to the harm vignette and their responses to the help vignette is due entirely to pragmatic factors” (Knobe 2004, p. 184).

By choosing this search methodology, the searches return articles with one or more of the indicator pairs. One of the limitations of the JSTOR Data for Research corpus (<http://dfr.jstor.org/>) is that we are unable to identify how many times each indicator pair shows up. As such, the results returned by the searches can only show the prevalence of articles containing each type of argument, rather than the prevalence of each type of argument within articles. To get a better idea of prevalence relative to the number of articles within the JSTOR corpus, we focused on the ratio of search results to total philosophy articles within the corpus and the percentage of matching articles in the database. We chose to look at the years 1840 to 2012 because early years often have so few articles that results are not reliable.

3. Results

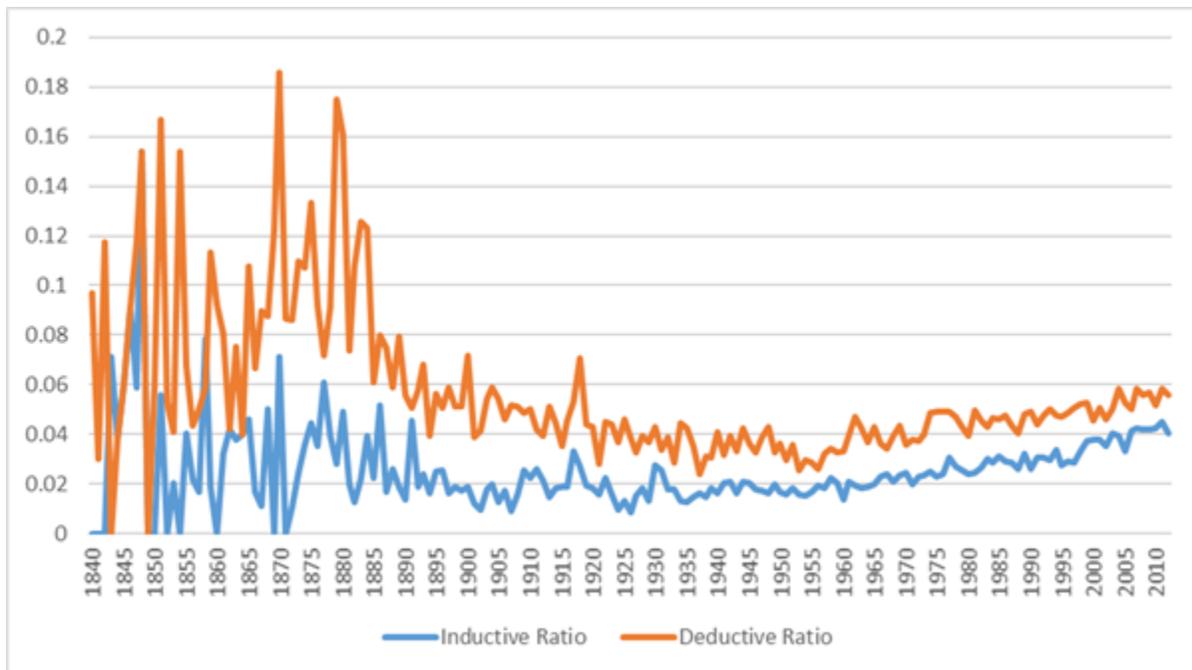
In searches allowing 3 words between indicator terms, the overall percentage of matches is low, with an average ratio of 3.67% for deductive indicators and 1.81% for inductive indicators. In addition, the ratio of articles advancing deductive arguments is almost always higher than those advancing inductive arguments (see Figure 1). The ratios are converging over time, with more recent years seeing small differences between inductive and deductive ratios. To test this, we looked at a linear model of the differences. The equation for the line was $y = -0.0003x + 0.0463$. Because the slope of the line is negative, we can say the overall trend is that the difference between inductive and deductive ratios is in fact decreasing over time.

Figure 1. Percentage of philosophy articles in the JSTOR corpus with inductive or deductive indicators within 3 words of each other.



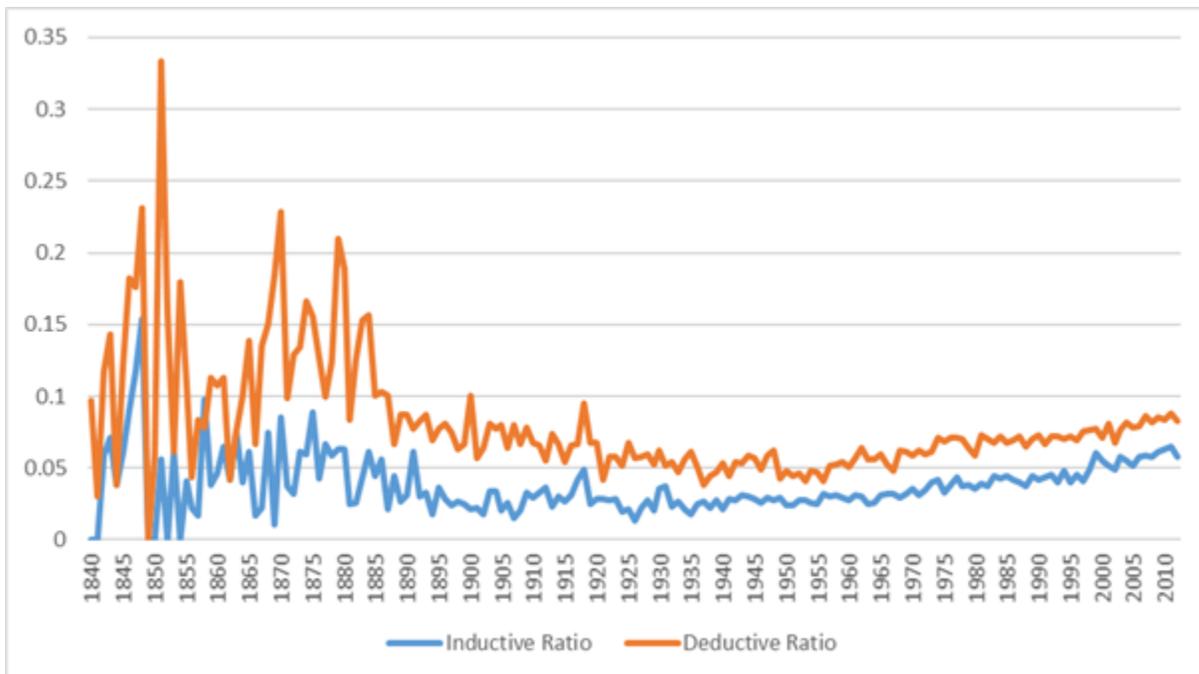
As we would expect, searches allowing 6 words between the indicator words resulted in higher overall ratios (see Figure 2). In this case, the average percentage of articles advancing inductive arguments is 2.59% and the average deductive ratio is 5.67%. Consistent with the less permissive search, we see that the difference in ratios is still declining over time. We again developed a linear model of the ratio differences over time with the linear equation $y = -0.0003x + 0.0591$. The negative slope again indicates that the percentage of articles advancing inductive arguments is converging with the percentage of articles advancing deductive arguments. The overall decrease in difference is consistent with the results from the less permissive search.

Figure 2. Percentage of philosophy articles in the JSTOR corpus with inductive or deductive indicators within 6 words of each other.



The most permissive search we performed allowed for up to 10 words between the indicator words. This search returned similar trends as the less permissive searches while increasing the overall percentage of matching articles (see Figure 3). The average percentage of articles advancing inductive arguments is 3.81% and the average percentage of articles advancing deductive ones is 8.13%. Again, we modelled the difference in these percentages over time. The result is another linear model, this time with the equation $y = -0.0004x + 0.0762$, indicating that the two ratios are converging over time with the smallest differences being the most recent.

Figure 3. Percentage of philosophy articles in the JSTOR corpus with inductive or deductive indicators within 10 words of each other.



Our last set of tests involved rigorously comparing the inductive ratios to the deductive ratios over time. As evidenced above, the deductive ratios tend to be higher but we also tested this hypothesis using paired t-tests. For the least permissive search (maximum 3 words between indicators), the percentage of articles advancing deductive arguments is statistically higher than the percentage of articles advancing inductive arguments, $t(172) = -8.90$ and one-sided $p = 3.69 \times 10^{-16}$. These results are consistent for searches allowing 6 words between indicators, $t(172) = -13.57$ and one-sided $p = 2.79 \times 10^{-29}$, and searches allowing 10 words between indicators, $t(172) = -15.77$ and one-sided $p = 1.53 \times 10^{-35}$. Overall, this indicates that the percentage of articles advancing deductive arguments is higher than the percentage of articles advancing inductive arguments.

4. Discussion

To sum up, our key findings are the following:

1. The ratio, and likewise percentage, of philosophy articles advancing deductive arguments is higher than the percentage of articles advancing inductive arguments.
2. The ratios and percentages of philosophy articles advancing deductive arguments and those advancing inductive arguments are converging over time.
3. The difference between inductive and deductive ratios, and their percentages, is declining over time.

On the “armchair philosophy” picture of philosophical methodology, finding 1 is to be expected. As we mentioned in Section 2, if philosophy is indeed a priori, and in the business of discovering necessary truths from the armchair, then we would expect philosophers to frame their arguments using mostly deductive indicators, such as ‘necessarily’, ‘certainly’, and ‘conclusively’, rather

than inductive indicators, such as ‘probably’, ‘plausibly’, and ‘likely’. This is because, unlike inductive arguments, deductive arguments are the sort of arguments whose premises, if true, “provide *conclusive* support for their conclusions,” “*necessitate* the truth of the conclusion,” and “are *indefeasible*” (Fogelin and Sinnott-Armstrong 2005, p. 251; emphasis added).

However, findings 2 and 3 are not quite what we would expect if philosophy is a priori, and in the business of discovering necessary truths from the armchair. It is not the fact that philosophy articles sometimes advance inductive arguments that is unexpected on the “armchair philosophy” picture. As we mentioned in Section 2, inductive arguments are now “allowed” in philosophy (Gutting 2009, p. 11). Rather, what is unexpected, from the vantage point of the “armchair philosophy” picture, is that the ratios and percentages of articles advancing deductive arguments and those advancing inductive arguments are converging over time. In other words, the results of our experimental study reveal a trend in philosophical methodology. The trend is that the difference between the ratios of articles advancing deductive arguments and those advancing inductive arguments is decreasing over time. To put it another way, it looks like deductive arguments are gradually losing their status as the dominant form of argumentation in philosophy.

These findings, namely, 2 and 3, are not quite consistent with those of Knobe (2015). Knobe (2015) compared two samples of papers in philosophy of mind, one from 1960 to 1999 and another from 2009 to 2013, and found that the percentage of papers using purely a priori methods declined from 62% in the 1960-1999 sample to 12% in the 2009-2013 sample. On the other hand, our data show that, overall, the ratio of philosophy articles advancing deductive arguments is higher than those advancing inductive arguments, although the difference between the ratios of articles advancing deductive arguments and those advancing inductive arguments is

decreasing over time. In that respect, it is important to note that our sample is much larger than Knobe's samples, for our searches are not restricted to philosophy of mind and include, not only the most highly cited papers in Google Scholar Metrics and Web of Science, but all philosophy articles in the JSTOR corpus from 1840 to 2012. For these reasons, Knobe's (2015) data do not warrant any conclusions about trends in philosophical methodology, whereas our data support our claim that the overall trend is that the difference between the ratios of articles advancing deductive arguments and those advancing inductive arguments is decreasing over time.

Of course, our data do not tell us what might be the cause (or causes) of this trend. So we can only speculate that the cause might have something to do with the increasing popularity of empirical methods in philosophy, such as experimental philosophy, and the like. That is why we think that further studies are needed in order to understand the trend in philosophical methodology that our experimental study has uncovered.

5. Conclusion

In this paper, we set out to empirically test the "armchair philosophy" picture of philosophical methodology, according to which philosophy is a priori, and in the business of discovering necessary truths from the armchair. We would expect the "armchair philosophy" picture to be reflected in philosophical practice in terms of the sort of arguments philosophers advance. In particular, we would expect philosophers to advance mostly deductive, rather than inductive, arguments, which are the sort of arguments whose premises, if true, necessitate the truth of their conclusions.

The results of our experimental study show that the percentage of philosophy articles advancing deductive arguments is higher than the percentage of articles advancing inductive

arguments, which is what we would expect from the vantage point of the “armchair philosophy” picture. However, contrary to what we would expect to see reflected in philosophical practice if the “armchair philosophy” picture is accurate, our results show that the percentages of articles advancing deductive arguments and those advancing inductive arguments are converging over time and that the difference between inductive and deductive ratios is declining over time. This trend suggests that deductive arguments are gradually losing their status as the dominant form of argumentation in philosophy. We propose that further studies are needed to understand the cause (or causes) of this trend.

Acknowledgments

We are grateful to an anonymous reviewer of *Metaphilosophy* for helpful comments on an earlier draft.

References

Adams, F. and Steadman, A. (2004). Intentional Action in Ordinary Language: Core Concept or Pragmatic Understanding? *Analysis* 64(2): 173-181.

Atkinson, R. F. (1957). J. S. Mill’s “Proof” of the Principle of Utility. *Philosophy* 32(121): 158-167.

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

Bealer, G. (1998). Intuition and the Autonomy of Philosophy. In M. R. DePaul and W. Ramsey (eds.), *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry* (pp. 201-240). Lanham, MA: Rowman & Littlefield.

Copi, I. M., Cohen, C., and McMahon, K. (2011). *Introduction to Logic*. Fourteenth Edition. New York: Prentice Hall.

Deutsch, M. (2015). *The Myth of the Intuitive: Experimental Philosophy and Philosophical Method*. Cambridge, MA: MIT Press.

Fogelin, R. J. and Sinnott-Armstrong, W. (2005). *Understanding Arguments*. Seventh Edition. Belmont, CA: Thomson Wadsworth.

Gutting, G. (2009). *What Philosophers Know: Case Studies in Recent Analytic Philosophy*. New York: Cambridge University Press.

Henderson, D. and Horgan, T. (2013). On the Armchair Justification of Conceptually Grounded Necessary Truths. In A. Casullo and J. C. Thurow (eds.), *The A Priori in Philosophy* (pp. 111-133). Oxford: Oxford University Press.

Horvath, J. and Grundmann, T. (2012). Introduction. In J. Horvath and T. Grundmann (eds.), *Experimental Philosophy and Its Critics* (pp. 1-10). New York: Routledge.

Hurley, P. J. (2016). *Logic: The Essentials*. Stamford, CT: Cengage Learning.

Knobe, J. (2004). Intention, Intentional Action and Moral Considerations. *Analysis* 64(2): 181-187.

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

Kriegel, U. (2016). Philosophy as Total Axiomatics: Serious Metaphysics, Scrutability Bases, and Aesthetic Evaluation. *Journal of the American Philosophical Association* 2(2): 272-290.

Levi, I. (2005) Inductive Inference as Ampliative and Non-monotonic Reasoning. In R. van der Meyden (ed.), *Proceedings of the 10th Conference on Theoretical Aspects of Rationality and Knowledge* (pp. 177-192). National University of Singapore: ACM Digital Library.

McConnell-Ginet, S. (1982). Adverbs and Logical Form: A Linguistically Realistic Theory. *Language* 58(1): 144-184.

Melnyk, A. (2008). Conceptual and Linguistic Analysis: A Two-Step Program. *Nous* 42(2): 267-291.

Mill, J. S. (1871). *Utilitarianism*. Fourth Edition. London: Longmans, Green, Reader, and Dyer.

Mizrahi, M. (2012). Why the Ultimate Argument for Scientific Realism Ultimately Fails. *Studies in History and Philosophy of Science Part A* 43(1): 132-138.

Moore, G. E. (1959). *Philosophical Papers*. Second Impression 1963. London: Routledge.

Restall, G. (2009). A Priori Truths. In J. Shand (ed.), *Central Issues of Philosophy* (pp. 37-50). Malden, MA: Wiley-Blackwell.

Ronald, K. and Roskelly, H. (2001). Untested Feasibility: Imagining the Pragmatic Possibility of Paulo Freire. *College English* 63(5): 612-632.

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

Sankey, H. (2016). *Scientific Realism and the Rationality of Science*. New York: Routledge.

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

Williamson, T. (2016). Philosophical Criticisms of Experimental Philosophy. In J. Sytsma and W. Buckwalter (eds.), *A Companion to Experimental Philosophy* (pp. 22-36). Malden, MA: Wiley-Blackwell.

Worrall, J. (1989). Structural Realism: The Best of Both Worlds? *Dialectica* 43 (1-2): 99-124.