Forthcoming in Knobe & Nichols (eds.) *Experimental Philosophy*, Vol.2

Gender and Philosophical Intuition[1]

Wesley Buckwalter
City University of New York Graduate Center

Stephen Stich
Rutgers University

In recent years, there has been much concern expressed about the underrepresentation of women in academic philosophy in the English speaking world. A full explanation of this troubling phenomenon is likely to be quite complex since there are, almost certainly, many factors that contribute to this gender disparity. Our goal in this paper is to call attention to a cluster of phenomena that may be contributing to the underrepresentation of women in philosophy, though until now these phenomena have been largely invisible. The findings we review indicate that when contemporary American and Canadian women and men with little or no philosophical training are presented with standard philosophical thought experiments, in many cases their intuitions about these cases are significantly different. We suspect that these differences could be playing a significant role in shaping the demography of the profession. But at present this is *only* an hypothesis, since we have no evidence that bears directly on the causal relation between the gender gap in academic philosophy and the facts about intuition that we will recount. In future work, we plan to focus on that causal link. However, we believe that the facts we report about gender differences in philosophical intuitions are both important and disturbing, and that philosophers (and others) should begin thinking about their implications both for philosophical pedagogy and for the methods that philosophers standardly use to support their theories. It is our hope that this paper will help to launch conversations on these issues within the philosophical community and beyond.

Here is how we plan to proceed. In section 1 we briefly review some of the data on the underrepresentation of women in academic philosophy. In section 2 we explain how we use the term 'intuition', and offer a brief account of how intuitions are invoked in philosophical argument and philosophical theory building. The third section is the longest and most important.

---

[1] We are very grateful to Helen Beebee, Fiery Cushman, Carrie Figdor, Ori Friedman, Joshua Greene, Geoffrey Holtzman, Jonathan Livengood, Edouard Machery, Shaun Nichols, Molly Paxton, Jennifer Saul, Christina Starmans, Justin Sytsma, Valerie Tiberius and Liane Young, all of whom shared their data with us and provided clear and careful answers to our many questions about their work. Special thanks to our research assistant, Michael Sechman. We would also like to thank Louise Antony, Helen Beebee, Michael Bishop, Ned Block, Paul Bloom, Tamar Gendler, Joshua Knobe, Tania Lombrozo, Brian Leiter, Edouard Machery, Ron Mallon, Shaun Nichols, Richard Nisbett, Jesse Prinz, Jennifer Saul, Christina Starmans, Valerie Tiberius, Virginia Valian and Jonathan Weinberg, as well as members of the MERG (Metro Experimental Research Group) lab, the Moral Psychology Research Group, and the Society for Philosophy and Psychology for insightful comments on earlier versions of this material.

It is in that section that we set out the evidence for gender differences in philosophical intuition.[2] We also mention some evidence about gender differences in decisions and behaviors that are (or at least should be) of considerable interest to philosophers. In the fourth section, our focus switches from facts to hypotheses. In that section we explain how differences in philosophical intuition might be a hitherto neglected part of the explanation for the gender gap in philosophy. The fifth section is a brief conclusion.

1. Some Data on the Gender Gap in Academic Philosophy

In a powerful and much discussed paper about the challenges confronting women who pursue academic careers in philosophy, Sally Haslanger (2008) included a table indicating the number and percent of women faculty in the top twenty American philosophy departments, as ranked by the *Philosophical Gourmet Report* (more commonly known as the *Leiter Report*).[3] Using data reported by Leiter, Haslanger computed that 19.5% of the tenured and tenure track faculty in these departments were women. Checking and updating Leiter's data, Haslanger found that the percentage of women faculty was actually a bit lower: 18.7%. Since these data were collected in 2006, we thought it would be interesting to see if things had changed. Using Leiter's rankings in the 2009 *Philosophical Gourmet Report* and current philosophy department websites, we found that the situation has changed very little (Table 1).[4] Using the same method, we looked at the top four Leiter-ranked philosophy departments in Canada and the top four in Australasia. By our count there are 216 regular faculty in these eight departments, 55 of whom (25.4%) are women.[5] The numbers cited thus far are from highly ranked philosophy departments. For a more comprehensive measure, we turned to the report on women in philosophy in the UK by Professors Helen Beebee and Jennifer Saul (2011), written for the

---

[2] More accurately, what the evidence indicates is that there are gender differences in philosophical intuitions *among contemporary residents of the U.S.A. and Canada.* As we emphasize in section 3.9, whether these differences exist in other groups is currently unknown. Since repeating the italicized phrase makes for awkward prose, we will often omit it. But the reader should keep in mind that the qualification is always intended.

[3] http://www.philosophicalgourmet.com/

[4] The websites of the philosophy departments mentioned in this section were accessed on various days from March 8 through March 19, 2010. In assembling Table 1, we made no attempt to correct the information we found on departmental websites, though in several cases we had good reason to believe that the websites did not accurately reflect the current department membership. In a number of cases, we had to make decisions about whether to count people who had various sorts of split or part-time appointments. We consulted with Professor Haslanger, and tried to apply the same rules of thumb that she had employed. But, inevitably, this table reflects a number of "judgment calls."

[5] Here again we had to make quite a few "judgment calls". The people we counted were on lists with a variety of labels including "permanent faculty," "main faculty," "continuing positions" and "academic staff".

British Philosophical Association and the Society for Women in Philosophy joint committee for women in philosophy. According to the figures provided by 38 Heads of Department (including the largest departments) in response to a questionnaire, there are 448 permanent staff (Lecturers, Senior Lecturers, Readers and Professors) in these 38 departments, 24% of whom are women.

Table 1. Gender Ratios in Tenure-Track Positions by Faculty Rank in United States Philosophy Departments Ranked 1-20 in the 2009 Leiter Report as of March 2010.

Department	Women	Total	Percentage	Full Professor			Associate			Assistant		
				W	T	%	W	T	%	W	T	%
NYU	5	30	17%	1	24	4%	1	2	50%	3	4	75%
Rutgers	5	33	15%	3	25	12%	2	8	25%	0	0	NA
Princeton	4	23	17%	0	12	0%	1	4	25%	3	7	43%
Pittsburgh	3	22	12%	1	17	6%	0	2	0%	2	3	67%
Michigan	4	23	17%	2	15	13%	1	3	33%	1	5	20%
Harvard	5	19	26%	5	13	39%	0	1	0%	0	5	0%
MIT	4	16	25%	3	11	27%	0	3	0%	1	2	50%
Yale	8	24	33%	5	17	29%	1	1	100%	2	6	33%
Stanford	5	23	22%	2	14	14%	3	8	37%	0	1	0%
Berkeley	5	21	24%	3	13	23%	1	5	20%	1	3	33%
UCLA	3	18	17%	2	12	17%	1	4	25%	0	2	0%
UNC	3	19	16%	3	13	23%	0	4	0%	0	2	0%
Columbia	7	26	27%	5	20	25%	0	1	0%	2	5	40%
Arizona	6	23	26%	1	11	9%	5	11	45%	0	1	0%
CUNY	12	48	25%	10	42	24%	2	6	33%	0	0	NA
Notre Dame	6	39	15%	2	24	8%	2	8	25%	2	7	29%
Brown	4	17	24%	3	12	25%	0	1	0%	1	4	25%
Cornell	4	15	27%	1	8	13%	2	5	40%	1	2	50%
USC	3	21	14%	1	13	8%	2	6	33%	0	2	0%
UT-Austin	3	29	10%	2	21	10%	0	4	0%	1	4	25%
U.S.A. Total	*99*	*489*	*19.8%*	*55*	*337*	*16.3%*	*24*	*87*	*27.6%*	*20*	*65*	*30.7%*

For a different kind of data, we consulted the *PhilPapers* internet survey of philosophers conducted by David Chalmers, David Bourget and associates in November 2009.[6] The survey was taken by 3226 respondents, including 1803 philosophy faculty members and/or PhDs and 829 philosophy graduate students. Though respondents were based in over 29 countries, about 79% of those who indicated a "country of primary affiliation" were based at institutions in the USA, the U.K., Australia, Canada and New Zealand. Among the respondents, 3013 specified their gender. There were 2525 males and 488 females. So in the *PhilPapers* survey, 16.2% of respondents who indicated their gender were women. Though there is room for debate about which of these measures is most informative or most useful, the general picture is quite clear. Women are seriously under-represented in academic philosophy.

[6] Available at http://philpapers.org/surveys/.

2. What Are Philosophical Intuitions and What Role Do They Play in Philosophical Argument ?

There is a lively debate in the philosophical literature on the use of intuitions as evidence in philosophy (Bealer 1998, Devitt 2006 & 2009, Gendler 2007, Goldman 2007, Jackson 1998, Kornblith 1998, Ludwig 2007, Sosa 2007 & 2009, Stich 2009, Weinberg, Nichols & Stich 2001, Weinberg 2007, Williamson 2004 & 2007). In this debate there are, broadly speaking, two families of views about how intuitions should be characterized. One family uses the terms 'intuition' and 'philosophical intuition' quite inclusively. For authors who adopt this view, intuitions are whatever contemporary philosophers have in mind when they use the term 'intuition' reflectively. Writers in the other family propose much narrower definitions, which require that philosophical intuitions have quite restrictive phenomenological, epistemic or psychological properties. The second family is significantly more quarrelsome than the first, and the definitions proposed often pick out what appear to be quite different classes of psychological phenomena. In this paper, we will adopt the terminological strategy endorsed by the first family; we will use the terms 'intuition' and 'philosophical intuition' quite liberally. On this terminological issue, we share the view expressed by Timothy Williamson: "Although we could decide to restrict the term 'intuition' to states with some list of psychological or epistemological features, such a stipulation would not explain the more promiscuous role the term plays in the practice of philosophy." (Williamson 2007, 218)

Appeal to intuition has played a central role in Anglo-American philosophy over the last 50 years. In a typical episode, a philosopher will describe a real or (more commonly) an imaginary situation and ask whether some of the people or objects or events in the situation described exhibit some philosophically interesting property or relation:

- Is the action described *morally wrong*?
- Does the person described *know* that he won't win the lottery?
- When the speaker in the story uses the word 'water' does the word *refer* to H_2O?
- Does the "Chinese Room" really *understand* the story?

When things go well, both the philosopher and his audience will agree that the answer is intuitively obvious, and that will be taken to be *evidence* for or against some philosophical thesis. Readers acquainted with contemporary "analytic" philosophy will have encountered many examples of this practice; readers who are unfamiliar with this literature will find lots of examples in section 3. However, it is worth noting that while using the term 'intuition' as a label for people's spontaneous responses to philosophical thought experiments is a relatively new phenomenon,[7] the practice itself goes all the way back to the beginnings of Western philosophy. In a famous passage in Plato's *Republic*, Cephalus proposes an account of justice on which what justice requires is speaking the truth and paying one's debts. Socrates responds with a thought

[7] Hintikka (1999) notes that the use of the term 'intuition' in contemporary philosophy became much more common as philosophers became acquainted with Chomsky's work in linguistics where 'intuition' is used as a label for spontaneous judgments about the grammatical properties of sentences that speakers are asked to consider.

experiment and a question: "Suppose that a friend when in his right mind has deposited arms with me and he asks for them when he is not in his right mind, ought I to give them back to him?" He then proceeds to answer his own question: "No one would say that I ought to or that I should be right in doing so, any more than they would say that I ought always to speak the truth to one who is in his condition."[8] Cephalus agrees. In contemporary terminology, Socrates' thought experiment is designed to elicit the intuition that returning the weapons and speaking the truth is not morally required in this situation, and it succeeds admirably both for Cephalus and for many modern readers.

In many traditional and contemporary philosophical projects, when an intuition is invoked it is assumed that the *propositional content* of the intuition is likely to be true, and thus that the proposition can be used as evidence. Philosophical theories that are compatible with the content of the intuition are supported, and philosophical theories that are incompatible with the content of the intuition are challenged. So, for example, if we have the intuition that a character in a Gettier thought experiment does not know that p (where p is a true proposition that he believes and is justified in believing), this is evidence against the justified-true-belief account of knowledge (Gettier 1963; for an example see section 3). Similarly, in a Magistrate and the Mob thought experiment (Smart 1973; for an example see section 3), if we have the intuition that it would be morally impermissible for the Chief of Police to frame an innocent man to stop a riot, that is evidence against some versions of act utilitarianism. Though a few recent writers have denied that philosophical intuitions are used as evidence in this way (Ichikawa 2008, Deutsch 2010, Cappelen forthcoming), they are very much in the minority. Far more common is the view voiced by Kripke:

> Some philosophers think that something's having intuitive content is very inconclusive evidence in favor of it. I think it is very heavy evidence in favor of anything, myself. I really don't know, in a way, what more conclusive evidence one can have about anything, ultimately speaking. (Kripke 1980, p. 42)

In contemporary philosophy, the explicitly stated goal of many projects is conceptual analysis. In these projects, intuitions play a somewhat different role. Rather than assuming that the content of the intuition is likely to be true, it is typically assumed that intuitions are evidence about the nature of the concept being analyzed. One version of this assumption has been stated very clearly by Alvin Goldman.

> [T]he evidential status of application intuitions is of the constitutively grounded variety.[9] It's part of the nature of concepts (in the personal psychological sense) that possessing a concept tends to give rise to beliefs and intuitions that accord with the contents of the concept. If the content of someone's concept F implies that F does (doesn't) apply to example x, then that person is disposed to intuit that F applies (doesn't apply) to x when the issue is raised in his mind. (Goldman 2007, p. 4)

---

[8] Plato (1892), I, 331, p. 595.

[9] "Application intuitions" is Goldman's term for "intuitions about how cases are to be classified".

Though there is much more to be said about the ways in which intuitions are used as evidence in philosophy, for our purposes this brief discussion should suffice. The bottom line is that in philosophy intuitions are often taken to be evidence relevant either to the truth or falsity of a philosophical theory that purports to characterize some philosophically important phenomenon (like knowledge or reference or moral permissibility), or to an account of some philosophically important concept.[10]

3. Some Evidence for Gender Differences in Philosophical Intuition

We first became interested in gender differences in philosophical intuition after learning of a study in which Christina Starmans and Ori Friedman (2009) found dramatic differences in the intuitions reported by undergraduate men and women on Gettier-style thought experiments. Intrigued by these data, we set out to determine whether they were isolated results or whether there are other gender differences in philosophical intuition. However, this is a challenging task because researchers who study philosophical intuitions often do not collect or analyze demographic variables such as gender. So our strategy was threefold. First, we contacted a number of researchers who had done work on philosophical intuition and asked whether they were aware of published or unpublished data that could throw light on the existence and scope of gender differences in response to philosophical thought experiments. In addition to sending us references and copies of papers in press, several of the colleagues we contacted reported that while they had collected gender data, they had not bothered to analyze it. At our request, they agreed to do so. Second, Buckwalter reanalyzed the data of an earlier study he had conducted with James Beebe to see if there was a gender effect that they had not noticed, and ran an additional follow-up experiment in a neighboring domain. Third, we launched a series of new studies to determine whether there are gender differences in intuitive responses to some well known philosophical thought experiments for which gender data had not previously been collected. We will report our findings in the order just recounted.

3.1 Starmans & Friedman: Gettier Cases

Starmans and Friedman presented undergraduate participants at the University of Waterloo in Canada with a number of pen and paper Gettier style thought experiments. In one study, participants ($N = 140$; 84 men, 56 women) read the following vignette.

> Peter is in his locked apartment, and is reading. He decides to have a shower. He puts his book down on the coffee table. Then he takes off his watch, and also puts it on the coffee table. Then he goes into the bathroom. As Peter's shower begins, a burglar silently breaks into Peter's apartment. The burglar takes Peter's watch, puts a cheap plastic watch in its place, and then leaves. Peter has only been in the shower for two minutes, and he did not hear anything.

[10] Though intuitions clearly play an important role in contemporary philosophy, they are *not* the only source of evidence for philosophical theories, and some contemporary philosophers make little or no use of them.

After reading the text, participants were presented with three comprehension check questions to determine whether they understood and recalled the details of the story. They were then asked the test question:

> ***Does Peter really know that there is a watch on the table, or does he only believe it?***

where they could respond with either 'really knows' or 'only believes'. To provide a basis for comparison, a control group of participants were given the same vignette, but were instead asked the test question about the book rather than the watch. Though there were no significant gender differences when participants were asked whether Peter knows that there is a book on the table (a large majority of both genders said that he really knows) there was a huge gender difference when participants were asked about the watch. Only 41% of the male participants said that Peter really knows that there is a watch on the table, while 71% of the female participants said that Peter really knows ($p < 0.05$, Fisher's exact test). Concerned that the gender of the protagonist might be playing a role in generating these results, Starmans and Friedman ran another study using a slightly different vignette in which the central protagonist was female. In this version, the objects involved were a wedding ring and a fork. Once again, the results were striking: 36% of male participants said the female protagonist really knows in the Gettier condition, while 75% of the female participants said that she really knows ($N = 112$, 54 men, 58 women, $p < 0.01$, Fisher's exact test). These results indicate that, at least in some groups of undergraduates, intuitive responses to Gettier cases can be highly variable and that men and women students can respond very differently to some prototypical Gettier-style thought experiments.

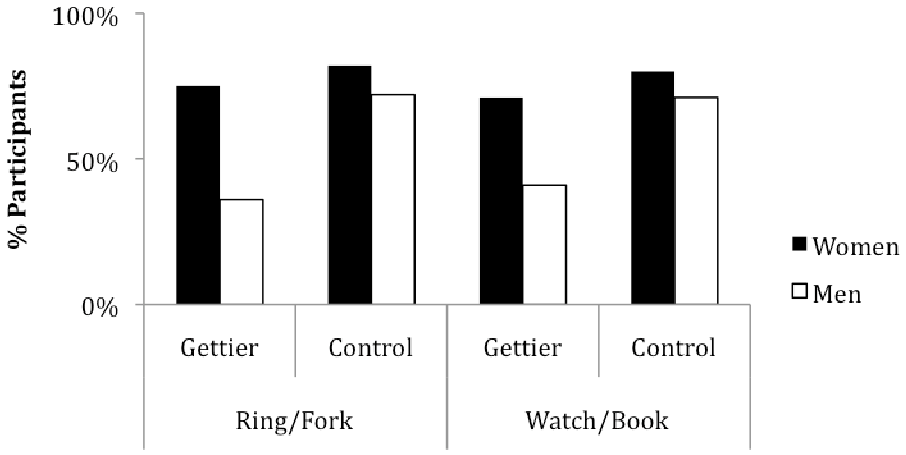

Figure 1. Percent of Participants Attributing Knowledge in Gettier Study.

Since their original study, Starmans and Friedman (personal communication) have conducted a number of additional experiments using different Gettier vignettes and different populations of participants, including people recruited in public places in New York City. Among that latter group, they found that roughly half of both male and female participants

attributed knowledge to the protagonists in their Gettier vignettes. We are currently conducting a series of further studies of Gettier intuitions in both student populations and in the general public. Preliminary results show a bewildering pattern in which some vignettes evoke gender differences in some groups and some do not. It is clear that there is still a lot to learn.

3.2: Holtzman: Compatibilism, Physicalism & Dualism Cases

When we notified colleagues that we were interested in gender differences, the first researcher we heard from was Geoffrey Holtzman who had been conducting a series of online studies of philosophical thought experiments (Holtzman, in preparation a, b).[11] In one of these studies, he elicited participants' intuitions ($N = 192$; 102 men, 90 women) about the following compatibilism thought experiment:

> Suppose Scientists figure out the exact state of the universe during the Big Bang, and figure out all the laws of physics as well. They put this information into a computer, and the computer perfectly predicts everything that has ever happened. In other words, they prove that everything that happens, has to happen exactly that way because of the laws of physics and everything that's come before. ***In this case, is a person free to choose whether or not to murder someone?***[12]

As shown in Figure 2, Holtzman found a very substantial difference between the responses of men and women participants. While 63% of women responded that a person in the situation described is free to choose whether or not to murder someone, only 35% of men gave that response ($d = 0.58$, $p < 0.0005$, Fisher's exact test, all experiments two-tailed).

[11] Holtzman used SurveyMonkey for these experiments, a commercially available web-based survey tool.

[12] In all of the Holtzman studies we discuss, participants were asked to reply either 'Yes' or 'No'. Only participants who reported that they had taken no philosophy courses were considered.

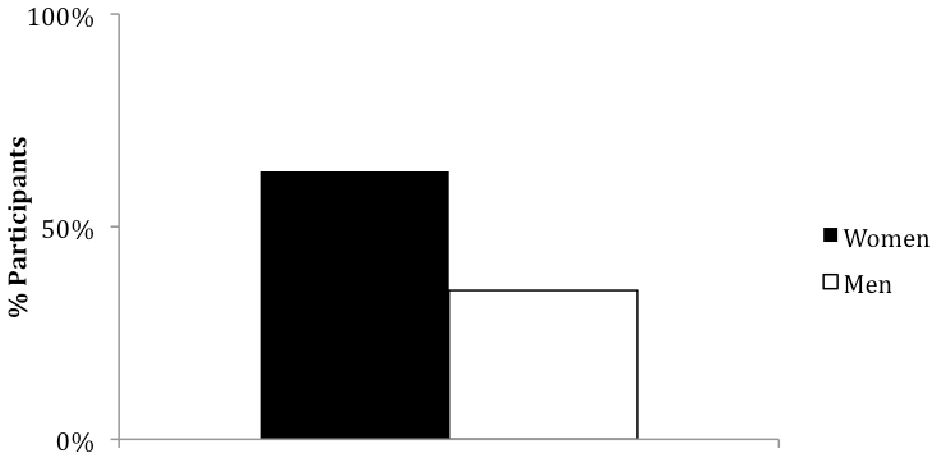

Figure 2. Percent Answering "Yes" in Compatibilism Study.

In a second study (N = 195; 93 women, 102 men), Holtzman asked participants to read an intuition probe inspired by Frank Jackson's famous "Mary" case (Jackson 1982 & 1986).

> Suppose you meet a man from the future who knows everything there is to know about science. He tells you that he doesn't like apples, and says that though he has never eaten one, he has figured out what apples taste like just by studying the relevant science. ***Could he know what apples taste like without ever having eaten one?***

Among male participants, 39% said "yes", but among female participants, only 17% said "yes" (d = 0.50, p < 0.005, Fisher's exact test).

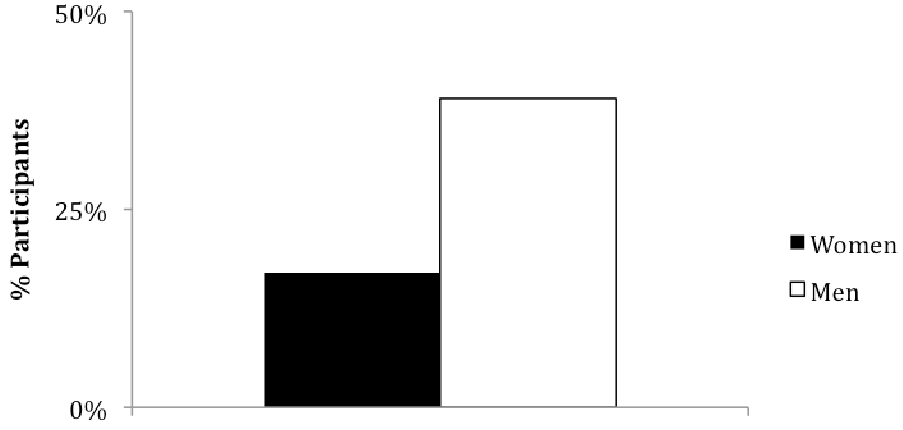

Figure 3. Percent Answering "Yes" in Physicalism Study.

Participants in a third study (N = 185; 87 women; 98 men;), were presented with a version of a well known "dualism" thought experiment.

> Suppose neurologists are able to identify every part and every connection in the human brain. Working with a team of computer scientists, they then build a robot that has a complete electronic replica of the human brain. ***Could this robot experience love?***

Here Holtzman found a smaller, but significant difference. Among male participants, 79% said "yes", among female participants, 62% said "yes" (d = 0.37, p = 0.016, Fisher's exact test).

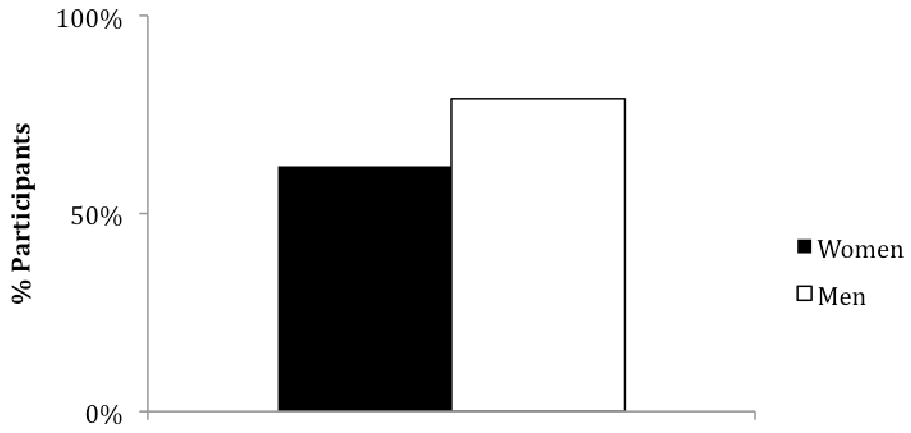

Figure 4. Percent Answering "Yes" in Dualism Study.

It is important to note that Holtzman also collected data on participants' intuitions about a number of other philosophical thought experiments and found no significant gender differences.[13]

3.3. Cushman: The Violinist and The Magistrate & the Mob

Fiery Cushman was one of the researchers who agreed to look for gender effects in data he had collected online in collaboration with Liane Young. One study in which he found them used a version of one of contemporary philosophy's most famous thought experiments, the "Violinist" case first introduced into the literature by Judith Jarvis Thomson (1971) in a widely

[13] In Holtzman's studies, the three cases we have discussed were presented along with six other vignettes for which no significant gender differences were found. See Holtzman (under review) for details on these six vignettes. With the exception of the dualism study, significance values in the experiments we have recounted remain at the $p < 0.05$ level after correcting by a factor of 9.

discussed paper on abortion. In this experiment, Cushman and Young presented participants (*N* = 298; 176 men, 122 women) with the following vignette:

> Jill wakes up one morning and finds a strange man next to her in bed, plugged into her kidney. A man from the Society for Music Lovers introduces himself and explains to Jill that she has been plugged into a famous violinist who is dying of kidney failure. Without Jill's help, the violinist will die. The man from the Society explains that Jill must stay plugged into the violinist for 9 months in order for him to recover and survive. Jill pulls the plug and the violinist dies.
>
> ***Jill's pulling the plug was:***

Participants were asked to respond on a scale from 1 to 7, with 1 labeled "Forbidden," 4 labeled "Permissible," and 7 labeled "Obligatory." What Cushman found was that men were more likely to say that Jill's pulling the plug was permissible, while women tended to consider this action forbidden.[14] The surprising results are depicted in Figure 5.

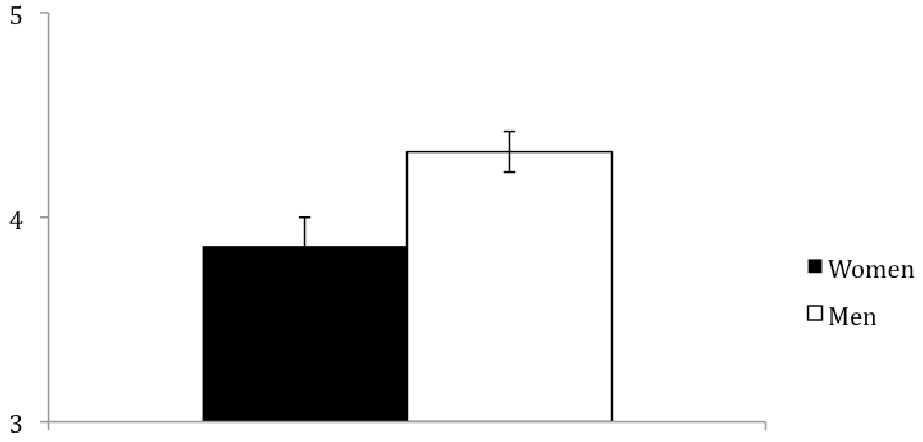

Figure 5. Mean Judgment in Violinist Study (the scale ran from 1 to 7).

A second case in which Cushman found a significant gender effect was a version of the "Magistrate and the Mob" thought experiment made prominent by Smart (1973). Participants (*N* = 529; 380 men, 149 women) read the following:

> Steve is the police chief of a large city. A particularly volatile political situation has erupted into violence with the assassination of a candidate, and citizens are rioting in the streets, demanding an arrest be made. The situation will worsen unless Steve can produce a perpetrator; people have already looted numerous

[14] For male participants the mean was 4.32, *SD* = 1.39; for female participants the mean was 3.86, *SD* = 1.57, (*d* = 0.31). An independent samples t-test reveals a significant difference between these two groups, $t(296) = 2.65$, $p < 0.01$.

stores, overturned cars, brutally attacked bystanders, and set fire to a government building. The police department has Steve's 14-year-old brother in custody for petty theft and drunkenness. He could easily fabricate sufficient evidence against his young brother to satisfy the public and stop the riots. Steve chooses not to frame his teenage brother and the violent attacks escalate.

The choice Steve made was:

Participants were asked to respond with a horizontal sliding bar, where the leftmost side was anchored with "Good" and the rightmost side was anchored with "Bad." Numerical scores were then assigned to those responses on a scale from -225 to +225. This means that lower scores indicate that participants thought the action was good, while higher scores indicate that participants thought the action was bad. As can be seen in Figure 6, while both groups tended to think that Steve's choice was a good one, the results suggest that men judged it was better for Steve *not* to have framed his teenage brother than women did.[15]

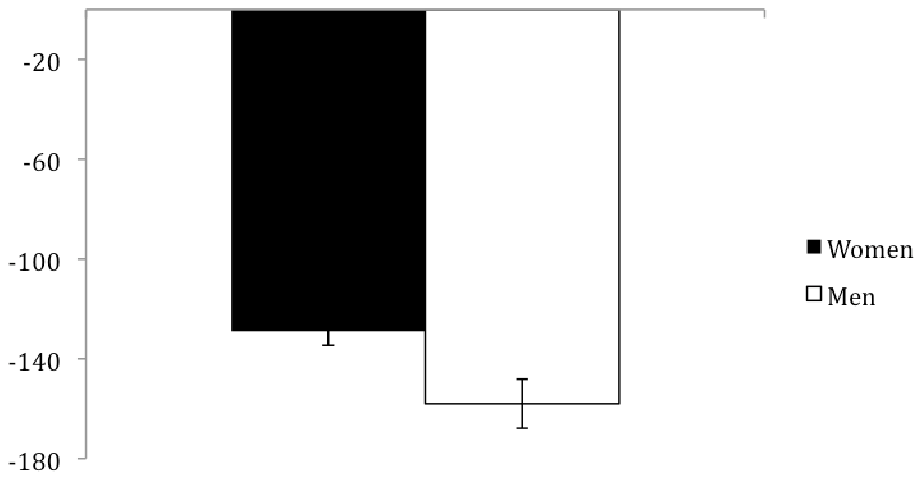

Figure 6. Mean Judgment in Magistrate and Mob Study (the scale ran from -225 to +225).

Like Holtzman, Cushman also reported that there were a number of cases in which gender differences were not found.

3.4. Zamzow and Nichols: A Trolley Case

A third colleague who responded to our inquiry was Shaun Nichols who provided a copy of a recently published paper co-authored with Jennifer Zamzow (Zamzow & Nichols, 2009). The paper includes a report of a pen and paper study in which undergraduate participants were presented with a bystander version of the trolley dilemma in which five people can be saved if a

[15] For male participants the mean was -158, $SD = 120.39$; and for female participants the mean was -129, $SD = 108.36$, ($d = 0.25$). A significant main effect was obtained for gender, $F(1,521) = 7.40$, $p < 0.01$.

train is diverted to a sidetrack where it will kill a different innocent person. Some participants were told that the individual on the side track was "a stranger," others were told that it was "a 12 year old boy". The text of their vignettes were as follows:

> You are taking your daily walk near the train tracks and you notice that the train that is approaching is out of control. You see what has happened: the driver of the train saw five people working on the tracks and slammed on the brakes, but the brakes failed and the driver fainted. The train is now rushing toward the five people. It is moving so fast that they will not be able to get off the track in time. You happen to be standing next to a switch, and you realize that the only way to save the five people on the tracks is to throw the switch, which will turn the train onto a side track, thereby preventing it from killing the five people. However, there is a stranger [or, in the other vignette: a 12-year-old boy] standing on the side track with his back turned, and if you proceed to throw the switch, the five people will be saved, but the person [boy] on the sidetrack will be killed.

Participants, each of whom saw only one version of the vignette, were asked the extent to which they agreed with several statements including: ***"It is morally acceptable for me to pull the switch."*** Responses were collected on a 7 point scale with 1 labeled "strongly disagree" and 7 labeled "strongly agree". Zamzow and Nichols found that men judged that the killing of a stranger is *less* morally acceptable than did women.[16] However, when the person on the side track was described as a 12 year old boy, this pattern was reversed. In that case, men tended to judge that the killing was *more* morally acceptable than women did.[17] Zamzow and Nichols also explored participants' judgments about trolley cases in which the person on the side track is described as either a brother or a sister of the participant. They found that men judged killing one's brother to be less morally acceptable than did women, but that women judged that killing one's sister is less morally acceptable than did men.[18]

---

[16] In the stranger case, the mean response among male participants was 4.21, $SD = 1.93$, and the mean among female participants was 4.95, $SD = 1.07$, ($d = 0.50$).

[17] In the 12 year old boy case, the mean response for male participants was 4.87, $SD = 1.71$, and the mean for female participants was 4.26, $SD = 1.79$, ($d = 0.35$). A two-way between-subjects analysis of variance was conducted to evaluate the effect of condition (either stranger or 12 year old boy) and gender on participant responses. The interaction of these two factors approached significance $F(1, 85) = 3.46$, $p = 0.07$.

[18] In the killing your brother case, the mean judgment for male participants was 3.41, $SD = 1.67$, and the mean for female participants was 4.33, $SD = 1.35$, ($d = 0.59$). In the killing your sister case, mean judgments for male participants was 4.40, $SD = 2.13$, and the mean for female participants was 3.78, $SD = 1.58$, ($d = 0.33$). A two-way between-subjects analysis of variance reveals a significant interaction effect between these two factors $F(1, 95) = 4.45$, $p < 0.05$.

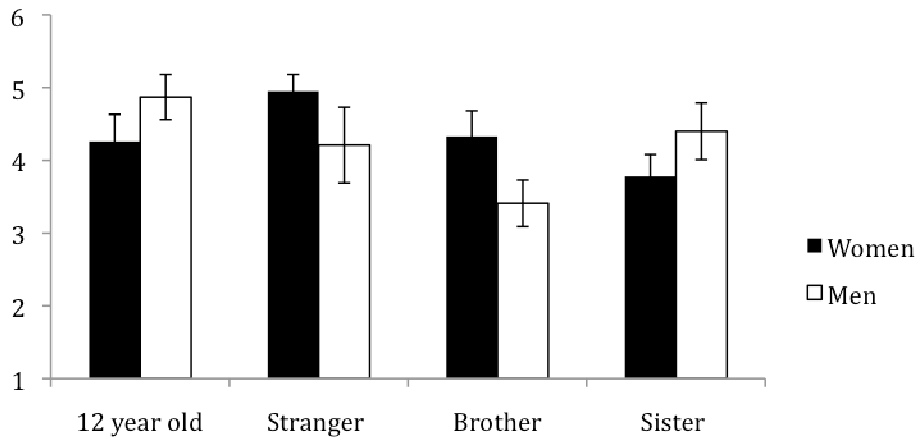

Figure 7. Mean Agreement in Trolley Study (the scale ran from 1 to 7).

Interestingly, Zamzow and Nichols also asked participants what they *would do*, when confronting this dilemma and found *no* significant difference between males and females.

3.5. Pizarro, Uhlmann & Bloom: Moral Responsibility and Causal Deviance

Yet another colleague who responded to our inquiry about gender differences was David Pizarro, who reported that on reanalyzing some of the data that he and colleagues had collected (Pizarro et al. 2003) he found an intriguing gender difference that they had not previously noticed. To understand the motivation for this experiment, a bit of background is needed. In typical cases of intentional action, an agent forms the intention to accomplish a goal, comes up with a plan for doing it, and then carries out that plan. But many philosophers have been intrigued by another sort of case, where the agent forms the intention to bring about a goal, and does bring it about, but not in the way the agent planned. Rather, in these cases, the agent's intention causes a chain of events quite different from those the agent had planned. Cases of this sort are sometimes said to exhibit "causal deviance" (Searle, 1983). Pizarro and colleagues were interested in whether people's intuitions about the *moral* properties of an action would be affected if the action were causally deviant. To find out, participants in this experiment read each of the four vignettes below, counterbalanced for order. The first two describe a morally positive action, the second two a morally negative action. In each pair, the link between the intention and the action is normal in the first vignette and "deviant" in the second.

> *Positive Normal*: Tom is walking in the park when he sees a man choking on a sandwich. Tom, full of nervousness because of his intention to save the man's life, runs over and performs the Heimlich maneuver. The man coughs up the sandwich he had been choking on and his life is saved.
>
> *Positive Deviant*: Tom is walking in the park when he sees a man choking on a sandwich. Tom, full of nervousness because of his intention to save the man's life, runs over to perform the Heimlich maneuver. However, Tom's nervousness leads him to have an epileptic seizure. (Had Tom not had the epileptic seizure, he would have carried out the

Heimlich maneuver and saved the choking man's life.) By chance, the epileptic fit happens to lead Tom's arms to squeeze on the man's chest, causing the man to cough up the sandwich and saving his life.

Negative Normal: Tom lies in wait for his enemy, who had stolen his life savings. As soon as his enemy appears, Tom, nervous because of his intention to kill the man, pulls out a gun and shoots his enemy dead.

Negative Deviant: Tom lies in wait for his enemy, who had stolen his life savings. As soon as his enemy appears, Tom, nervous because of his intention to kill the man, pulls out a gun. However, Tom's nervousness triggers an epileptic seizure. (Had Tom not had an epileptic seizure, he would have shot his enemy dead.) By chance, the epileptic fit leads Tom to squeeze the trigger, and the bullet happens to hit and kill his enemy.

After each vignette, participants were asked to judge (1) how moral or immoral the agent was, (2) how much blame or praise the agent should receive for his actions, and (3) how positively or negatively the agent should be judged. Responses were assessed on a nine-point scale ranging from -4 to +4. In the blame question, -4 was labeled *extreme blame*, 0 was labeled *neither blame nor praise*, and +4 was labeled *extreme praise*. The labels for the other questions were similar. Scores for the negative actions were multiplied by -1. Thus while higher numbers in the positive cases mean that participants judged those agents to be more praiseworthy, etc., higher numbers in the negative cases mean that participants judged the characters more blameworthy. Since within-subjects responses to the three questions that were asked about each of the four cases were highly correlated, they were averaged to create a "moral sanction index."

As Pizarro and colleagues had predicted, in both the positive and the negative cases, responses were less extreme in the deviant cases than in the normal cases. Participants thought that the action in the Positive Normal case was better than the action in the Positive Deviant case, and they thought the action in the Negative Normal case was worse than the action in the Negative Deviant case.[19] What Pizarro and colleagues did not predict, and had not noticed until we asked about gender effects, was that "the discounting of blame for the Negative Deviant act is driven by the women, and the discounting of praise for the Positive Deviant act is driven by the men" (Pizarro, personal communication).[20] The reanalyzed data that Pizarro provided are represented in Figure 8. (Recall that higher scores in negative cases represent a higher

---

[19] Pizarro et al (2003) provide the following technical details. "In order to test the hypothesis that individuals discounted responsibility for causally deviant actions, a 2 (causal condition: deviant vs. normal) x 2 (positive vs. negative act), repeated measures ANOVA was conducted. As predicted, there was a main effect for experimental condition, $F(1, 25) = 18.13$, $p < 0.001$, such that individuals discounted moral responsibility for acts that were causally 'deviant.'"

[20] For women: Negative Deviant ($M = 2.67$, $SD = 1.00$), Negative Normal ($M = 3.17$, $SD = 1.05$); paired-sample t-test, $t(13) = -2.88$, $p = 0.01$. Positive Normal ($M = 2.93$, $SD = 0.80$), Positive Deviant ($M = 3.10$, $SD = 0.85$); $t(13) = -1.53$, $p = 0.15$. For men: Negative Deviant ($M = 2.58$, $SD = 1.19$), Negative Normal ($M = 2.86$, $SD = 1.23$); $t(11) = -1.39$, $p = 0.19$. Positive Normal ($M = 2.94$, $SD = 1.01$), Positive Deviant ($M = 2.39$, $SD = 0.86$); $t(11) = -3.35$, $p = 0.01$.

blameworthiness index, while higher scores in positive cases represent a higher praiseworthy index.)

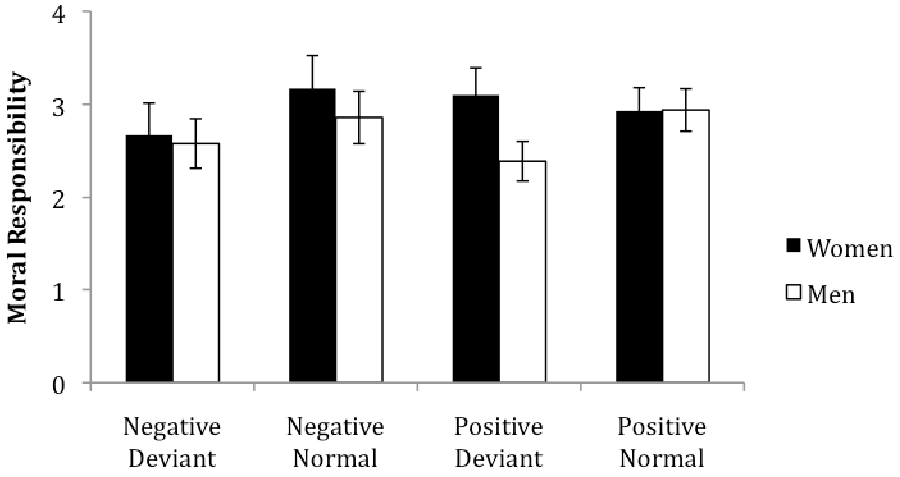

Figure 8. Moral Sanction Indexes in Causal Deviance Study.

3.6. Beebe & Buckwalter: The Epistemic Side-Effect Effect

Intrigued by some of the findings we have been recounting, Buckwalter decided to re-analyze the data he and James Beebe had collected, which had demonstrated a surprising extension of the side-effect effect discovered by Joshua Knobe (Beebe and Buckwalter 2010). While Knobe's original studies focus on attributions of intention, Beebe and Buckwalter's study explores attributions of knowledge. Knobe (2003) constructed a pair of vignettes which described the circumstances of a chairman of a company receiving some news from a subordinate about a new program they could adopt to increase profits. However, instituting this program has a side effect: in one case it will harm the environment and in another case it will help the environment. In both cases, the chairman does not care about any possible side effect to the environment; his only concern is profit. So in both cases the chairman tells his subordinate to start the new program, and in each case the environment is affected as predicted. Knobe's surprising finding was that when asked about the actions of the chairman in the harm case, 82% of participants agreed that the chairman intentionally harmed the environment, whereas in the help case 77% of subjects denied that the chairman intentionally helped the environment. The result suggests that people are more likely to say that an agent brought about a side-effect *intentionally* when that side-effect is bad than when it is good. This asymmetry in participant responses has been widely replicated (Knobe 2010).

Beebe and Buckwalter wanted to know whether moral considerations like those invoked in Knobe's studies affect people's intuitions about whether people have knowledge. To investigate this question, they conducted a pen and paper experiment to see if the original side-effect effect asymmetry persisted when participants were asked about what the chairman knows. Their hypothesis was that given the same evidence in these cases, participants would be less likely to say an agent knows that an action will bring about a specified side effect when the side

effect is good, and more likely to attribute knowledge when the side effect is bad. Using a between-subjects experimental design, undergraduate participants ($N = 749$) were given either the help or harm version of the following vignette:

> The vice-president of a company went to the chairman of the board and said, "We are thinking of starting a new program. We are sure that it will help us increase profits, and it will also (*help / harm*) the environment." The chairman of the board answered, "I don't care at all about (*helping / harming*) the environment. I just want to make as much profit as I can. Let's start the new program." They started the new program. Sure enough, the environment was (*helped / harmed*).

On a seven-point scale from -3 to 3 (where -3 was anchored with "the chairman didn't know" and 3 was anchored with "the chairman knew"), participants were asked, "***Did the chairman know that the new program would (help/harm) the environment***". The result was that the degree to which participants attribute knowledge to the chairman was significantly different between conditions.[21] Even though there is equally strong evidence in both help and harm conditions that the chairman's action would bring about a certain side-effect, participants were significantly less likely to agree that the chairman knows an action will bring about that side effect when the effect is good, and more likely to attribute that knowledge when the side-effect is bad. Beebe and Buckwalter call this asymmetry between knowledge attribution in help and harm conditions "the epistemic side-effect effect".[22]

One might think that the interesting thing about the epistemic side-effect effect finding is not that knowledge attribution is close to ceiling in the harm condition, but rather that participants are much less willing to attribute knowledge in the help condition. However, when Buckwalter re-analyzed the data he found that that men (405 participants) and women (340 participants) answer differently in the help condition, and that much of the difference between the harm and the help condition can be attributed to this gender difference. The mean differences between the way men and women responded between conditions is represented in Figure 9; women were more likely than men to say that an agent *does not know* an action will bring about the side effect when that effect is good.[23] This study of the epistemic side-effect effect reveals another significant gender effect in epistemic intuitions.

---

[21] In the help condition the mean response was 0.91, $SD = 2.09$, in the harm condition the mean was 2.25, $SD = 1.50$. An independent samples t-test reveals a significant difference between these two groups, $t(747) = -10.126$, $p < 0.001$.

[22] For more on the epistemic side-effect effect, see Schaffer & Knobe (forthcoming) and Beebe & Jensen (forthcoming).

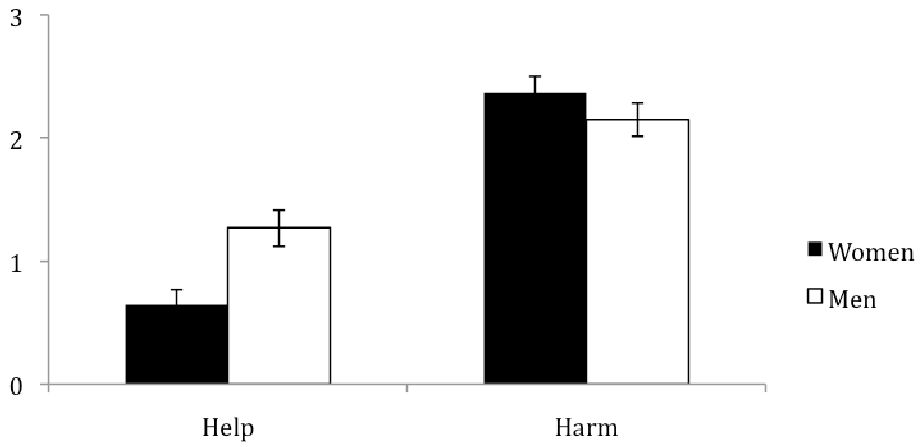

Figure 9. Mean Judgment in Epistemic Side-Effect Effect Study (the scale ran from -3 to +3).

3.7: Buckwalter: Absence Causation

Intuitions about intention and knowledge are not the only ones that exhibit Knobe's side-effect effect. Similar results have been reported for a wide range of intuitions, including intuitions about causation (Hitchcock & Knobe 2009; Knobe & Fraser 2008; for reviews see Knobe 2010 and Knobe et al. forthcoming). However, none of this work has tried to determine whether the gender effect on knowledge judgments recounted in the previous section generalizes to other sorts of intuitions. To explore the issue, Buckwalter designed an experiment focused on people's intuitions about cases in which an agent's failure to act facilitates either a positively valenced or a negatively valenced outcome (Buckwalter manuscript a, b). Cases like this are sometimes described as examples of "absence causation".

In a between subjects 2 x 2 experimental design, each participant (N = 415; 251 female, 160 male, 4 unreported) received one of four vignettes that are similar in structure to Knobe's chairman cases. The vignettes differed in a pair of ways. First, two of them make it explicit that the protagonist's failure to act is intentional, while in the other two the protagonist's failure to act is not intentional. Second, in two of the vignettes the outcome is positively valenced, while in the other two the outcome is negatively valenced. These are the two vignettes in which the protagonist's failure to act is intentional:

> *Intentional-Help*: P&G has just purchased all of the utility companies in a small, rural community. The previous owner neglected the power plants, and many of the pipes are leaking a certain chemical into the local town's water supply. This chemical is full of natural nutrients for plants, and will surely double the harvest. *When the president of P&G found out about the leaks he thought to himself, "This is great! If I don't fix the leaks, all of the local crops will flourish." In hopes that it would help the crops, he decides to do nothing.* Sure enough, the crops flourished and the harvest was doubled.

Intentional-Harm: P&G has just purchased all of the utility companies in a small, rural community. The previous owner neglected the power plants, and many of the pipes are leaking a certain chemical into the local town's water supply. This chemical is toxic to plants, and will surely cut the town's harvest in half. *When the president of the company found out about the leaks he thought to himself, "This is great! If I don't fix the leaks, all of the townspeople's crops die, and they will have no choice but to buy all their food from us. Think of all those profits." In hopes that it would harm the crops, he decides to do nothing.* Sure enough, the crops died, and the townspeople lost half of the harvest.

In the other two vignettes, *Unintentional-Help* and *Unintentional-Harm*, the information about the president's mental states is omitted, and the italicized text is replaced by the following sentence: *The president of P&G had no idea about the leaks, and no idea about how the chemical might affect the townspeople's crops.* After reading the vignettes, participants were asked a pair of questions. The first question was:

How much (*praise/blame*) does the president of P&G deserve?

Participants responded on a seven-point scale anchored by *Not At All Blameworthy/Praiseworthy* and *Very Blameworthy/Praiseworthy*. In the second question, they were asked to indicate their level of agreement with the following statement:

By not fixing the pipes, the president of P&G is a cause of the (improved/diminished) harvest.

Participants responded on a seven-point scale, anchored by *Strongly Disagree* and *Strongly Agree*.

Not surprisingly, participants judged that the president deserved a lot of blame in the harm conditions, but deserved much less praise in the help conditions. They also attributed more praise or blame in the intentional conditions than in the unintentional conditions.[24] The responses to the causation question exhibited a similar pattern. Participants were much more likely to treat the president as a cause when the outcome was bad than when it was good, and when it was intentional rather than unintentional.[25] But what makes this study important for our purposes is that there is a significant gender difference in responses to the causation question. Women were much more likely than men to agree that the president was a *cause* of the outcome

[24] A two-way between-subjects analysis of variance was conducted to evaluate the effect of intentionality and valence. A significant main effect was obtained for intentionality, $F(1, 410) = 63.92$, $p < 0.001$. The main effect for valence was also significant, $F(1, 410) = 230.86$, $p < 0.001$.

[25] Means and standard deviations for *causal agreement*: Intentional-Help ($M = 3.89$, $SD = 1.88$), Intentional-Harm ($M = 5.50$, $SD = 1.59$), Unintentional-Help ($M = 3.51$, $SD = 1.97$), Unintentional-Harm ($M = 4.50$, $SD = 1.71$). A significant main effect was obtained for intentionality, $F(1, 411) = 15.17$, $p < 0.001$. The main effect for valence was also significant, $F(1, 411) = 54.83$, $p < 0.001$.

when that outcome was bad, and more likely to disagree that the president was a cause when the outcome was good.[26] The results, displayed in Figure 10, are remarkably similar to the pattern of results in the epistemic side-effect effect study displayed in Figure 9.[27]

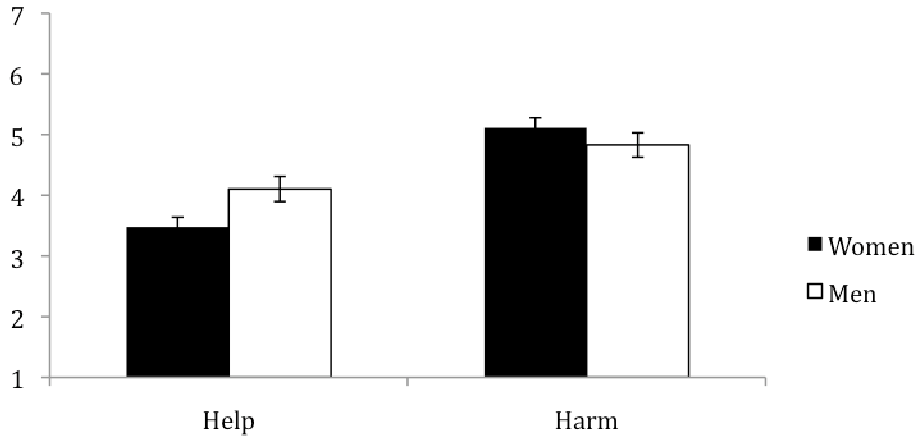

Figure 10. Mean Agreement in Absence Causation Study (the scale ran from 1 to 7).

3.8: Buckwalter & Stich: The Brain in the Vat, Twin Earth, The Chinese Room and The Plank

We have now completed a series of experimental studies looking at other familiar philosophical thought experiments. This research was conducted on the internet using Amazon's Mechanical Turk, recruiting participants from across the United States.[28] In all of the studies we discuss in this section, participants were presented with the text of a thought experiment, asked a comprehension check question and a test question, and then were asked to fill out a short, eight-item demographic questionnaire.[29] Importantly, one of the items included on the demographic

---

[26] Women (Help $M = 3.48$, $SD = 1.87$, Harm $M = 5.11$, $SD = 1.67$) and men (Help $M = 4.1$, $SD = 1.98$, Harm $M = 4.83$, $SD = 1.79$). A two-way between-subjects analysis of variance was conducted to evaluate the relationship between gender and valence on participant responses in the causation cases. A significant interaction was found between these factors $F(1, 407) = 6.11$, $p < 0.05$.

[27] For further discussion of the relation between these two sets of findings, see Buckwalter manuscript a.

[28] $N = 1836$; 715 men, 1090 women, 37 unreported; 48% under 30 years of age; 78% self-identified as white non-Hispanic; 50% reported that they hold less than a bachelor's degree.

[29] Our eight item demographic questionnaire collected information about gender, age, education level, philosophical training, native language, race, religiosity, and income level.

questionnaire asked about previous philosophical training. In the results we discuss below, we focus on the data from people who reported that they have taken no philosophy courses. The reason for this will become clear in section 4.[30] What we have found reinforces the observation we made earlier: in this population of American participants there are quite significant gender differences in responses to philosophical thought experiments in some cases, while in other cases there aren't.

The first case we'll discuss is a version of the old philosophical chestnut, The Brain in the Vat (Brueckner 2008). The text of the scenario we used was borrowed from an earlier study in which Nichols, Stich & Weinberg (2003) found a substantial difference between the intuitions of participants who had taken few or no philosophy courses and participants who had taken three or more. This is the vignette participants were asked to read:

> George and Omar are roommates, and enjoy having late-night 'philosophical' discussions. One such night Omar argues, "At some point in time, like, the year 2300, the medical and computer sciences will be able to simulate the real world very convincingly. They will be able to grow a brain without a body, and hook it up to a supercomputer in just the right way so that the brain has experiences exactly as if it were a real person walking around in a real world, talking to other people. The brain would believe it was a real person walking around in a real world, except that it would be wrong. Instead it's just stuck in a virtual world, with no actual legs to walk and with no other actual people to talk to. And here's the thing: how could you ever tell that it isn't really the year 2300 now, and that you're not really a virtual-reality brain? If you were a virtual-reality brain, after all, everything would look and feel exactly the same to you as it does now! George thinks for a minute, and then replies: "But, look, here are my legs". He points down to his legs. "If I were a virtual-reality brain, I wouldn't have any legs really, I'd only just be a disembodied brain. But I know I have legs, just look at them! So I must be a real person, and not a virtual-reality brain, because only real people have real legs. So I'll continue to believe that I'm not a virtual-reality brain."
>
> George and Omar are actually real humans in the actual real world today, and so neither of them are virtual-reality brains, which means that George's belief is true.

After answering a comprehension check question designed to be sure that they had understood the story, participants saw the sentence:

---

[30] In addition to excluding data from participants who have taken one or more philosophy courses, participants were also eliminated if they did not select English as their native language, if they completed the studies in less than 30 seconds, or if they did not correctly answer simple comprehension check questions. Participants whose IP address indicated that they were not located in the United States were excluded from participation.

George knows that he is not a virtual-reality brain.

They were then asked to indicate their agreement or disagreement on a seven-item scale, with the leftmost anchor labeled "Completely Disagree," the midpoint labeled "In Between" and the rightmost anchor labeled "Completely Agree". The result, shown in Figure 11, was that women were significantly more likely than men to agree that George knows that he is not a virtual reality brain.[31]

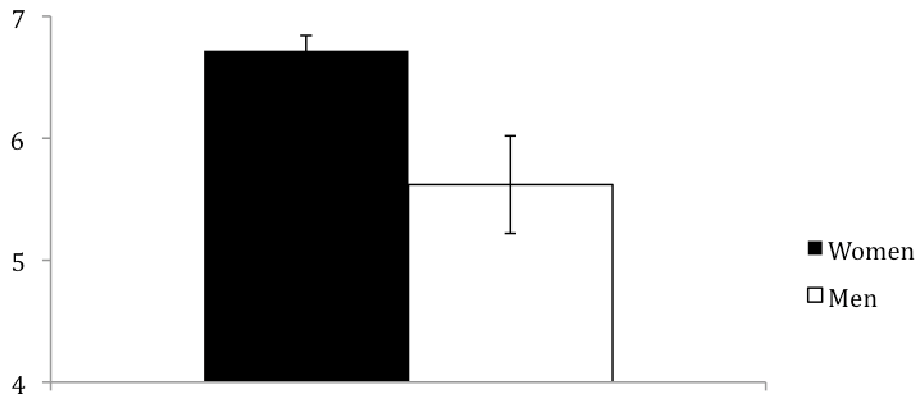

Figure 11. Mean Agreement in Brain in Vat Study (the scale ran from 1 to 7).

Twin-Earth scenarios, first used by Hilary Putnam (1973), have been become a staple of debate in the philosophy of mind and the philosophy of language (Pessin & Goldberg 1996). This is the version that we presented to participants:

> Suppose that elsewhere in the universe there is a planet called "Twin-Earth". Twin-Earth looks exactly like our Earth in virtually all respects. It is populated by twin equivalents to every person and thing here on our Earth, and even revolves around a star that appears to be exactly like our sun.
>
> Oscar grows up here on our Earth, while someone exactly like Oscar, who we can call "Twin-Oscar", lives on Twin-Earth. Oscar and Twin-Oscar both go through life having the same experiences, and both perceive their environment in exactly the same way. They look and act completely alike, and even experience the same emotions.

[31] $N = 63$; 24 men, 39 women. The mean response from male participants was 5.62, $SD = 1.97$, the mean for female participants was 6.72, $SD = 0.76$, ($d = 0.81$). An independent samples t-test reveals a significant difference between these two groups, $t(61) = -3.12$; $p < 0.01$.

> In fact, there is only one difference between these two planets. The difference is that on Earth the stuff that fills the lakes and rivers and that people and animals drink is H_2O, while on Twin Earth, the stuff that fills the lakes and rivers and that people and animals drink is another chemical compound, XYZ, that to the naked eye looks completely indistinguishable from the H_2O on Earth. H_2O and XYZ also taste exactly the same, and both have the ability to quench thirst and to sustain life.
>
> However, Oscar and Twin-Oscar both live before the development of modern science, and they have no idea about chemistry or molecular composition. When they go for a swim, both Oscar and Twin-Oscar point to the liquid in the lake and call it "water" even though on Earth that liquid is made up of H_2O, and on Twin-Earth it is made up of XYZ.

Again, there was a comprehension check question, and then participants were asked:

> ***When Oscar and Twin-Oscar say "water" do they mean the same thing, or different things?***

Responses were solicited on a seven-point scale, the leftmost anchor labeled "they mean different things," the midpoint labeled "in between" and the rightmost anchor labeled "they mean the same thing". The result was that women were more likely to give the Putnamian answer that Oscar and Twin-Oscar mean different things when they say "water" (Figure 12).[32]

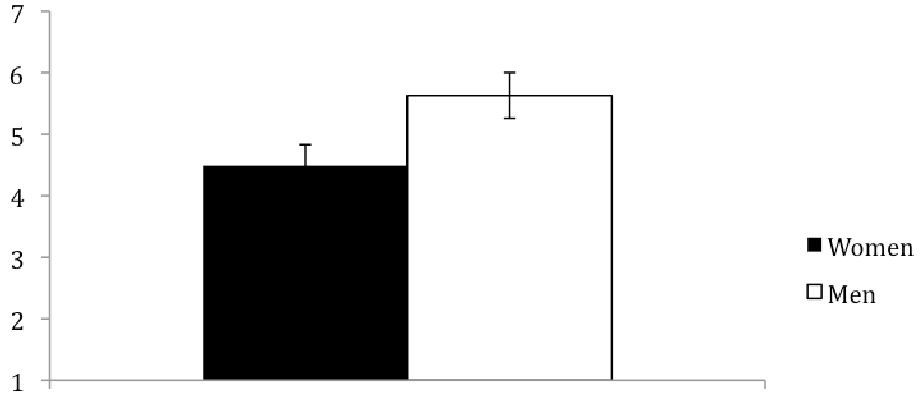

Figure 12. Mean Judgment in Twin Earth Study (the scale ran from 1 to 7).

---

[32] $N = 84$; 35 men, 49 women. The mean response for male participants was 5.63, $SD = 2.21$, the mean for female participants was 4.49, $SD = 2.42$, ($d = 0.49$). An independent samples t-test reveals a significant difference between these two groups, $t(82) = 2.21$, $p < 0.05$.

Searle's Chinese Room thought experiment (Searle 1980) has been widely discussed in the philosophy and cognitive science literature and has found its way into a number of introductory textbooks. In our study, we asked participants to read the following version of the Chinese Room scenario.

> Jenny is a native English speaker who can only speak English. She is locked in a room full of boxes of Chinese symbols, together with an instruction manual written in English for manipulating the symbols. People from outside the room send in notes on pieces of paper with Chinese symbols written on them, which unknown to Jenny, are questions in Chinese. Jenny's job is to look through her manual until she finds the symbols that look exactly like the ones written on the pieces of paper. When she finds that string of symbols, the manual will tell her what new string of symbols to write down, and send to the people outside the room.
>
> By following the instructions in the manual, Jenny is able to give the correct answers to the questions. The system consisting of Jenny and the instruction manual that she is using can be thought of as an unusual sort of computer. Jenny gets so good at following the instructions in the manual, that from the point of view of any one outside the room who speaks Chinese, her responses are absolutely indistinguishable from those of Chinese speakers.

After responding to a comprehension question participants were asked whether they agreed with the following statement.

> ***The computational system consisting of Jenny and her instruction manual understands the Chinese written on the notes.***

As in the Brain in the Vat case, responses were on a seven-item scale with the leftmost anchor labeled "Completely Disagree," the midpoint labeled "In Between" and the rightmost anchor labeled "Completely Agree". The result, shown in Figure 13, was that men were more likely than women to agree that the computational system understands the Chinese written on the notes.[33]

[33] $N = 127$; 54 men, 73 women. The mean response for male participants was 4.13, $SD = 2.47$, the mean for female participants was 3.25, $SD = 2.36$, ($d = 0.37$). An independent samples t-test reveals a significant difference between these two groups, $t(125) = 2.05$, $p < 0.05$.

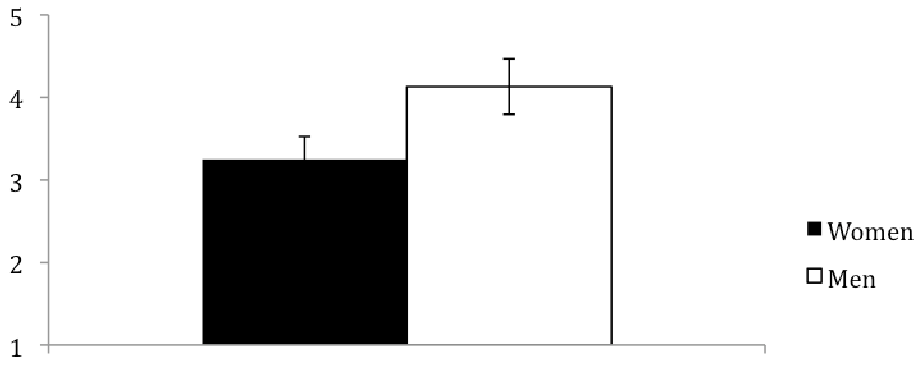

Figure 13. Mean Judgment in Chinese Room Study (the scale ran from 1 to 7).

In a fourth study, we looked at participants' moral intuitions, using a version of the "Plank of Carneades" thought experiment. Participants saw the following vignette:

> There are two shipwrecked sailors, Jamie and Ricki. They both see a small plank that can only support one of them and both of them swim desperately towards it. Jamie gets to the plank first. Ricki, who is stronger and is going to drown, pushes Jamie off and away from the plank and, thus, ultimately, causes Jamie to drown. Ricki gets on the plank and is later saved by a rescue team.

After responding to a comprehension question participants were asked:

How morally blameworthy is Ricki for what he did?

Participants answered on a seven-item scale, with the leftmost anchor labeled "not at all blameworthy" the midpoint labeled "in between" and the rightmost anchor labeled "extremely blameworthy." The result was that while both groups were above the midline, on average women judged the character in the vignette more morally blameworthy (Figure 14).[34]

[34] N=110; 37 men, 73 women. The mean response for male participants was 4.95, SD = 2.07, the mean for female participants was 5.64, SD = 1.35 (d = 0.42). An independent samples t-test reveals a significant difference between these two groups, $t(108) = -2.13$, $p < 0.05$.

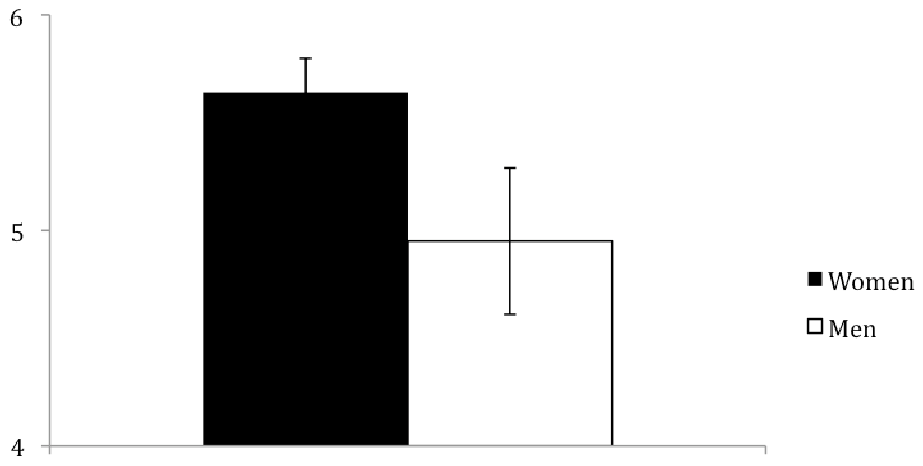

Figure 14. Mean Judgment in Plank Study (the scale ran from 1 to 7).

3.9. Similar findings in Social Psychology and Experimental Economics

Our theme in section 3 has been that in the Americans and Canadians who participated in the experiments we have reviewed there are indeed gender differences in intuitive responses to many philosophical thought experiments, some of which are large, unexpected and dramatic. Though as we have noted repeatedly, there are also a number of studies of philosophical intuitions that do not find gender differences. To the best of our knowledge, there is currently no good way of predicting where these gender differences will be found. It has been our experience that philosophical audiences often find results indicating gender differences in intuition to be quite surprising. But perhaps they shouldn't. For the picture we have been sketching is broadly consistent with a substantial body of research in psychology (Eagly 1995; Halpern 2011) and experimental economics (Croson & Gneezy 2009; Eckel & Grossman 2008a & 2008b), that has reported gender differences in studies of a number of preferences, decisions and behaviors that are (or should be) of considerable interest to philosophers. Though this is not the place to review this literature in detail, a single example may help to convey its interest and importance.

Several studies of prosocial behavior have shown that women are considerably more generous than men in double-blind dictator games. In these games, one person (the "dictator") must choose how to divide a fixed amount of money between himself and an anonymous recipient. A common finding is that female dictators give recipients about twice as much as male dictators (Eckel & Grossman 1998, Vesterlund 2001, Rigdon et al. 2009). But there is an important exception to this rule. If there is even the slightest hint that participants are being observed, the gender disparity disappears. In one fascinating study, Rigdon at al. (2009) presented participants (N = 113; 51 men, 62 women) with either the "face" or the "control" version of the task displayed in Figure 15.[35]

[35] "Face" stimuli like those used in Rigdon et al. (see figure 15) have been shown to activate the

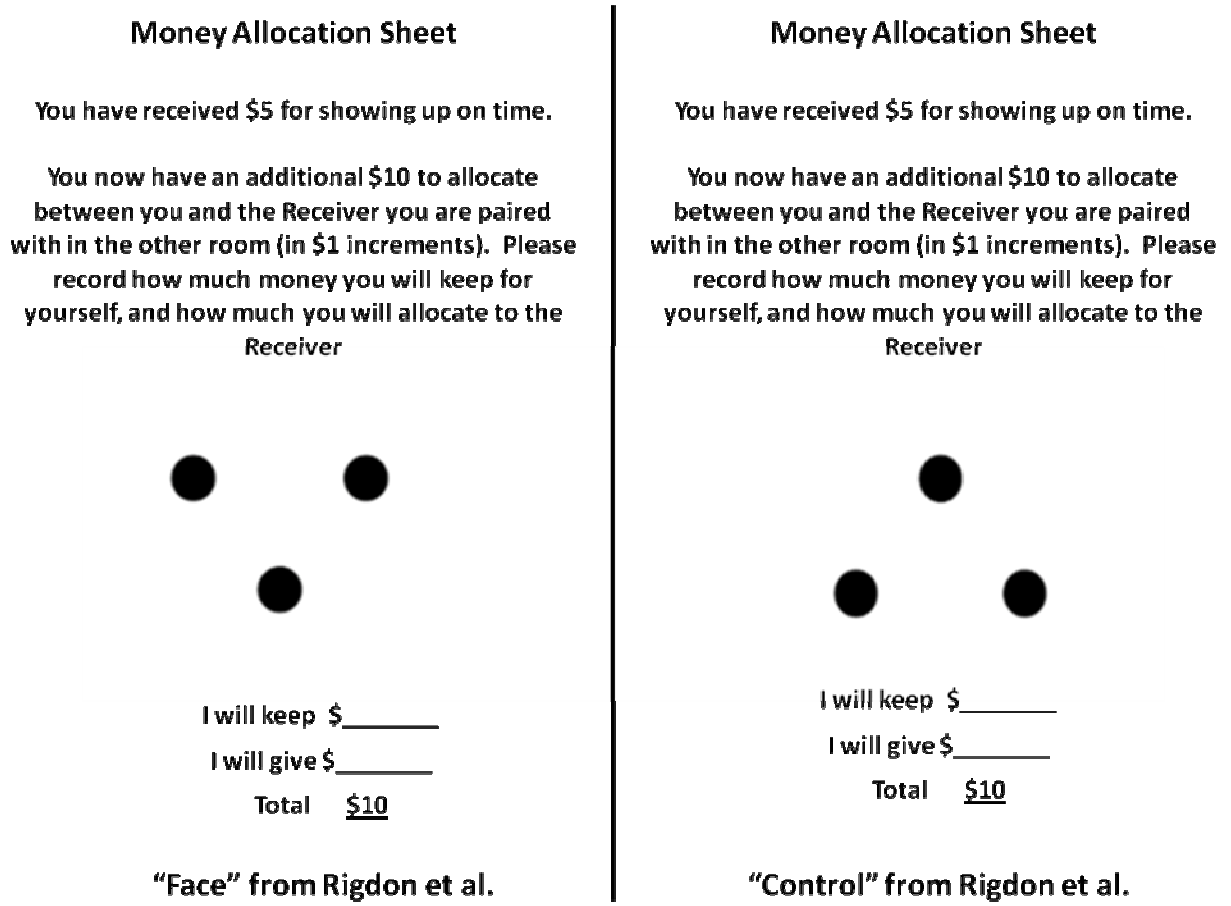

Figure 15. Stimuli in Dictator Game Study.

Figure 16 shows the percent of participants who transferred a dollar or more in each condition. Though the manipulation had no significant effect on the women, it had a powerful effect on the men. In the control condition, only 37% of men transferred $1.00 or more, but in the presence of a weak social cue – three dots arranged to look like a face – 79% of men transferred $1.00 or more to the "Receiver" they had been paired with ($p = 0.006$). Rigdon et al. also found that the three dot "face" dramatically increases the average amount that male participants give – from $1.41 in the control condition to $3.00 in the face condition, while having no significant effect on female participants who gave $2.12 in the face condition, and $2.79 in the control condition).[36]

---

fusiform face area (FFA) of the brain (Bednar and Miikkulainen 2003).

[36] Rigdon et al. report that they conducted a Logit analysis to see the extent to which gender influences the amount transferred in the face condition. Regressing the amount (where amount = 1 if transfer > average transfer) on gender (where gender = 1 if male), they found that males are 3.35 times more likely than females to fall in this category ($p = 0.048$). The difference in average transfers by female participants across treatments is not significant ($p = 0.1592$); the

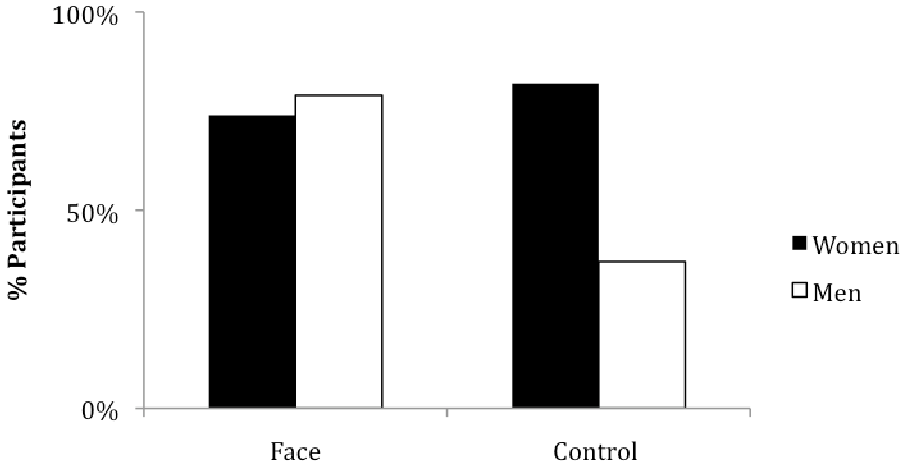

Figure 16. Percent Transfer (>$1) in Dictator Game Study.

3.10. Conclusions, Caveats and Further Questions

Our central claim has been that in the North American populations that have been studied there are significant differences between men and women in intuitive responses to some philosophically important thought experiments. While much work remains to be done, we believe we have made a reasonably persuasive case for this conclusion. In section 3.8, we saw that there are also significant differences between the genders in some philosophically interesting behaviors and choices, like those studied in various economic games. It is important to note that in a variety of other domains psychologists have either failed to detect gender differences or have detected only very weak gender differences.[37] However, we do not think that the absence of strong gender differences in many areas of mental processing provides plausible grounds to doubt the variation in philosophical intuitions that has been our concern. Instead, we think, it shows that the differences between men and women are subtle, complex, and – for the moment, at least – often quite unpredictable.

A natural question to raise at this point is: What is the *explanation* for the gender differences in philosophical intuition? On our view, however, it would be premature to venture an answer. We will need to know much more about gender differences in intuition (and behavior

---

difference in proportion of female participants who send $1.00 or more is not significantly different across the treatments (p = 0.4223).

[37] See, for example, Hyde's paper, "The Gender Similarity Hypothesis,"(2005) which offers a meta-analysis in defense of the claim that on many different psychological variables, men and women are quite similar.

and choices) before making a serious attempt to explain them. One of the most important questions that remains unanswered is whether these differences in intuition are a culturally local phenomenon. As we have taken pains to emphasize, all the participants in the studies we have discussed were located in North America, and many of them were university students. This is not at all unusual. In a recent and very influential paper, Henrich, Heine and Norenzayan (2010) point out that a substantial majority of experimental participants in recent behavioral science experiments are Americans, and a majority of these are university students. Indeed, "a randomly selected American undergraduate is more than 4,000 times more likely to be a research participant than is a randomly selected person from outside the West" (ibid, 63). However, Henrich and colleagues go on to show that when cross-cultural data are available it often turns out that Americans are outliers, and American university students are outliers among the outliers! So without data from cross-cultural studies, we think it would be unwise to assume that the intuition differences we have reported would be found in European participants, and even more unwise to assume that they would be found in Asians or Africans or other groups around the globe. Moreover, when cross-cultural studies of philosophical intuitions are done, there is no reason to suppose they will all turn the same way. It is entirely possible that some gender differences we have recounted are pan-cultural while others are culturally local.

There are many other important questions that also remain unanswered: Do some gender differences in intuition reflect deep cognitive or affective differences between the genders, or do they all arise from relatively superficial semantic factors, or pragmatic factors, or local norms of self-presentation – or from something else entirely? Here again the answer is that we do not know. Are there different explanations for the gender differences in intuition in different parts of philosophy? Is the explanation for the gender differences in moral intuitions different from the explanation for the gender differences in metaphysical intuitions? Indeed, is the explanation for the gender differences in one sort of moral intuition different from the explanation for the gender differences in another sort of moral intuition? Once again, we do not know. We believe that all of these questions are addressable using the techniques of contemporary cognitive science, and that this sort of experimental philosophy should be high on the agenda of philosophers interested in understanding their own discipline and explaining why men and women sometimes have different philosophical intuitions. Much more work will be needed before we understand gender differences in philosophical intuition.

Another question raised by the finding that men and women sometimes have significantly different philosophical intuitions is: What implications do these gender differences have for philosophical methodology and philosophical pedagogy? We believe that the implications are both important and pervasive. But defending that view requires a long, careful and systematic argument that will have to be postponed for another occasion.[38]

4. How Gender Differences in Philosophical Intuition Might Help to Explain the Gender Gap in Academic Philosophy

[38] See Stich (in preparation).

In this section, we turn from facts to hypotheses. Our goal is to outline a series of conjectures about ways in which the sorts of gender differences in philosophical intuition that we have been exploring, and differences in intuition that may not be associated with gender, might be part of the explanation for the gender gap in academic philosophy. Along the way, we will present some evidence that is compatible with these hypotheses, though much more evidence will be required before we can have confidence that the hypotheses are correct. Before beginning, we want to make it very clear that if hypotheses like those we will discuss turn out to be true, it most definitely would not follow that differences in intuition are the *only* factor leading to the gender gap in philosophy. Quite to the contrary, we think that historical, sociological and economic factors are also very likely to be part of the explanation, and that gender based discrimination and sexist attitudes and behavior are also important contributing factors (Haslanger, 2008; Saul, forthcoming). All of this would be fully compatible with our hypotheses. We have chosen to focus our research on differences in philosophical intuition not because we think it is the only factor involved, or the most important, but because until very recently these differences were almost entirely unrecognized. Moreover, as we will argue in this section, when combined with one of the standard methods invoked in doing and teaching analytic philosophy, they have the potential to generate *unconscious* and *unintentional* biases against women. If we are to develop efficacious strategies for combating the underrepresentation of women that is rampant in our profession, it is crucial that we understand how these biases might arise, assess how large their impact has been, and begin to think about ways to lessen their influence.

We will start with a brief elaboration of a point made in section 2. When a philosopher invokes a philosophical intuition in a philosophical argument, the intuition (or, more accurately, the propositional content of the intuition) is typically being used as *evidence*. Philosophers rarely argue that the propositional content of an intuition they are invoking is true. Rather, they take the propositional content of the intuition to be *obvious*, and they use the proposition as a premise in the argument they are constructing. So, for example, while philosophers (and philosophy instructors) can and do offer arguments from the intuitively supported *premise* that the protagonist in a Gettier-style thought experiment does not know that p, to the conclusion that the justified-true-belief account of knowledge is false, philosophers rarely even attempt to argue that the protagonist does not know that p. If an interlocutor were to deny that the premise is true and insist that the protagonist *does* know that p, the philosopher might take steps insure the interlocutor has understood the story and has not ignored or forgotten some of the crucial details. But if the interlocutor has understood the vignette and has not ignored important details, it is far from clear what else the philosopher could say to convince him, since philosophers generally assume that it is obvious that the Gettier protagonist does not know that p.[39] In this respect, as Sosa has noted, the role of philosophical intuition is similar to the role of observation or perception in providing evidence for scientific theories.

> [T]he way intuition is supposed to function in epistemology and in philosophy more generally … is by analogy with the way observation is supposed to function

[39] As Ned Block reminded us, not all that long ago it was a common practice for philosophers to dismiss people who didn't share their intuitions by saying that they have "a tin ear".

> in empirical science. Empirical theories are required to accord well enough with the deliverances of scientific observation.[40] (Sosa 2007, 107)

But now consider the predicament of a young woman in a philosophy class, who (like 71% - 75% of women in the Starmans & Friedman study) does not find it obvious that the characters in Gettier vignettes do not have knowledge of the relevant proposition. Rather, her intuitions tell her that the Gettier characters *do* have knowledge, though her instructor, whether male or female, as well as a high percentage of her male classmates, clearly think she is mistaken. Different women will, of course, react to a situation like this in different ways. But it is plausible to suppose that some women facing this predicament will be puzzled or confused or uncomfortable or angry or just plain bored. Some women may become convinced that they aren't any good at philosophy, since they do not have the intuitions that their professors and their male classmates insist are correct. If the experience engenders one or more of these alienating effects, a female student may be less likely to take another philosophy course than a male classmate who (like 59% - 64% of the men in the Starmans & Friedman study) has the "standard" intuitions that their instructor shares. That male student, unlike the majority of his female classmates, can actively participate in, and perhaps enjoy, the project of hunting for a theory that captures "our" intuitions.

If these speculations are on the right track, then as students in philosophy courses are repeatedly exposed to the practice of using intuitions as evidence, we should expect to find enrollments of women dropping off. The more courses a woman takes, the more likely it is that she will be exposed to thought experiments on which her intuitions and those of her instructor diverge – and the more likely it is that she will decide not to take another course. Is this the case? Though gender coded enrollment data are not readily available, we have been able to obtain reliable data on the gender ratios in philosophy courses during a recent 10 year period at Rutgers University, a large state university with a highly ranked philosophy department.[41] As can be seen in Figure 17, in the 100-level introductory courses, the percent of female and male students is almost equal: 46.2% female and 53.8% male.[42] But at *each* higher level, the percent of women goes down: 40.38% at the 200-level, 36.50 at the 300-level, 29.31 at the 400 level, and 26.2% in 500 thru 800 level graduate courses.[43]

---

[40] The analogy, as Jonathan Weinberg (2007) notes, is rather too favorable toward intuition, since in the case of perceptual disagreements there are often well-established procedures for adjudication.

[41] Our thanks to Barry Qualls, Vice-President for Undergraduate Education at Rutgers, and Kenneth Iuso, the Rutgers University Registrar, for making these data available.

[42] Students at Rutgers are not required to take an introductory philosophy course.

[43] Percentages calculated from 79,904 enrollments (47,013 men, 32,891 women) during the target period. Though we claim that these enrollment data are *consistent* with our hypothesis, we want to be very clear that we do not claim that they support our hypothesis over a variety of other hypotheses that might be offered about the causal mechanisms responsible for the underrepresentation of women in philosophy. At the 2011 meeting of the Society for Philosophy

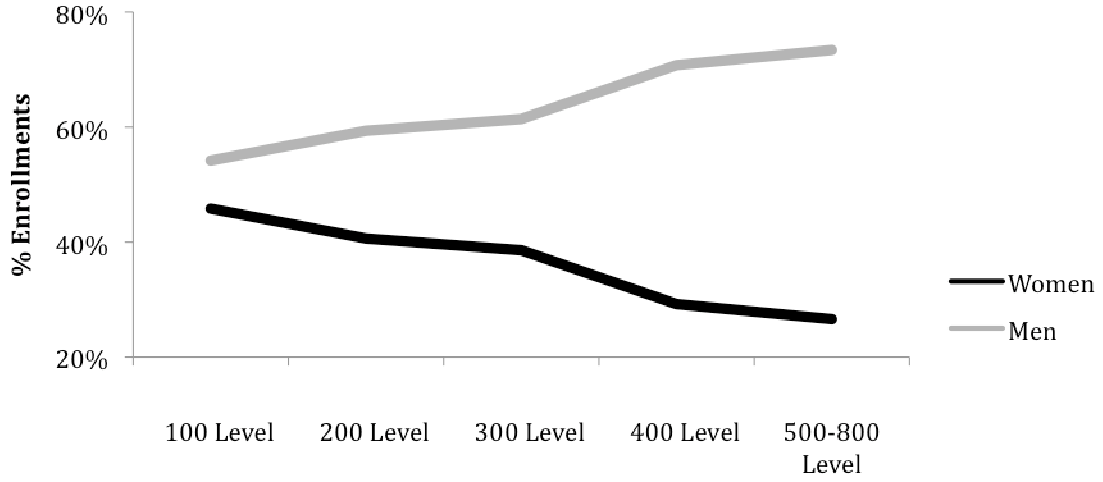

Figure 17. Percent Enrollment in Undergraduate Philosophy Courses at Rutgers, The State University of New Jersey, from Spring 1999 to Spring 2010.

Of course, the hypothesis we are urging is only plausible if students are, often enough, exposed to thought experiments that tend to evoke different intuitions in men and women. Are they? One way to address this question is to look at the most widely used philosophy textbooks. This is easier said than done, since most major textbook publishers treat their sales data as proprietary information that is not publicly available. As a workaround, we contacted the philosophy textbook editor of a leading publishing house and asked him to provide his professional opinion about which texts published by both his firm and his competitors had the largest sales. We then examined these texts to see which philosophical thought experiments

---

and Psychology, Molly Paxton, Carrie Figdor and Valerie Tiberius presented data from an ongoing study that are broadly consistent with this pattern. They found that the largest statistically significant drop in the proportion of women in philosophy occurs between taking an introductory class in philosophy and declaring a major in philosophy. For details of this valuable and sophisticated study, see Paxton et al. (in preparation). Additional data from the University of Minnesota, provided by Valerie Tiberius, confirmed that there is a precipitous drop in the proportion of women taking philosophy courses between introductory level courses, where about 45% of the students are women, and intermediate level courses, where only about 28% are women. Tiberius also provided University of Minnesota enrollment data for eleven other departments: Anthropology, Biology, Computer Science, Electrical Engineering, English, History, Journalism, Mechanical Engineering, Political Science, Psychology and Sociology. None of them exhibited the pattern found in Philosophy. Obviously *something* is discouraging women students from continuing in philosophy beyond the introductory level, and whatever it is is not present in many other disciplines.

were most common.[44] The five most common, in decreasing order, were versions of (or close relatives of)

1. The Violinist
2. The Brain in the Vat
3. Compatibilism (thought experiments about free will in a deterministic universe)
4. Physicalism (inspired by Jackson's Mary-case)
5. The Chinese Room

This was a relatively crude exercise, to be sure, but the results are suggestive nonetheless. Note that gender differences in intuition have been found in versions of all of the philosophical thought experiments on this list. So there is reason to conclude, albeit tentatively, that many students are indeed exposed to these thought experiments and that the more philosophy courses they take, the more philosophical thought experiments they encounter.

What we are suggesting is that part of the gender gap in academic philosophy can be explained as a *selection effect*. Students come to philosophy with somewhat different intuitions about many standard philosophical thought experiments, and as we have shown, in many cases there are statistically significant differences between women's intuitions and men's. However, most of the faculty members who get to say which intuitions are correct (and "obvious") are now, and always have been, men. So women students are more likely than men students to find that their intuitions about the thought experiments discussed in their philosophy classes are at odds with those of their instructor. If it is indeed the case that students (of either gender) are less likely to continue in philosophy if their intuitions do not accord with those of their instructor, then all the elements of a powerful and cumulative selection effect are in place – a selection effect which "filters out" a greater proportion of women than of men.

We are not the first to suggest the existence of an intuition-based selection effect in philosophy, though, to the best of our knowledge, it has not previously been linked to gender. In his well known paper, "Reflection on Reflective Equilibrium," Robert Cummins (1998) wrote:

> The Putnamian take on these [Twin-Earth] cases is widely enough shared to allow for a range of thriving intramural sports among believers. *Those who do not share the intuition are simply not invited to the games*. This kind of selection allows things to move forward, but it has its price…. We must take care that such agreement about intuitions as there is is not merely a selection effect. This is easier said than done, since it is all too easy for insiders to suppose that dissenters just do not understand the case. *If we are honest with ourselves, I think we will*

---

[44] We searched over 25 leading introduction to philosophy, introduction to ethics, contemporary moral issues, philosophy of mind, philosophy of language, epistemology, and metaphysics textbooks and readers published by Cengage, McGraw Hill, Oxford University Press, and Pearson Press.

> *have to confront the fact that selection effects like this are likely to be pretty widespread in contemporary philosophy*. (116, emphasis added)

What we would add to this is that, more often than not, those who are filtered out by these intuition-based selection effects are women.

One concern about our selection effect hypothesis that has been raised frequently by philosophical colleagues with whom we have discussed our work is that in a number of the cases of gender difference that we report in section 3, it is not clear that it is the women who have the intuitions that most professional philosophers take to be incorrect. What fuels this worry is that, while there is substantial data showing gender linked variation in non-philosophers' intuitions about various philosophical thought experiments, there is little or no reliable data concerning professional philosophers' intuitions in these cases. If the differences in non-philosophers' intuitions are going to be part of the explanation of why women are more inclined than men to stop taking philosophy classes, then the female intuitions in these studies should be more likely than the male intuitions to diverge from the consensus among professional philosophers. However, since there is relatively little data about the intuitions of professional philosophers, it is not clear that the evidence assembled in section 3 exhibits this pattern.

While there is little data available indicating how professional philosophers might respond to the specific thought experiments discussed in section 3, many colleagues have offered conjectures. Most of them agree that the women's intuitions reported in the Gettier, Violinist, and Magistrate and Mob cases are at odds with the dominant intuitions of professional philosophers in these cases. Though they are less confident, a number of colleagues are also inclined to think that the women's intuitions differ from those professional philosophers would have in the Dualism and Epistemic Side-Effect Effect experiments. However, in three of the cases presented in section 3 – Compatibilism, the Brain in the Vat and Twin Earth – most of the colleagues we have spoken with speculate that the women's intuitions in the studies are *more* consistent with the dominant view among professional philosophers and the men's intuitions are less consistent.

This is an important challenge.[45] But, for two very different reasons, we are not persuaded that it undermines our selection effect hypothesis. The first is that, as already noted, in most cases we really do not know what the dominant intuition is among professional philosophers. It has been suggested that the *PhilPapers* survey cited in section 1 could help to shed light on this matter. However, that survey asked about respondents' philosophical *views* rather than about their intuitions in response to specific thought experiments. So, for example, the Free Will question in the *PhilPapers* survey simply asked, " Free will: compatibilism, libertarianism, or no free will?" Results indicated that 59% "accepted or leaned toward" compatibilism, 13.7% libertarianism, 12.2% no free will, 14.9% other. We doubt that information of this sort gives us a reliable way of predicting what participants in the *PhilPapers* survey would have said in response to any detailed compatibilism thought experiment. For, as

---

[45] Our thanks to the many colleagues who have proposed versions of this challenge. Special thanks to Michael Bishop for a particularly clear and detailed statement of the argument.

the burgeoning experimental philosophy literature exploring intuitions about free will and responsibility makes clear, people's intuitions in this area are exquisitely sensitive to subtle features of the vignette. With a relatively small and apparently inconsequential change in wording, a vignette that evokes mostly compatibilist intuitions will sometimes evoke mostly incompatibilist intuitions.[46] Moreover, to assess the selection effect hypothesis, what we really need to know is not which intuitions are dominant among professional philosophers but which intuitions are dominant among the subset of professional philosophers (and philosophers-in-training) who teach the first few philosophy courses to which students are exposed. And since there are typically many quite different thought experiments used to gather evidence for or against a philosophical theory, we need to know which versions are actually used in the classroom. So, contrary to what our critics have suggested, it is far from clear that the men in the studies reviewed in section 3 often have intuitions that are incompatible with those they are likely to encounter from their philosophy instructors.

Clearly, the response we have just offered is a two-edged sword. If we do not have adequate information about which intuitions are favored by people teaching philosophy courses, it undercuts the plausibility of the objection we have been considering, but it also undercuts the plausibility of our claim that an intuition based selection-effect contributes to the gender disparity in philosophy. However, our second response suggests that one edge of the sword is significantly sharper than the other. The core idea of our second response is that difference in intuitions can interact with a cluster of other factors in ways that will make it more likely that women will be discouraged from continuing in philosophy. The factors we have in mind have been a main focus of the widely discussed work of Carol Dweck and her associates.[47]

The story starts with the finding that there is an important difference in the way people, both children and adults, think about intelligence and intellectual ability. Dweck and colleagues have used a variety of terms for this distinction. In a recent book aimed at a popular audience, she calls them the "fixed mindset" and the "growth mindset". People with the fixed mindset view intellectual ability in general and more focused abilities in specific subjects as "a gift – an ability that you simply have or you don't" (Dweck, 2006, 3). Those with the growth mindset view both general intelligence and ability in specific subjects as malleable and cultivatable – "something that builds on an initial ability and expands through practice and dedication."(ibid.) The distinction is important because it correlates with a surprising range of additional phenomena, particularly in the domain of education (Dweck, 2000). The most important of these, for our purposes, is that "viewing intellectual ability as a gift (a fixed entity) led students to question that ability and lose motivation when they encountered setbacks. In contrast, viewing intellectual abilities as qualities that could be developed led them to seek active and effective remedies in the face of difficulty" (Dweck, 2006, 3). In an experiment reported in Licht & Dweck (1985), it was found that inducing confusion, by including some intentionally puzzling text in an otherwise straightforward series of written lessons, had a major impact on the success of fixed mindset 5th graders in mastering the remaining, sensible, material. But inducing

[46] See, for example, Nahmias, Coates & Kvaran (2007) and Feltz, Cokely & Nadelhoffer (2009).

[47] We are indebted to Tania Lombrozo for calling Dweck's work to our attention. For an excellent – though now slightly dated – overview of Dweck's work, see Dweck (2000).

confusion had no impact on the success of growth mindset children in mastering the rest of the material. This, and a number of similar findings, some of them involving university students (Grant & Dweck, 2003) indicate that fixed mindset individuals do not deal well with material that engenders puzzlement or confusion.

What makes all of this relevant to our current concerns is that there is also a substantial body of evidence indicating that the fixed mindset view of intellectual ability is significantly more common in females than in males (Dweck & Gilliard, 1975; Dweck & Bush, 1976). So if, as we have suggested, having philosophical intuitions that one's instructor takes to be obviously mistaken (though he typically does not and cannot explain why) engenders puzzlement and confusion, and if fixed mindset individuals react to this by becoming convinced that they are not good at philosophy and losing their motivation to pursue it, then we would expect fixed mindset individuals to drop out of philosophy more frequently than growth mindset individuals, when they find that they have what their instructors take to be the wrong intuitions. And the majority of these fixed mindset individuals will be women. Thus even if it is the case that women and men are about equally likely to have intuitions that are at odds with those of their instructors, women are more likely than men to conclude that they are no good at philosophy.

To make matters worse, Dweck and her colleagues have also found that the debilitating effect of puzzlement and confusion interacts with IQ differently in males and females. Among males, those with higher IQs did better when confronted with confusing material than those with lower IQs. But among females, the pattern was reversed. The brighter they were, the more likely they were to be negatively impacted by confusing material (Licht & Dweck, 1984, 634). So the impact of having intuitions that clash with those of one's instructors may tend to selectively discourage *bright women*. Thus when intuitions play a significant role in philosophical education, as the level of the course and the difficulty of the material increases, we might expect the men to excel, to get more encouragement from their instructors, and thus to be more inclined to continue in philosophy, while the women will be more inclined to look elsewhere. If this story, or even a substantial part of it, is more or less on the right track, then differences in philosophical intuition will play a role in unleashing a cascade of events that increasingly skews the gender distribution in philosophy courses toward men and away from women.

The goal of this section has been to develop some hypotheses about ways in which differences in philosophical intuition – both those that are associated with gender and those that are not –might play a role in explaining the underrepresentation of women in philosophy. Here is how we see the state of play. It is plausible to assume that students who have intuitions that conflict with those of their instructor often find the experience puzzling, confusing and alienating. If so, then in those cases where the majority of women have intuitions that are in conflict with the intuitions of the majority of philosophy instructors (and the majority of males in their classes), a straightforward selection effect would lead to more women deciding not to continue in philosophy. If cases like this predominate in contemporary philosophical education, then a simple selection effect hypothesis might be all we need to link differences in philosophical intuition with the underrepresentation of women. But since we have relatively little reliable data about the intuitions of philosophy instructors, it is not clear that women have intuitions that are inconsistent with those of their instructors significantly more often than men in thought

experiments that are commonly invoked in philosophy classes. However, even in the cases where more men than women have the professionally disparaged intuition, the data assembled in section 3 suggests that there will still be *lots* of women who have the that disparaged intuition. Indeed, one of the most striking facts about the data we have presented is how much *disagreement* there is in intuition, not just between the genders but within them.[48] Dweck's work suggests that when more men than women have the intuition that their instructors take to be mistaken, it is still the case that more women may be driven from philosophy, since when confronted with confusing or puzzling material women are more likely than men to conclude that they have little ability in that domain, and the brighter a woman is, the more likely it is that she will be susceptible to this effect. As the acute reader will have noticed, what is doing the work in our Dweck-inspired hypotheses is not *gender* differences in intuition but *differences* in intuition *tout court*. This might lead one to wonder whether philosophical practice and philosophical education should rely so heavily on undefended, and typically undefendable, intuitions that *some* people take to be *obviously* true while many others do not. We think this is an excellent question which one of us has tried to address elsewhere.[49]

5. Conclusion

In this paper, we have had a pair of goals. The first was to call attention to a fact about the sorts of intuitions that philosophers have relied on as evidence from antiquity to the present. In some cases – or perhaps in many – men and women in our culture tend to have different philosophical intuitions. Until the recent blossoming of experimental philosophy, these differences had been almost entirely unrecognized. Though there are now more than a dozen studies in which gender differences in philosophical intuitions have been found, there is no obvious pattern in these findings and no good way to predict where gender differences will be found in the many cases that have not yet been studied. Much work remains to be done before we have an accurate account of the contours of gender differences in philosophical intuition and a good explanation of why they exist.

Our second goal was to propose a cluster of hypotheses about how differences in intuition, both those that are associated with gender and those that are not, might play a role in explaining the egregious underrepresentation of women in philosophy. We have set out several ways in which differences in philosophical intuition, along with the practice of using intuitions as evidence, might lead to unconscious and unintentional bias against women. Bias of this sort, if it exists, is surely not the only factor contributing to the underrepresentation of women in philosophy, though it might well amplify and exacerbate some of the other causes of the gender gap in philosophy. In order to develop effective strategies for combating the gender disparity, we need to know whether the hypotheses we have proposed are true and, if they are, how large a role the bias they generate is playing in generating and sustaining the lamentably skewed

[48] Arguably, the existence of widespread disagreement in philosophical intuitions is the single most consistent finding to emerge from a decade of work in experimental philosophy. For some further discussion of this point, see Mallon et al. (2009).

[49] Stich (in preparation).

demography of academic philosophy. In research currently underway, we hope to make some progress toward answering these questions.

References

Bealer, G. (1998). Intuition and the autonomy of philosophy. In M. DePaul and W. Ramsey (eds.), *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*. Lanham, MD: Rowman and Littlefield. 201-239.

Beebe, J. & Buckwalter, W. (2010). The epistemic side-effect effect. *Mind and Language*, 25: 474-498.

Beebe, J. & Jensen, M. (forthcoming). Surprising connections between knowledge and action: The robustness of the epistemic side-effect effect. To appear in *Philosophical Psychology*.

Beebee, H. & Saul, J. (2011). Women in Philosophy in the UK: A Report. www.bpa.ac.uk/policies.

Bednar, J., and Miikkulainen, R. (2003). Learning innate face preferences. *Neural Computation*, 15: 1525-1557.

Buckwalter, W. (unpublished manuscript, a). Gender and epistemic intuition. City University of New York, Graduate Center.

Buckwalter, W. (unpublished manuscript, b). Solving the puzzle of causation by absence: Experiments in metaphysics. City University of New York, Graduate Center.

Brueckner, T. (2008). Brains in a vat. *The Stanford Encyclopedia of Philosophy (Fall 2008 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2008/entries/brain-vat/>.

Cappelen, H. (forthcoming). *Philosophy Without Intuitions*. Oxford: Oxford University Press.

Croson, R. and Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, 47: 448-474.

Cummins, R. (1998). Reflections on reflective equilibrium. In M. DePaul and W. Ramsey (eds.), *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*. Lanham, MD: Rowman and Littlefield. 113-127.

Deutsch, M. (2010). Intuitions, counter-examples and experimental philosophy. *Review of Philosophy and Psychology*, 1: 447-460.

Devitt, M. (2006). Intuitions. In V. Gomez Pin, J. Galparaso and G. Arrizabalaga (eds.), *Ontology Studies Cuadernos de Ontologia: Proceedings of VI International Ontology Congress*, 169-76.

Devitt, M. (2009). On determining what there isn't. In D. Murphy and M. Bishop (eds.), *Stich and His Critics*. Oxford: Blackwell. 46-61.

Dweck, C. and Gilliard, D. (1975). "Expectancy statements as determinants of reactions to failure: Sex differences in persistence and expectancy change," *Journal of Personality and Social Psychology*, 32, 1077-1084.

Dweck, C. and Bush, E. (1976). Sex differences in learned helplessness: I. Differential debilitation with peer and adult evaluators. *Developmental Psychology*, 12: 147-156.

Dweck, C. (2000). *Self-Theories: Their Role in Motivation, Personality, and Development*. New York: Psychology Press.

Dweck, C. (2006). Is math a gift? Beliefs that put females at risk. In S. Ceci and W. Williams, (eds.), *Why Aren't More Women in Science? Top Researchers Debate the Evidence*. Washington, D.C.: American Psychological Association. Also available on Prof. Dweck's website at: https://www.stanford.edu/dept/psychology/cgi-bin/drupalm/system/files/cdweckmathgift.pdf Page references in the text are to the website version.

Dweck, C. (2008). *Mindset: The New Psychology of Success*. New York: Ballantine Books.

Eagly, A. (1995). The science and politics of comparing women and men. *American Psychologist*, 50: 145-158.

Eckel, C. and Grossman, P. (1998). Are women less selfish than men? Evidence from dictator experiments. *Economic Journal*, 108: 726 -735.

Eckel, C. and Grossman, P. (2008a). "The difference in the economic decisions of men and women: Experimental evidence," in C. Plott and V. Smith (eds.), *Handbook of Experimental Economics Results*, Amsterdam: North-Holland / Elsevier. 509-519.

Eckel, C. and Grossman, P. (2008b). Sex and risk: Experimental evidence. In C. Plott and V. Smith (eds.), *Handbook of Experimental Economics Results*, Amsterdam: North-Holland / Elsevier. 1061-1073.

Feltz, A., Cokely, E. and Nadelhoffer, T. (2009). Natural compatibilism versus natural incompatibilism: Back to the drawing board. *Mind and Language*, 24: 1-23.

Gendler, T. (2007). Philosophical thought experiments, intuitions, and cognitive equilibrium. *Midwest Studies in Philosophy*, 31: 68-89.

Gettier, E. L. (1963). Is justified true belief knowledge? *Analysis* 23: 121-3.

Goldman, A. (2007). Philosophical intuitions: their target, their source, and their epistemic status. *Grazer Philosophiche Studien,* 74: 1-26.

Grant, H. and Dweck, C. (2003). Clarifying achievement goals and their impact. *Journal of Personality and Social Psychology*, 85: 541-553.

Halpern, D. (2011). *Sex Differences in Cognitive Abilities*, 4th Edition. Florence, KY: Psychology Press, Taylor & Francis.

Haslanger, S. (2008). Changing the ideology and culture of philosophy: Not by reason (alone). *Hypatia*, 23: 210-223.

Henrich, J., Heine, S. and Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33: 61-83.

Hintikka, J. (1999). The emperor's new intuitions. *Journal of Philosophy*, 96: 127–147.

Hitchcock, C. & Knobe, J. (2009). Cause and norm. *Journal of Philosophy*, 106: 587–612.

Holtzman, G. (under review). Do Personality Effects Mean Philosophy Is Intrinsically Subjective?

Holtzman, G. (in preparation, a). Why are men incompatibilists?

Holtzman, G. (in preparation, b). The psychology of physicalism.

Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist*, 60: 581-592..

Ichikawa, J. (2008). *Imagination and Epistemology*, Ph.D. dissertation submitted to the Graduate School of Rutgers University, Chapter 5.

Jackson, F. (1982). Epiphenomenal qualia. *Philosophical Quarterly*, 32: 127–136.

Jackson, F. (1986). What Mary didn't know. *Journal of Philosophy*, 83: 291–295.

Jackson, F. (1998). *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. Oxford: Clarendon Press.

Knobe, J. (2003). Intentional action and side effects in ordinary language. *Analysis,* 63: 190-193.

Knobe, J. (2010). Person as scientist, person as moralist. *Behavioral and Brain Sciences*, 33: 315-329.

Knobe, J. & Fraser, B. (2008). Causal judgment and moral judgment: Two experiments. In W. Sinnott-Armstrong (ed.), *Moral Psychology*, vol. 2: *The Cognitive Science of Morality: Intuition and Diversity*. Cambridge, MA: MIT Press.

Knobe, J., Buckwalter, W., Nichols, S., Robbins, P., Sarkissian, H., & Sommers, T. (forthcoming). 'Experimental Philosophy.' *Annual Review of Psychology*.

Kornblith, H. (1998). The role of intuition in philosophical inquiry: An account with no unnatural ingredients. In M. DePaul and W. Ramsey (eds.), *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*. Lanham, MD: Rowman and Littlefield. 129-141.

Kripke, S. (1980). *Naming and Necessity*. Cambridge, MA: Harvard University Press.

Licht, B., and Dweck, C. (1984). Determinants of academic achievement: The interaction of children's achievement orientations with skill area. *Developmental Psychology*, 20: 628-636.

Livengood, J., Sytsma, J., Feltz, A., Scheines, R., and Machery, E. (forthcoming). Philosophical temperament. *Philosophical Psychology*.

Ludwig, K. (2007). The epistemology of thought experiments: First vs. third person approaches. In *Midwest Studies in Philosophy*, *Philosophy and the Empirical*, 31: 128–159.

Mallon, R., Machery, E., Nichols, S. and Stich, S. (2009). Against arguments from reference. *Philosophy and Phenomenological Research*, 79: 332-356.

Nahmias, E., Coates, D., and Kvaran, T. (2007). Free will, moral responsibility and mechanism: Experiments on folk intuition. *Midwest Studies in Philosophy*, 31: 214-242.

Nichols, S., Stich, S. and Weinberg, J. (2003). Meta-skepticism: Meditations on ethno-epistemology. In S. Luper (ed.), *The Skeptics*. Aldershot, U.K.: Ashgate Publishing. 227-247.

Paxton, M., Figdor, C. and Tiberius, V. (in preparation). Quantifying the gender gap: An Empirical study of the underrepresentation of women in philosophy.

Pessin, A. and Goldberg, S. (eds.), (1996). *The Twin Earth Chronicles: Twenty Years of Reflection on Hilary Putnam's "The Meaning of Meaning"*. Armonk, NY: M E. Sharpe.

Plato (1892). The *Republic* in *The Dialogues of Plato*, Vol. I (B. Jowett, Trans.). New York: Random House.

Pizarro, D., Uhlmann, E., & Bloom, P. (2003). Causal deviance and the attribution of moral responsibility. *Journal of Experimental Social Psychology*, 39: 653–660.

Putnam, H. (1973). Meaning and reference. *Journal of Philosophy*, 70: 699-711.

Rigdon, M., Ishii, K., Watabe, M. and Kitayama, S. (2009). Minimal social cues in the dictator game. *Journal of Economic Psychology*, 30: 358-367.

Saul, J. (forthcoming). Unconscious influences and women in philosophy. To appear in F. Jenkins & K. Hutchison (eds.), *Women in Philosophy: What Needs to Change?*

Schaffer, J. and Knobe, J. (forthcoming). Contrastivism surveyed. To appear in *Nous*. doi: 10.1111/j.1468-0068.2010.00795.x

Searle, J. (1980). Minds, brains and programs. *Behavioral and Brain Sciences*, 3: 417–457.

Smart, J. (1973). An outline of a system of utilitarian ethics. In J. Smart and B. Williams (eds.), *Utilitarianism: For and Against*. Cambridge: Cambridge University Press. 3-47.

Sosa, E. (2007). Experimental philosophy and philosophical intuition. *Philosophical Studies*, 132: 99-107.

Sosa, E. (2009). A defense of the use of intuitions in philosophy. In M. Bishop and D. Murphy (eds.), *Stich and His Critics*. Oxford: Blackwell. 101-112.

Starmans, C. and Friedman, O. (2009). Is knowledge subjective? A sex difference in adults' epistemic intuitions. Poster presented at the 6th Biennial Meeting of the Cognitive Development Society, San Antonio, TX, October 16-17, 2009. Abstract available at http://www.cogdevsoc.org/prog2009/CDS09Program.pdf.

Stich, S. (2009). Reply to Sosa. In M. Bishop and D. Murphy (eds.), *Stich and His Critics*. Oxford: Blackwell. 228-236.

Stich, S. (in preparation). Experimental philosophy and the bankruptcy of the great tradition.

Thomson, J. (1971). A defense of abortion. *Philosophy and Public Affairs*, 1, 1: 47-66.

Vesterlund, L. (2006). Why do people give? In R. Steinberg and W. Powell (eds.), *The Nonprofit Sector*, 2nd edition, New Haven: Yale University Press.

Weinberg, J. (2007). How to challenge intuitions empirically without risking skepticism. *Midwest Studies in Philosophy*, 31: 318 – 343.

Weinberg, J., Nichols, S. and Stich, S. (2001). Normativity and epistemic intuitions. *Philosophical Topics*, 29: 429-60.

Williamson, T. (2004). Philosophical intuitions and skepticism about judgment. *Dialectica*, 58: 109-53.

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

Zamzow, J. and Nichols, S. (2009). Variations in ethical intuitions. *Philosophical Issues* 19: 368-388.