Potential Sources of Gender Disparities in Philosophy: An Examination of Philosophical Subfields and Their Relation to Gender[1]

Molly Paxton

[1] Thanks to the American Philosophical Association for the grant that funded work on this project. Thanks also to Joseph Paxton for comments and statistical advise. Finally, thanks to David Bourget and David Chalmers for providing the data on which most of this report was ultimately based.

Introduction

This report addresses two empirical questions relevant to issues regarding gender disparities within professional philosophy: the structure of philosophical subfields (or "Areas of Specialization") and differential gender disparities within subfields. Issues regarding gender disparities within philosophy have recently received increased attention (APA Committee on the Status of Women, 2014). However, to this point, the nature of these disparities has been addressed mainly at the level of professional philosophy as a whole[2], without particular concern for whether some philosophical subfields may display greater disparities than others. In order to more effectively target potential remedies to the issue of gender disparities in philosophy, it would be beneficial to know whether gender disparities differ across subfields.

In doing so, it will first be necessary to identify what these subfields are. There is an existing resource for identifying philosophical subfields—viz., the *Philosophical Gourmet Report*, which is compiled and edited by the philosopher Brian Leiter, in consultation with an advisory board of approximately fifty philosophers from around the English-speaking world (as of the most recent *Report*).[3] Specifically, the *Gourmet Report* contains a *Breakdown of Programs by Specialties*, in which experts in each of a list of approximately thirty subfields provide ratings for the strength of each department in their subfields.

The scope of the *Gourmet Report* is broad, and a new version is released every few years. However, the *Report* does not describe how the list of subfields within the *Breakdown of Programs by Specialties* was created. It is not implausible that this list is based largely on the impressions of the editor, perhaps in consultation with the advisory board. If so, this would not necessarily be a problem, as the editor and advisory board have relevant expertise when it comes to identifying the structure of the discipline. However, a cursory glance at the *Curricula Vitae* (CVs) of most any arbitrarily chosen set of professional philosophers will reveal that self-reported Areas of Specialization cannot be sorted cleanly and completely into the *Gourmet Report's* subfield list.

Although this fact would be no surprise to the editors, it does suggest that there may be room for improvement in the *Report's* subfield list. How might we go about making such improvements?

Alternative Categorization Strategies

One way to supplement such lists of philosophical subfields is to look to philosophers' self-reported *Areas of Specialization*. It is plausible that these Areas are not independent of each other. For example, if a given philosopher is interested in "Perception," it would not be surprising to find that they are also interested in "Theories of Mental Content." These Areas of Specialization will therefore tend to co-occur, allowing one to reduce what at first may have appeared to be two different Areas of Specialization into a single Area—e.g., Philosophy of Mind. Statistical methods such as Principal Components Analysis (PCA) can be used to identify

---

[2] And even then, mainly within the English-speaking world.

[3] Shortly after this report was completed, Leiter announced that he would be joined by a co-editor for the 2014 version of the *Gourmet Report*, and will be stepping down thereafter.

such co-occurrences, thereby reducing a large and potentially unwieldy list of Areas to a more coherent and manageable list. Such a list could be compiled without reference to the judgments of any given philosopher, and could thus be used to supplement existing lists of philosophical subfields, such as the one provided by the *Gourmet Report*.

But is such a method practical, given available Area of Specialization data? Preliminary studies suggest that PCA and similar statistical data reduction methods may indeed be useful in suggesting empirically-derived philosophical subfields by calculating the co-occurrence of philosophers' self-reported Areas of Specialization (Paxton, 2013). These subfields can be thought of as "higher-level" Areas of Specialization. To take another example: if the analysis detects a highly regular pattern of co-occurrence between categories such as "Phenomenology" and "Existentialism," PCA would indicate this in the form high "loadings" for each of these categories on a single *factor* or *component*—i.e., a higher-level or latent category. This would tell us that subjects who indicated a specialization in Phenomenology often also indicated a specialization in Existentialism and vise versa. This would then allow us to reduce the data by grouping Phenomenology and Existentialism into a single category, perhaps called "Continental Philosophy." While some of these Areas may be intuitive to many philosophers—such as the examples used here—this method has the potential to reveal patterns among Areas of Specialization that are currently not recognized in existing lists.

Using such statistical methods to map distinctive patterns among philosophers' Areas of Specialization may allow us to arrive at a more informative list of philosophical subfields than has been available previously. Obtaining a manageable list of subfields in this way would then allow us to provide a more accurate characterization of which subfields are most affected by gender disparities and other forms of underrepresentation. Furthermore, such a list of empirically-derived subfields would allow more accurate data collection for any future studies or surveys within the discipline.

Below, I describe the statistical methods used to generate this list, with the goal of examining gender disparities within these subfields.

Method

To address the questions outlined above, I first attempted to create an archival database of philosophers in the United States and Canada. To create this database, I began by downloading a large number of CVs ($n \approx 500$) that professional philosophers in these countries made available online, either through their personal websites or through their department websites. Personal websites were found by following links from department websites. Department websites were found using Keith DeRose's list of websites for PhD-granting philosophy departments (DeRose, 2007). Because this list was last updated in 2007, many links in this list were broken. For these, I performed a Google search for "Philosophy" followed by "[University Name]." A majority of the time, this method yielded a current link. Department websites that could not be found using Google were removed from the list.

Of the CVs that were located in this way, approximately half clearly listed Areas of Specialization. These Areas of Specialization were grouped by individual, such that the resulting dataset contained rows corresponding to individual philosophers, and columns corresponding to their Areas of Specialization. This dataset was then submitted to a Principal Components Analysis using varimax rotation, which is a standard method for deriving components.

Unfortunately, the Principal Components Analysis failed to converge on a reliable solution with this dataset. Reliability was determined using a split-half method in which a random half of the data were selected, a PCA was performed on these data, and the process was repeated for the other half of the data. The resulting solutions were manually examined, and were found to overlap only minimally. In other words, the list of components (empirically-derived subfields) yielded by the first random half of the data had little in common with the list of components yielded by the second random half of the data.

One way to increase the reliability of such an analysis is to collect additional data. In this case, the archival method of data collection was an extremely time-intensive process, taking more than six months to complete for approximately 500 CVs. Even still, I decided to collect additional CVs to ensure reliable results.

While doing so, another dataset became available, which immediately made more reliable analyses possible without the collection of additional CVs. Specifically, the *PhilPapers Surveys* was a recent data collection effort designed to determine the philosophical positions held by a large sample of professional philosophers (Bourget & Chalmers, 2013).[4] Incidentally, the survey had also asked respondents to list their Areas of Specialization, in order to analyze philosophical positions by AOS. Given that this dataset exceeded what would have been possible to collect using prior archival methods in a reasonable amount of time, I decided to abandon further archival data collection in favor of the *PhilPapers Surveys* data.[5]

The survey authors had previously conducted PCA and factor analyses on philosophers' beliefs, but had not examined the AOS data closely, as this was not their focus. Consequently, I repeated the split-half PCA method described above for determining reliability using the *PhilPapers Survey* AOS data. Manual inspection of the results from each random half of the data revealed far more overlap than was present in the much smaller archival dataset. Consequently, I proceeded to submit the full dataset to PCA. The results are described below.

Results

As outlined in the published *PhilPapers Surveys* article (Bourget & Chalmers, 2013), the sample was drawn from regular faculty members at "99 leading departments of philosophy" in the English-speaking world. The survey link was distributed via email to 1972 faculty members, and 931 responded. Consequently, self-selection effects are possible, but should be minimized by the large sample size.[6]

Of the 931 total respondents, 918 listed at least one AOS (Mean=2.31, SD=.82). Of these, 157 were female, 717 were male, and 44 did not specify their gender. This is consistent with past results on the "Gender Gap" in philosophy, which suggest that approximately 20% of philosophy

---

[4] I was aware of the *PhilPapers Surveys* at the outset of my archival data collection, as it was conducted in 2009. However, the survey data were not available at that time, as the authors had not yet completed their analyses of philosophical positions. In addition, it did not seem wise to try to duplicate their efforts. For this reason, archival methods were the only option at the outset for collecting the needed data.

[5] These data were kindly provided by David Bourget and David Chalmers.

[6] At the very least, the current sample is likely to be more representative of the target population than those of the average psychology study, for instance.

faculty are women (Paxton, Figdor, & Tiberius, 2012). In addition, the number of Areas listed did not differ as a function of gender (Males=2.33, Females=2.27; t(872)=.76, p=.45).

The mean age of the sample was 59.84 (SD=12.16). There was a significant negative correlation between age and the number of Areas listed (r=-.16, p<.001). The reasons for this association are not immediately clear. The strength of the association was nearly identical for males (r=-.16, p<.001) and females (r=-.16, p=.06).

Prior to conducting the Principal Components Analysis on the provided Areas of Specialization, I inspected the distribution of AOS frequencies. A total of 63 unique Areas were provided. Of these, 3 were listed by only one or two respondents. Because factor structures resulting from many low frequency responses are often unreliable (in the sense detailed above), I excluded these 33 Areas from further analysis:

Excluded Areas	Count
African/Africana Philosophy	2
Early Analytic	1
Environmental Philosophy	1
Feminist Epistemology	1
Formal Axiology	1
Formal Epistemology	1
History and Philosophy of Human Sciences	1
History of Ideas	1
history of modern	1
History of Philosophy & Science/Math	1
History of Science	1
History of Women's Ideas	1
Indo-Tibetan Philosophy	1
Kant	2
Linguistics	1
Moral Philosophy	1
Moral Psychology	2
Neurophilosophy	1
Philosophy of Art	1
Philosophy of Color	1
Philosophy of Computing and Information	2
Philosophy of Emotions	1
Philosophy of Medicine	1
Philosophy of Physics	1
Philosophy of Psychology	1
Philosophy of the Americas	2
Philosophy of the Body	1
Philosphy of Technology	1
Political Ecology	1
Political Philosophy	1
Pragmatism	1
Theories of Content and Concepts	1
Truth	1

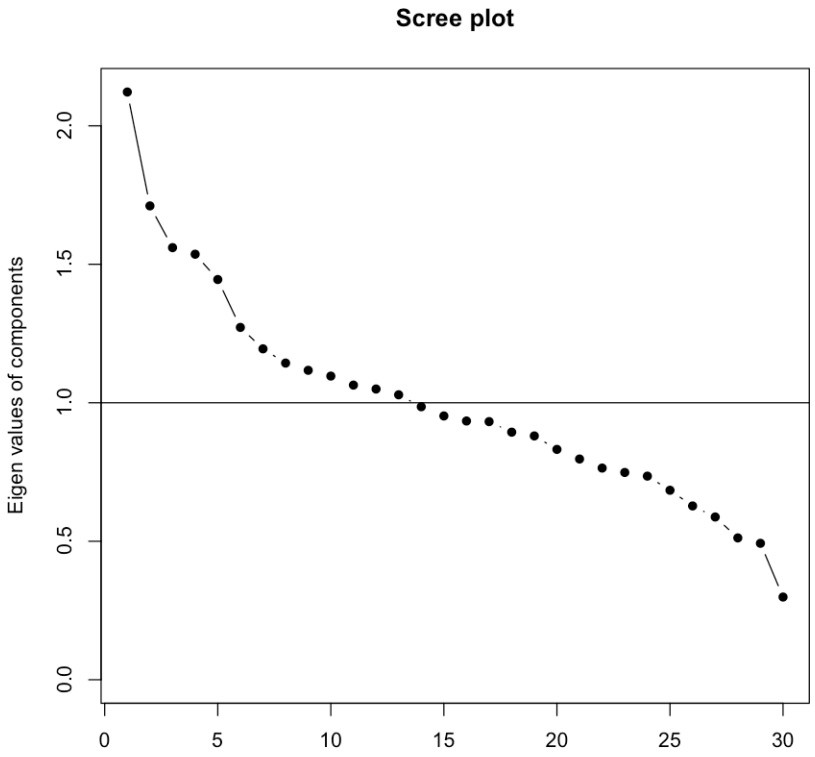
The remaining 30 Areas are the focus of the following analyses:

Included Areas	Count
17th/18th Century Philosophy	107
19th Century Philosophy	37
20th Century Philosophy	42
Aesthetics	38
Ancient Greek Philosophy	64
Applied Ethics	44
Asian Philosophy	10
Continental Philosophy	25
Decision Theory	31
Epistemology	159
General Philosophy of Science	96
Logic and Philosophy of Logic	94
Medieval and Renaissance Philosophy	17
Meta-Ethics	102
Metaphilosophy	21
Metaphysics	233
Normative Ethics	139
Philosophy of Action	43
Philosophy of Biology	38
Philosophy of Cognitive Science	57
Philosophy of Gender, Race, and Sexuality	17
Philosophy of Language	171
Philosophy of Law	28
Philosophy of Mathematics	35
Philosophy of Mind	190
Philosophy of Physical Science	61
Philosophy of Probability	19
Philosophy of Religion	47
Philosophy of Social Science	21
Social and Political Philosophy	100

Principal Components Analysis

A Principal Components Analysis was performed using data from these 30 Areas. A *scree plot* was generated in order to determine the number of components (here, empirically-derived subfields) to select, or *extract*, from the PCA. This scree plot (shown below) provides a visualization of the components (x-axis) and their corresponding *eigenvalues*. Note that eigenvalues indicate the proportion of the variance in the raw data accounted for by each component, net of the other components. For each component, the total variance it accounts for can be determined by taking its eigenvalue and dividing by the total number of components. In

this case (n=30 components), the first component has an eigenvalue of approximately 2, and thus accounts for about 0.07 or 7% of the total variance (2/30=.07).



One common method for determining the number of components to extract is to draw a threshold line corresponding to an eigenvalue of 1.0. You then sum the number of components above this threshold, and extract that number of components. This is often referred to as the "Kaiser criterion." However, this and other purely quantitative methods for determining the number of components to extract are somewhat arbitrary, insofar as they lead you to ignore the content of the components (here, the particular subfields that would be combined into each component), and can differ substantially in their results based on the total number of items that go into the analysis.

For this reason, I employed a two-step approach here. In the first step, the Kaiser criterion was applied for purely practical purposes, in order to yield a manageable number of components to examine. This yielded 13 components. I then performed a PCA using varimax rotation as mentioned above. The items (i.e., Areas) within the resulting 13 components were manually examined to determine whether their content provides information not available in existing lists of philosophical subfields (e.g., the aforementioned *Gourmet Report*).

A given self-reported AOS was included in a given component (empirically-derived AOS) just in case it was a relatively good predictor of that component, as indicated by the *loadings* within the factor pattern matrix output by the PCA. These *loading* are correlations between each variable (self-reported AOS) and component (empirically-derived AOS). Thus, when this correlation is relatively large, it indicate that the variable is a relatively good predictor of the component.

How does one determine what constitutes a *relatively good* predictor? There is again no conventional way of doing so quantitatively: as is often the case, any given threshold can seem

arbitrary. One reasonable policy is thus to choose a threshold based on what has provided sensible results given past experience with similar analyses. On this basis, I chose a threshold of .5 (on a 0-1 range), keeping in mind the choice of a different threshold could yield different conclusions.

Each of the 13 components (empirically-derived AOS), along with their constitutive variables (self-reported AOS) and the loadings of the variables on their components can be found immediately below:

Component 1
- Meta-ethics (.7)
- Normative ethics (.73)

Component 2
- Logic and Philosophy of Logic (.8)
- Philosophy of Mathematics (.79)

Component 3
- General Philosophy of Science (.71)
- Philosophy of Physical Science (.8)

Component 4
- Philosophy of Action (.61)
- Philosophy of Law (.62)

Component 5
- 20th Century Philosophy (.7)
- Continental Philosophy (.71)

Component 6
- Medieval and Renaissance Philosophy (.54)
- Philosophy of Religion (.73)

Component 7
- Philosophy of Gender, Race, and Sexuality (.71)
- Social and Political Philosophy (.65)

Component 8
- Decision Theory (.62)
- Epistemology (.53)
- Philosophy of Probability (.54)

Component 9
- 17th/18th Century Philosophy (.59)
- 19th Century Philosophy (.54)
- Aesthetics (.6)

Component 10
- Philosophy of Biology (.78)

Component 11
- Ancient Greek Philosophy (.79)
- Asian Philosophy (.52)

Component 12
- Philosophy of Social Science (.81)

Component 13
- Metaphilosophy (.73)

Perhaps the most obvious point to be made about these components is that they do not include all 30 variables entered into the analysis. This is in part because the remaining variables (e.g., Applied Ethics) were relatively poor predictors of these 13 components. The fact that this analysis also omits Areas of Specialization that are often regarded as central (e.g., Metaphysics) may raise concerns about its ability to supplement existing lists of philosophical subfields.

However, these results may yet be informative. For example, it is interesting to note that Component 6 would have us combine Medieval and Renaissance Philosophy with the Philosophy of Religion into a single subfield. Of course, there is an obvious sense in which these two Areas of Specialization come apart--e.g., one could do Medieval Philosophy for it's own sake from a historical perspective without engaging in substantive debates in the Philosophy of Religion. However, a notable consistency in the present results comes from the 2011 *Gourmet Report* (the most recent version at the time of writing) where the top 6 department in the *Breakdown of Programs by Specialty* for the Philosophy of Religion are a proper subset of the top 9 departments for Medieval Philosophy. This might lead one to believe that, while it is true that one can do Medieval Philosophy without engaging in the Philosophy of Religion, it is more difficult to engage in the latter without knowledge of the former.

In addition, it is worth noting that Philosophy of Physical Science loads with General Philosophy of Science in Component 3, while Philosophy of Biology (Component 10) and Philosophy of Social Science (Component 12) do not load with General Philosophy of Science. This may reflect the fact that, until quite recently, Philosophy of Science essentially *was* Philosophy of Physical Science. Thus, the main way that one can claim an AOS in General Philosophy of Science is if one is interested in Philosophy of Physical Science--otherwise you must think of yourself as more of a specialist (in Biology, Social Science, and the like).

Other interesting implications could likely be drawn from this analysis. However, contrary to prior results that gave us reason to be optimistic about the possibility of arriving at empirically-derived philosophical subfields (Paxton, 2013), the present results with a much larger dataset do not seem to support this optimism, particularly given the absence of core philosophical subfields in the latent components. Although the detailed reasons for this are not immediately clear, one possibility is that philosophers who list certain core subfields as an AOS (e.g., Metaphysics) also tend to list a *variety* of other subfields (e.g., Philosophy of Religion, Philosophy of Mind, Philosophy of Language, and so on). However, because there is no *particular* subfield that reliably co-occurs with the core subfield, the core subfields does not load with other subfields on a component, resulting in the absence of the core subfield on the list of components.

Future work on the structure of the discipline using larger datasets and different statistical methods will be needed to make further progress on this issue. In the interest of providing informative conclusions regarding gender differences between philosophical subfields--which was the other main goal of this report--the analyses that follow will focus on the original list of 30 variables (self-reported AOS) rather than the 13 components (empirically-derived AOS).

Gender Differences

When examining gender differences as a function of self-reported AOS, I took two partially distinct approaches. The first approach was to ask whether each AOS differs from a 50/50 gender ratio, or what might be called "gender parity." However, because the overall gender split in philosophy is known to be far from parity (see above), these analyses are very likely to

reveal a near uniform lack of parity across subfields. For this reason, I also conducted a second set of analyses that asked whether the gender balance in each AOS differed from the overall "base rate" across all subfields (viz., 82% male to 18% female). These base rate analyses are likely to be more sensitive to differences across subfields, as they ask how similar each subfield is to the gender-imbalanced profession, rather than asking whether each subfield is itself gender-imbalanced.

Both sets of analyses were conduced using a binomial test, which allows you to ask whether the distribution of observations in two categories (here "male" and "female") differs from a specified distribution. The specified distribution in the first case was 50% female (50% male), while in the second case it was 18% female (82% male). The results of both the gender parity and base rate analyses follow.

As expected, the first set of analyses reveal a near uniform lack of gender parity across subfields. The only subfields that do not have significantly more men than women are: Applied Ethics (40% female, p=.75); Asian Philosophy (36% female, p=.12); and Philosophy of Gender, Race, and Sexuality (81% female). Only the latter exhibited a significant majority of women (p=.02).

Focusing on the base rate analyses reveals only a slightly more nuanced picture. The following subfields have significantly fewer women than one would expect, even given the low base rate: Logic and Philosophy of Logic (10% female, p=.04); Metaphysics (12% female, p=.01); and Philosophy of Physical Science (5% female, p=.009). There were no women in the sample who claimed an AOS in Philosophy of Probability, although this result fell just shy of significance (p=.06) due to the relatively small sample size for that AOS (n=19). In contrast, there are significantly more women than one would expect given the overall base rate in Asian Philosophy (36% female, p=.004) and Philosophy of Gender, Race, and Sexuality (81% female, p<.001).[7]

Conclusion

Although the present attempt to supplement existing lists of philosophical subfields was largely inconclusive, analyses of gender disparities across self-reported Areas of Specialization do allow us to draw clear conclusions which largely obviate the need for concern over the structure of the field. Significant male majorities are present across philosophical subfields, with only three of thirty subfields resisting that trend. Furthermore, the vast majority of subfields (25/30) do not significantly differ from the gender base rate within the discipline as a whole. Finally, a small but nearly identical number of subfields have less women than the base rate (3) as have more women than the base rate (2). These results reveal that large gender disparities (over and above the lack of gender parity) are present across nearly the entire field. Together, these two sets of results strongly suggest that any potential remedy for gender disparities within philosophy would not benefit from targeting particular subfields, but would do well to target the discipline as a whole.

---

[7] Note that there results are not corrected for multiple comparisons. However, doing so would leave the substantive conclusions regarding the parity analyses largely the same, although the male majorities would no longer be statistically significant in the handful of borderline cases.

Bibliography

APA Committee on the Status of Women (Ed.). (2014). *Data on Women in Philosophy*. Retrieved October 2014, from http://www.apaonlinecsw.org/data-on-women-in-philosophy

Bourget, D., & Chalmers, D. J. (2013). What do philosophers believe? *Philosophical Studies*, *170*(3), 465–500.

DeRose, K. (Ed.). (2007). *U.S. Philosophy Departments*. Retrieved October 2014, from http://pantheon.yale.edu/~kd47/depts.htm

Paxton, M. J. (2013). 2012/2013 American Philosophical Association Pacific Division Conference Participation Data Report. *Unpublished Manuscript*.

Paxton, M. J., Figdor, C., & Tiberius, V. (2012). Quantifying the Gender Gap: An Empirical Study of the Underrepresentation of Women in Philosophy. *Hypatia*, *27*(4), 949–957.