

The Self-Reported Moral Behavior of Ethics Professors

Eric Schwitzgebel
Department of Philosophy
University of California
Riverside, CA 92521-0201
eschwitz at domain: ucr.edu

Joshua Rust
Department of Philosophy, Unit 8250
Stetson University
421 North Woodland Boulevard
DeLand, FL 32723
joshuarust at domain: gmail.com

March 16, 2011

The Self-Reported Moral Behavior of Ethics Professors

Abstract:

We examine the self-reported moral attitudes and moral behavior of 198 ethics professors, 208 non-ethicist philosophers, and 167 professors in departments other than philosophy on eight moral issues: academic society membership, voting, staying in touch with one's mother, vegetarianism, organ and blood donation, responsiveness to student emails, charitable giving, and honesty in responding to survey questionnaires. On some issues we also had direct behavioral measures that we could compare with self-report. Ethicists expressed somewhat more stringent normative attitudes on some issues, such as vegetarianism and charitable donation. However, on no issue did ethicists show significantly better behavior than the two comparison groups. Our findings on attitude-behavior consistency were mixed: Ethicists showed the strongest relationship between behavior and expressed moral attitude regarding voting but the weakest regarding charitable donation.

IRB approval: U.C. Riverside HRRB HS-09-002

Word Count: 12,434 (excluding title page, abstract page, footnotes, and figure captions), plus
3,797 words of footnotes and six figures

Keywords: ethics, moral psychology, moral behavior, attitude-behavior consistency,
experimental philosophy, applied ethics, vegetarianism, charity, voting

The Self-Reported Moral Behavior of Ethics Professors

1. Introduction.

Do philosophy professors specializing in ethics behave, on average, any morally better than do other professors? If not, do they at least behave more consistently with their espoused values? These questions have never been systematically studied.

One might find these questions interesting for several reasons, but our primary aim is to understand the relationship between moral reflection, especially the kinds of moral reflection practiced by professional philosophers, and moral behavior. Suppose that professional ethicists behave, overall, somewhat morally better than do comparison groups of professors not specializing in ethics. Suppose, for example, that ethicists donate more of their income to charity or eat a more vegetarian diet or donate blood more frequently than do other professors – and assume, furthermore, that such behavior is morally good. That discovery would seem to invite, though it would not compel, certain optimistic thoughts about the consequences of philosophical moral reflection. Alternatively, suppose that ethicists are no more likely to do any of these things or even, perhaps, less likely. That discovery would seem to invite, though it would not compel, certain pessimistic thoughts about the practical consequences of philosophical moral reflection. Likewise, suppose we found either that ethicists are more likely or that they are less likely to act in accord with their espoused norms than are other professors. That discovery, too, would fit more naturally with some views of the relationship between philosophical reflection and lived behavior than with other views.

Many philosophers, historically, have suggested that studying philosophical ethics – or at least studying it in the right way – will tend to improve moral behavior. Aristotle, for example,

hoped so; he said that the aim of studying ethics was “not the attainment of theoretical knowledge” but rather “to become good” (4th c. BCE/1962, 1103b). The ancient Confucians and Stoics sought personal improvement through philosophical reflection (as emphasized by Ivanhoe 1993/2000 and Hadot 1995). Kant sees moral philosophy as arming us with reason against the corrupting force of our wishes and inclinations (1785/2002, p. 206 [4:405]). Mill portrays philosophical disputation as crucial to maintaining the lively motivational force of ethical doctrines and religious creeds, which otherwise become hollow formularies (1859/2003, p. 118). Peter Singer aims to induce vegetarianism and charitable donation in his readers (e.g., 1972, 1975/2002).

Others have expressed doubts about the practical effects of philosophical moral reflection, at least as normally practiced. Bernard Williams (1985) and Annette Baier (1985) argue that philosophical ethics can undercut our intuitive sense of what’s right. Hans-Georg Moeller (2009) argues that philosophical moralizing tends to promote retributive violence. Richard Posner (1999) and Joshua Knobe and Brian Leiter (2007) argue that philosophical moral reasoning is largely post-hoc rationalization of antecedent opinions and is largely powerless to influence behavior. Despite this difference in focus, the pessimistic and optimistic strains in the tradition can be reconciled if philosophical moral reflection as *ordinarily* practiced, and *standard* philosophical moral education, is toxic or inert but a superior way is available. That is, perhaps, not an uncommon view – though of course there is no consensus about what the superior way might be.

The possibilities are manifold, the relationships complex, and systematic empirical data almost entirely lacking. To help fix ideas, consider as reference points the following four simple hypotheses about the relationship between philosophical moral reflection and practical moral

behavior. By “philosophical moral reflection” we mean simply whatever type of cognition philosophers tend to engage in when they bring their professional tools and competences to bear on moral issues.

Booster view: Philosophical moral reflection leads to the discovery of moral truths – either general moral truths that people tend to not to endorse absent such reflection (such as, perhaps, that eating meat is morally bad: see Section 7 below) or particular moral truths about specific situations that would not otherwise have been properly morally appreciated (such as that some particular behavior toward which one is tempted would be objectionably sexist). Such discoveries have a positive impact on moral behavior – though perhaps only on average, to a moderate extent, and in some areas. Furthermore, since it reveals connections between specific instances of moral behavior and general moral principles, philosophical moral reflection tends to increase the overall consistency between one’s broad moral attitudes and one’s practical moral behavior.

Epiphenomenalist view: Philosophical moral reflection is powerless to change moral behavior or moral attitudes, either for better or for worse. It is behaviorally inert and incapable of generating any new appreciation of moral truths. Whatever causal mechanisms generate our moral attitudes and behavior, they are uninfluenced by our practices of philosophical reflection. Thus, even heroic extremes of philosophical reflection will neither improve one’s moral behavior nor improve the alignment between general normative attitudes and practical behavior.

Rationalization view: Philosophical moral reflection tends to increase the consistency between attitudes and behavior, as the booster suggests, but it does so in the opposite causal direction than the booster suggests: The ethically reflective person’s attitudes shift to match her behavior rather than her behavior shifting to match her attitudes. The reflective person’s

practical behavior may be unaffected by such rationalizations (the *inert rationalization* view); or the tendency to rationalize may morally worsen philosophically reflective people by freeing them to act on immoral impulses that are superficially but unsatisfactorily justified by their reflections (the *toxic rationalization* view). On the inert rationalization view, for example, one will either steal or not steal a library book as a result of psychological processes uninfluenced by one's philosophical reflections, and then one will shape one's moral reflections to justify that incipient or recently past behavior. On the toxic rationalization view, one might feel an inclination to steal the book and act on that inclination as *a consequence of* a spurious moral justification for the theft.

Motivationally ineffectual discovery view: As in the booster view, philosophical moral reflection tends to lead to the discovery of moral truths. However, such discoveries have no consequences for the practical behavior of the person making those discoveries. Philosophical reflection might lead one to discover, for example, that it is morally wrong to eat the meat of factory-farmed mammals, but on this view one would continue to eat factory-farmed meat at exactly the same rate as one would have done absent any philosophical reflection on the matter.

To connect these views with our empirical research on the moral behavior of ethicists requires the following two assumptions, which we regard as empirically open but plausible to a first approximation. First: Professional ethicists tend to engage in philosophical moral reflection, in ways pertinent to practical issues that arise in their lives (i.e., not only at the highest level of abstraction or about far-fetched hypothetical scenarios), more often than do socially similar non-ethicists (and their philosophical moral reflection is at least as skillful). (Alternatively, one might hold that they reflect more skillfully and at least as often.) For example, lecturing on Kant on lying might increase ethicists' propensity toward philosophical reflection when tempted to lie

later in the day; and ethicists might on average more extensively consider the philosophical arguments for and against vegetarianism and charitable donation to famine relief. Second: Professional ethicists are similar to non-ethicist professors (either professors inside or outside of philosophy – we use both comparison groups) in all dimensions relevant to moral behavior *except* for their higher rates of philosophical moral reflection. For example, professional ethicists do not start out morally worse than non-ethicists and then use philosophical moral reflection to improve themselves up to average. With these two background assumptions in hand, we derive the following empirical predictions from the four views described above:

Moral behavior: The booster view predicts that ethicists, since they engage in more philosophical moral reflection than average, will behave morally better, on average, than do socially similar non-ethicists. The epiphenomenalist, inert rationalization, and motivationally ineffectual discovery views predict that ethicists and non-ethicists will behave similarly. The toxic rationalization view predicts that ethicists will behave morally worse.

Moral truth: If there are moral truths (or something similar enough to moral truths; we recognize that the metaethical issues are complex), the booster and motivationally ineffectual discovery views predict that ethicists will be better in touch with those truths than are non-ethicists, as a result of their greater tendency toward philosophical moral reflection. The epiphenomenalist view predicts that ethicists and non-ethicists will be similarly in touch or out of touch with moral truths. On the rationalization view, ethicists' moral judgments might be farther from the moral truth than non-ethicists', if the pressure of moral reflection toward attitude-behavior consistency draws ethicists disproportionately toward erroneous moral attitudes that superficially justify the bad behavior to which we are all tempted.

Attitude-behavior consistency: The booster view and the rationalization view predict that ethicists' higher rates of philosophical moral reflection will lead them to show greater attitude-behavior consistency than do non-ethicists (though the two views predict this result for different reasons). The epiphenomenalist view predicts that ethicists and non-ethicists will show similar levels of attitude-behavior consistency. On the motivationally ineffectual discovery view, ethicists might show less attitude-behavior consistency than do non-ethicists, if ethicists' skills help them escape the everyday self-serving rationalizations that veil immoral behavior behind false principle.

We emphasize that these four views are simplistic and underdeveloped. Many other and more complex views are possible and, given the complexity of the human mind, perhaps also more antecedently probable. These views don't do justice to the nuances of Aristotle, Mencius, Kant, Williams, etc. We present them only as starting points for reflection: With virtually no serious empirical data available on the relationship between philosophical moral reflection and real-world moral behavior, the range of plausible views remains relatively unconstrained. (There have been some studies of the effects of business ethics and medical ethics courses, but never to our knowledge with real-world moral behavior as the outcome measure.) Once data start coming in, we can begin to constrain the models, including possibly ruling out, or at least creating explanatory challenges for, the booster, the epiphenomenalist, the rationalization view, the motivationally ineffectual discovery view, or all four views.

2. An Overview of Our Research.

In several studies, we have begun to empirically explore the moral behavior of ethics professors. So far we have found little evidence that ethicists behave better. However, our

previous studies are neither broad-ranging enough nor univocal enough to create more than prima facie explanatory challenges for the booster, nor have they provided evidence on whether ethicists show lesser or greater attitude-behavior consistency than do socially similar non-ethicists.

At a Pacific Division meeting of the American Philosophical Association, we asked philosophers to anonymously evaluate the overall moral behavior of arbitrarily selected ethicists in their own departments and, for comparison, metaphysics and epistemology specialists in their own departments. The majority of respondents rated the ethicists no better than the M&E specialists (Schwitzgebel and Rust 2009). In another study, we examined publicly available voting records in five U.S. states (on the assumption that voting is a duty), and we found that ethicists and political philosophers voted neither more nor less often than did other professors (though Political Science professors did vote more often; Schwitzgebel and Rust 2010). In other studies, we have found ethicists neither more nor less likely than other professors to answer student emails (Rust and Schwitzgebel in preparation) or to behave courteously at philosophy conferences – though we did find that environmental ethicists left behind less trash in their meeting rooms (Schwitzgebel, Rust, Huang, Moore, and Coates forthcoming). Ethics books also appear to go missing from academic libraries more often than do other philosophy books (Schwitzgebel 2009).

In the research reported below, we extend our explorations into several new domains of behavior, using self-report in a multi-item questionnaire. The disadvantages of self-report, perhaps especially self-report of moral or immoral behavior, are well known; below we will discuss our strategies for addressing these disadvantages. The advantages of self-report are fourfold: First, it allows us to examine behavior on which it is difficult to obtain good non-self-

report data. We were especially interested in rates of vegetarianism and rates of overall charitable donation because those issues are widely discussed by ethicists, who often take explicit and concrete stances, with attitude and behavior highly variable among professors. Thus, vegetarianism and charitable donation seemed like promising areas to find substantial differences among groups' moral attitudes and behavior, if such differences exist. Second, a multi-item self-report questionnaire allows us to examine respondents' responses to several measures simultaneously, potentially revealing trends invisible in single-shot observational studies. Third, moral attitude – as opposed to moral behavior – is for the most part accessible only through self-report. And fourth, since we already possessed data on the actual behavior of some of our respondents, we were able to compare self-reported behavior with actual behavior, enabling three-way comparisons between expressed moral attitude, self-described moral behavior, and directly measured moral behavior; this also yielded a measure of honesty in respondents' survey replies – itself arguably a type of moral behavior.

From February to March, 2009, we sent surveys to 337 ethicists, 329 non-ethicist philosophers, and 314 non-philosophers – all tenured, tenure-track, or emeritus – from selected universities in five U.S. states: California, Florida, North Carolina, Minnesota, and Washington State.¹ We chose these five states because we had voting participation data for those states from a previous study (Schwitzgebel and Rust 2010). We classified philosophy recipients as ethicists or non-ethicists based on research interests stated on recipients' academic homepages. (See Schwitzgebel and Rust 2010 for details of the sampling and classification procedure.) All survey

¹ These numbers exclude 17 ethicists, 33 non-ethicist philosophers, and 44 non-philosophers from the targeted universities for whom we could not obtain valid email or postal addresses or who knew too much about our hypotheses to be useful respondents.

recipients received a preliminary email about the survey, followed a few days later by an emailed link to a web version of the survey (containing a unique tag for each participant). Unless they opted out of further communications, recipients who did not complete the survey received first an email reminder, then a hard copy of the survey with a return envelope, and then a final email reminder. We received returned surveys from 198 ethicists (59% of ethicist recipients), 208 non-ethicist philosophers (63% of recipients), and 167 non-philosophers (53% of recipients), for an overall response rate of 573 out of 980 (58%). Although responses were tagged with ID numbers to enable comparisons of behavioral data and survey response data, no identifying information about individuals, such as name or email address, was stored with the survey responses.

The survey consisted of three sections. The first section contained ten questions about normative attitude on nine moral issues: theft, academic society membership, voting, staying in touch with one's mother, vegetarianism, organ donation, blood donation, responsiveness to student emails, and charity. Nine of the ten questions asked the respondent to rate "the degree to which the action described is morally good or morally bad" by checking one circle on a nine-point scale from "very morally bad" (which we coded as 1) to "very morally good" (coded as 9), with the midpoint labeled "morally neutral" and the 3 and 7 points labeled "somewhat morally bad" and "somewhat morally good" respectively. In the hard copy of the survey, we cautioned respondents that

We recognize that it may be difficult to rate moral goodness and badness on a numerical scale, that different moral goods may be incommensurable, and that the goodness or badness of an action can vary with context. We encourage you to set aside such concerns as best you are able, interpreting the questions below as

straightforwardly as possible. You are welcome to clarify your answers, raise objections to the wording of the questions, etc., in the margins.

The web version of the survey contained the same caution, except concluding “At the end of this section there will be a comment field in which are you are welcome to clarify your answers, raise objections to the wording of the questions, etc.”. We read all such clarifications and critiques, but did not adjust responses in light of them; the clarifications and critiques brought to light no unpredicted systematic problems in the data. Respondents were also permitted to leave questions blank, anywhere in the survey. We chose to solicit responses on a scale of moral goodness rather than, say, a scale of rightness, permissibility, or requirement, because the badness-goodness scale seemed to us conceptually the simplest and the most symmetrical between the positive and negative sides.

The second part of the survey contained questions about the respondent’s own behavior on eight of the nine issues above, including further opportunity to comment in the margins or in web fields. The web version of the survey contained no back button that would have permitted respondents to retrospectively alter their normative responses in light of our questions about their personal behavior, though a few respondents commented that they would have liked to do so. We examined the hard-copy questionnaires for signs of alteration of the normative responses but found little evidence of such alteration. The second part of the survey concluded with a normative question about dishonesty in response to survey questions, followed by a self-report of

honesty or dishonesty. Despite the potential sensitivity of questions about respondents' personal behavior, respondents skipped very few questions in our survey.²

Philosophy respondents' questionnaires (both the ethicists' and the non-ethicists') contained a third part consisting of two questions. The first asked the respondent to report the level of abstraction at which she tends to consider ethical issues (metaethics, normative ethics, applied ethics, no ethics-related area among my specializations). The second asked the respondent to report what normative ethical view she finds broadly most appealing (deontological, consequentialist, virtue ethical, skeptical, no settled position).

Half of the survey recipients were given a charity incentive to complete the survey. We offered to donate \$10 to a charity of their choice from among Oxfam America, World Wildlife Fund, CARE, Make-a-Wish Foundation, Doctors Without Borders, or American Red Cross. (We kept our promise, donating a total of \$2,880 dollars to the selected charities.) The charity selection was the last question of the survey, set aside in a section of its own. The charity incentive was partly intended to enhance response rates. However, equally importantly, it allowed us to look for group differences in responsiveness to the charity incentive and in charity-response stability. (If a group gives similar responses to the charity questions regardless of whether they received the charity incentive, their responses are stable between the versions; if not, their responses are unstable.)

A full facsimile of the survey is available online at ****.

² Overall, only 25 respondents (4%) dropped out of the survey, completing the first question but reaching a point after which they completed no more questions. Most questions had response rates in the high 90-percents, typically 98%-100% excluding dropouts.

Thus, our study is multi-faceted – illuminatingly so, we hope. Proper presentation and analysis of the results will necessarily be somewhat lengthy.

3. Theft of \$1000.

As mentioned above, the survey began with ten normative questions about nine moral issues. Nine of the ten questions asked respondents to rate various actions on a nine-point scale from “very morally bad” (1) to “very morally good” (9), with “morally neutral” (5) in the middle. The first of these questions was “stealing \$1000 from a house where you are staying as a guest”. This first question was intended to help anchor the morally bad end of the scale, providing respondents with an implicit comparison point for their other responses and aiding interpretation of the scale. As expected, virtually all respondents rated this action on the morally bad side of the scale. All 573 respondents answered this question, and only one of those 573, an ethicist, rated it above 3 on our scale (that respondent rated it 5, “morally neutral”). 76% of respondents rated this action at the extreme end of the scale, that is, 1, “very morally bad”, and 96% rated the action either 1 or 2.

Non-philosophers were significantly more likely to rate the action at the extreme endpoint of the scale (92% rated it 1) than were either ethicists or non-ethicist philosophers (74% and 65% respectively; for ease of exposition, details of statistical tests will be confined to footnotes).³ This might reflect a real difference in opinion about the moral badness of theft, but

³ Philosophers 70% vs. non-philosophers 92%, $\chi^2 = 31.2$, $p < .001$. Ethicists 74% vs. non-ethicist philosophers 65%, $\chi^2 = 3.8$, $p = .052$. Unless otherwise specified, we employ an alpha level of .05 for statistical significance and .10 for statistical marginality.

we're inclined to think that it mostly reflects a scaling issue: Philosophers, accustomed to being presented with murderous moral thought experiments, may have been more likely than non-philosophers to have reserved the extreme bad end of the scale for particularly heinous deeds. Because of such scaling issues, as well as antecedent concerns about the meaningfulness of small-to-moderate differences in mean response on our scales, we emphasize qualitative differences in analyzing the normative questions – differences, that is, between rating the action anywhere on the morally bad side of the scale (1-4) vs. rating it as morally neutral (5) vs. rating it as morally good (6-9).

We did not ask for self-report of theft.

4. Membership in Academic Societies.

The second normative question asked respondents to rate “regularly paying membership dues to support one’s main academic disciplinary society (the APA, the MLA, etc., as appropriate)” on our nine-point scale from “very morally bad” to “very morally good”. We deliberately phrased this question somewhat leadingly, in terms of “support” by paying dues, to highlight the potential moral dimension of belonging to a disciplinary society. The three groups did not detectably differ in the percentage rating societal membership as morally good: 66% of ethicists did so, compared to 69% of non-ethicist philosophers and 65% of non-philosophers.⁴ Neither did the groups differ much in their mean responses on the scale: 6.3 for ethicists vs. 6.3

⁴ $\chi^2 = 0.9$, $p = .63$ (collapsing neutral and bad).

for non-ethicist philosophers and 6.6 for non-philosophers.⁵ Only three respondents rated membership on the bad side of the scale (two ethicists, one non-ethicist philosopher).

Part II of the questionnaire was labeled “Behavior” and began with the statement “The following questions pertain to YOUR OWN behavior on the issues previously presented.” The first question in Part II was

11. Are you currently a dues-paying member of your discipline’s main academic society?”

- yes
- no
- don’t recall

Non-philosopher respondents were marginally more likely to report societal membership: 87% did so, compared to 78% of ethicists and 78% of non-ethicist philosophers.⁶

Before distributing the surveys, we had gathered data on philosophy recipients’ membership in the American Philosophical Association by examining the APA membership list. (We emphasize that all data were coded to mask the identities of particular individuals and that all aspects of our research were approved in advance by the U.C. Riverside Human Research Review Board.) 68% of survey respondents appeared on the membership list (67% of ethicists,

⁵ The difference in mean rating is statistically significant by ANOVA ($F = 5.2, p = .006$), but given the scaling issues mentioned in Section 3, we don’t regard differences of less than 0.5 on the nine-point scale as meaningful – at least when the non-philosophers’ means are farthest from center.

⁶ $\chi^2 = 5.6, p = .06$; merging the two philosophy groups, $\chi^2 = 5.6, p = .02$; excluding 3 respondents answering “don’t recall”.

69% of non-ethicist philosophers), somewhat lower than the self-reported rate of 78%.

Assuming that the APA membership lists accurately reflect membership (there may be some small divergences), 44% of ethicist non-members falsely reported membership, as did 39% of non-ethicist members (not statistically different rates⁷) – seemingly reflecting a bias toward “socially desirable responding”, that is, toward giving answers that would appear to cast the respondent or the respondent’s group in a good light. In Section 11 below, we will discuss the general issue of response accuracy or honesty and whether it varies among the groups.

As one might expect, survey respondents – perhaps because they are more responsive to professional inquiries or more active in the research side of the profession – were more likely to be APA members than were non-respondents: 68% of survey respondents were APA members, compared to 52% of non-respondents.⁸ *Non-response bias* is the tendency for survey respondents to differ from survey non-respondents in one of the dimensions of interest. Accordingly, we note that positive attitudes toward societal membership may be somewhat overrepresented among survey respondents.

Among ethicists who rated membership on the morally good side of our nine-point scale, 75% were members (as measured by presence on the APA list), compared to 52% of ethicists who did not rate membership on the morally good side of the scale; for non-ethicist philosophers,

⁷ $\chi^2 = 0.3$, $p = .56$. Also, 6% of respondents underreported membership, perhaps some of which were emeritus members no longer paying dues but still entitled to appear on the membership list.

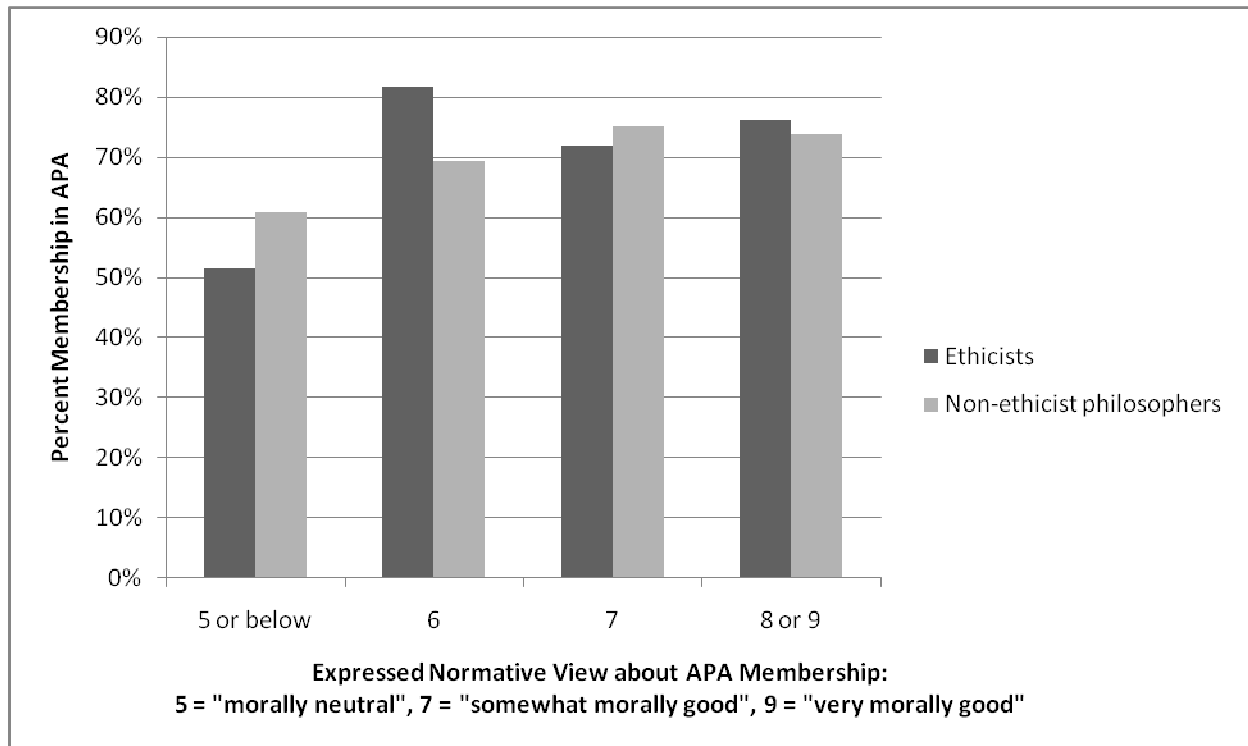
⁸ $\chi^2 = 18.7$, $p < .001$; ethicists 67% vs. 54%, $\chi^2 = 6.0$, $p = .01$; non-ethicist philosophers 69% vs. 49%, $\chi^2 = 13.6$, $p < .001$.

the corresponding spread was 73% vs. 61%.⁹ Thus, ethicists tended to show a stronger attitude-behavior relationship than did non-ethicist philosophers. However, this trend did not approach statistical significance.¹⁰ Figure 1 displays these results graphically. In Section 14, we will address the issue of whether ethicists show more consistency overall between their expressed attitudes and their directly measured or self-reported behavior.

⁹ Ethicists: $\chi^2 = 11.1$, $p = .001$ (for self-reported membership, $\chi^2 = 9.7$, $p = .002$); non-ethicist philosophers: $\chi^2 = 3.0$, $p = .09$ (for self-reported membership, $\chi^2 = 2.0$, $p = .15$).

¹⁰ We compare strength of attitude-behavior relationship by converting the two-by-two tables to correlations, then using Fisher's z_r conversion to test for differences in effect size (Rosenthal and Rosnow 1984/2008): ethicists $r = .24$ vs. non-ethicist philosophers $r = .12$, $p = .22$.

FIGURE 1: Relationship of expressed normative view about “regularly paying membership dues to support one’s main academic disciplinary society (the APA, the MLA, etc., as appropriate)” and appearance on the American Philosophical Association’s membership list. Increasing bar height from left to right reflects correlation between behavior and normative view.



In sum: Regarding societal membership, we found no substantial differences between ethicists and non-ethicist philosophers in normative attitude, self-reported behavior, directly measured behavior, or accuracy of response (which was low for APA non-members); there was a trend, but not statistically significant, for ethicists to show a stronger attitude-behavior relationship. Non-philosophers were somewhat more likely than philosophers to self-report membership in their disciplinary societies.

5. Voting in Public Elections.

Our third normative question asked respondents to rate “regularly voting in public elections” on our usual nine-point scale from “very morally bad” to “very morally good”. Later, in Part II, we asked three self-report questions on voting behavior:

12. Are you a United States citizen eligible to vote?

yes

no

13. How many public elections (including local, state, and national) would you estimate you have voted in since Jan. 1, 2000 (including foreign elections if not a U.S. citizen)?

enter a number _____

14. Did you vote in the November, 2008, U.S. general Presidential election?

yes

no

don't recall

The response groups did not detectably differ in the percentage who rated regularly voting as morally good: 90% of ethicists did so, compared to 85% of non-ethicist philosophers and 89% of non-philosophers.¹¹ (Only one respondent, a non-ethicist philosopher, rated voting as bad; the remaining 12% rated voting as morally neutral.) The groups did differ in mean response on the scale: Non-philosopher respondents gave a mean response of 7.8 on the nine-point scale, significantly higher than the 7.3 mean for ethicists and 7.1 for non-ethicist philosophers, but we would interpret this difference cautiously due to the scaling issues discussed in Section 3.¹²

Non-ethicist philosophers were somewhat less likely than either ethicists or non-philosophers to report being a U.S. citizen eligible to vote (84% vs. 91% and 93% respectively).¹³ Excluding non-citizens, all three groups reported voting at about the same rate: 9.2 times since the year 2000 for ethicists, compared to 9.0 times for non-ethicist philosophers and 9.8 times for non-philosophers.¹⁴ Virtually all respondents who reported being eligible to

¹¹ $\chi^2 = 2.4$, $p = .30$ (collapsing neutral and bad).

¹² ANOVA, $F = 16.0$, $p < .001$. Tukey pairwise comparisons find non-philosophers different from both ethicists and non-ethicist philosophers but finds no difference between ethicists and non-ethicist philosophers.

¹³ $\chi^2 = 8.2$, $p = .02$. All but four respondents who reported zero votes also reported being ineligible to vote in the U.S. In Schwitzgebel and Rust 2010 we report evidence from academic homepages that also suggests lower citizenship rates among non-ethicist philosophers.

¹⁴ ANOVA, $F = 0.7$, $p = .51$. All vote rate data is square-root transformed for a better approximation to normality.

vote in the U.S. also reported voting in the 2008 U.S. general Presidential election: an incredible 98% – all but six of the respondents. No respondent answered “don’t recall”.

For many of the survey respondents we had identity-masked voting participation data, derived from state records, from a previous study (Schwitzgebel and Rust 2010). We updated those data for the present study, using a coding scheme that converted names to unique identifiers (thus preventing us from drawing inferences about particular voters’ survey responses). Among survey recipients, the mean number of votes per year in state records was virtually identical for all groups: 1.09 for the ethicists, 1.12 for the non-ethicist philosophers, and 1.11 for the non-philosophers.¹⁵ The *variance* in voting rates was also virtually identical: Ethicists did not show a broader spread of voting habits (e.g., with an overrepresentation of ultra-conscientious voters balanced against an overrepresentation of merely sporadic voters).¹⁶

Survey respondents voted somewhat more frequently than non-respondents: 1.17 vs. 1.03 votes per year.¹⁷ Voter participation and survey participation might be related for any of a variety of reasons, such as illness, time demands, interest in having one’s voice heard, or primary

¹⁵ ANOVA, $F = 0.3$, $p = .76$, excluding recipients whose first state-recorded vote was in 2008 and recipients without a good name match in the voter rolls (see Schwitzgebel and Rust 2010 for methodological details).

¹⁶ SD .22 for ethicists, .22 for non-ethicist philosophers, and .21 for non-philosophers (Bartlett’s test for equal variances, $p = .49$).

¹⁷ T-test, $t = 3.9$, $p < .001$. The non-philosophers had the lowest response rates and correspondingly the largest non-response bias in voting rates: 1.20 vs. 1.01 (T-test, $t = 2.9$, $p = .005$).

residence in a different locale. We find the relatively small difference in vote rate between respondents and non-respondents encouraging on the issue of non-response bias.

Despite a tendency for Americans in general to substantially overreport voting (Bernstein, Chadha, and Mountjoy 2001; Holbrook and Krosnick 2010), we found no such tendency among our respondents. Virtually all of the respondents for whom we could find voting records and who reported voting in the 2008 U.S. general Presidential election did in fact vote in that election, reflecting an astounding 96% participation rate overall (compared to 62% participation among eligible U.S. citizens generally).¹⁸ Though the numbers are somewhat small, respondents for whom we had full voting records since the year 2000 tended to slightly underreport voting, reporting having voted in an average of 9.2 elections since the year 2000, while voter records show them actually to have voted in an average of 10.6 elections.¹⁹ (This underreporting might reflect difficulty remembering local elections and runoffs.)

¹⁸ U.S. participation rates from http://elections.gmu.edu/Turnout_2008G.html (accessed March 1, 2011). Only three ethicists, five non-ethicist philosophers, and two non-philosophers appeared to falsely report voting in the 2008 U.S. general election; Fisher's exact test, $p = .62$.

¹⁹ Paired-t, $t = -3.1$, $p = .003$. Only 88 respondents had full voting records since the year 2000, which required our having antecedently found an unambiguous name match in the states' voter database, that voter's continuous registration in that state since the year 2000, and complete state records for that voter since the year 2000 (this last criterion excluded all California respondents, since California did not provide data for local elections). Eight of the 30 ethicists for whom we had full voting records since the year 2000 overreported voting (27%), compared to seven of 24 non-ethicist philosophers (29%) and 15 of 34 non-philosophers (44%); given the

Among the 88% of respondents who expressed the view that voting is morally good, all three groups voted at virtually the same rate according to state records: 1.19 votes per year for ethicists, compared to 1.17 votes per year for both the other groups.²⁰ The 12% who rated voting morally neutral trended only slightly toward voting less often: 1.11 votes per year on average (with numbers too small for meaningful statistical comparison between ethicists and non-ethicists).²¹ However, ethicists showed a significant relationship between expressed normative attitude on the 1-9 scale and state-reported voting frequency, with a correlation of .36 – compared to only .14 for the non-ethicist philosophers and .01 for the non-philosophers.²² In other words, non-philosophers' expressed attitude toward voting was unrelated to their measured voting rate, whereas ethicists who expressed a strongly positive normative attitude toward voting tended to vote more frequently than ethicists who did not. Figure 2 displays the results.

small numbers involved this difference does not approach statistical significance ($\chi^2 = 2.5$, $p = .28$).

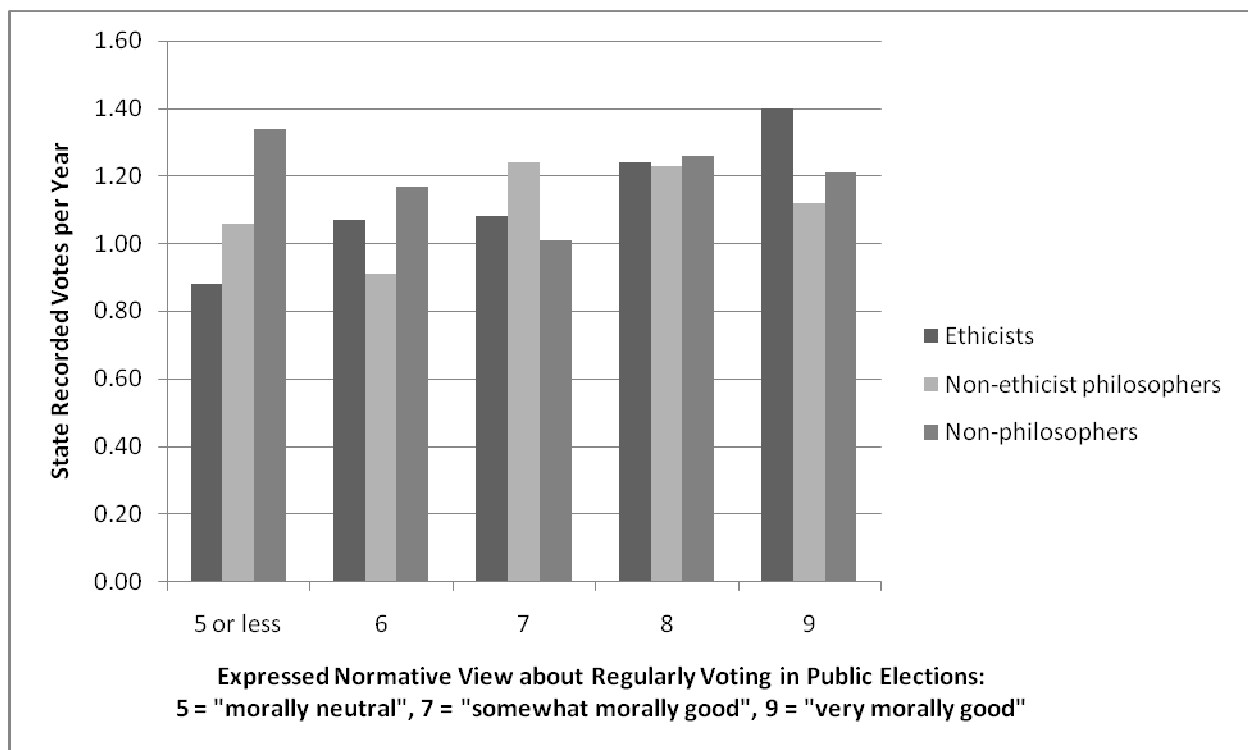
²⁰ ANOVA, $F = 0.1$, $p = .93$.

²¹ $t = 0.8$, $p = .42$.

²² Pearson's r ; $p < .001$, .18, and .91 respectively; merging all groups, $r = .18$, $p = .003$.

Correlations compared using Fisher's z_r conversion: ethicists vs. non-ethicist philosophers, $p = .08$ (marg. sig. two-tailed); ethicists vs. non-philosophers, $p = .01$.

FIGURE 2: Relationship of expressed normative view about “regularly voting in public elections” and voting rates derived from public records. Increasing bar height from left to right reflects correlation between behavior and normative view. Notice that only the darkest bars, for the ethicists, show this trend.



In sum, the groups differed little in overall normative view and had virtually identical voting behavior overall, but ethicists showed a stronger attitude-behavior relationship than the other groups, with a substantially higher observed correlation between state records of voting behavior and the strength of their expressed normative view.

6. Staying in Touch with Mom.

The fourth question of the survey asked respondents to rate “not keeping in at least monthly face-to-face or telephone contact with one’s mother” on the usual 1 to 9 scale from “very morally bad” to “very morally good”. Self-report of behavior occurred in Questions 15 and 16:

15. Over the last two years, about how many times per month on average have you spoken with your mother (face-to-face or on the phone)? (If your mother is deceased, consider how often you spoke during her last two years of life.)

- once (or less) every 2-3 months
- about once a month
- 2-4 times a month
- 5 times a month or more

16. About how many days has it been since your last face-to-face or phone contact with your mother?

enter a number in days (if mother deceased enter ‘999’) _____

The reader will be unsurprised to learn that for this question we had no direct measures of behavior. As with many of our behavioral self-report questions, we deliberately asked first a general question that would be hard to answer accurately and easy to fudge and a more specific

question that would be easier to answer and harder to fudge. We did this as a way of checking for a bias toward socially desirable responding, which we expected to reveal itself in disparities between the general and specific answers (see especially Section 7 below).

Similar percentages of all three groups rated not keeping in regular contact with one's mother on the bad side of our scale: 73% of ethicists, 74% of non-ethicist philosophers, and 71% of non-philosophers.²³ (3% rated it as morally good; our mothers tell us they mistakenly flipped the scale due to the negative phrasing of the question.) The groups also gave similar mean responses: mean 3.4 for ethicists, compared to 3.7 for non-ethicist philosophers and 3.3 for non-philosophers.²⁴

Non-philosophers reported the most frequent contact. If we accept the norm of at least monthly contact, non-philosophers reported the least counternormative behavior on the general question, with only 5% reporting contact less than once a month, compared to 11% of ethicists and 12% of non-ethicist philosophers. 83% of non-philosophers reported contact at least twice a month, compared to 70% of ethicists and 74% of non-ethicist philosophers.²⁵ On the specific question about days since contact, differences were more subtle, and not statistically significant, but in the same direction: 8% of non-philosophers reported that it had been more than 30 days since their last contact, compared to 10% of ethicists and 11% of non-ethicist philosophers

²³ $\chi^2 = 0.5$, $p = .79$ (collapsing neutral and good).

²⁴ These means are statistically different (ANOVA, $F = 4.3$, $p = .01$), but due to the scaling concerns discussed above we do not interpret differences of less than 0.5 on the scale as meaningful when non-philosophers' means are the farthest from center.

²⁵ $\chi^2 = 5.6$, $p = .06$ (marg. sig.); $\chi^2 = 8.0$, $p = .02$.

(excluding respondents reporting deceased mothers).²⁶ The log-transformed mean of the reported days since last contact plus one was 5.0 for the non-philosophers, 5.9 for the ethicists, and 6.7 for the non-ethicist philosophers.²⁷ These differences cannot be explained by group differences in age (there were none) or in gender (which did not relate to reported maternal contact). See Section 13 for more demographic analysis.

Like the voting question but unlike the societal membership question, respondents did not appear, overall, to show a social desirability bias – at least insofar as such a bias would emerge as a difference between the fudgeable general question and the presumably less fudgeable recent-instance question. Respondents reporting contact about once a month should, on average, report about 15 days since last maternal contact; they actually reported a log-transformed mean plus one of 12 days (38% reporting more than 14 days); those reporting 2-4 contacts per month had a mean of 5 days; and those reporting 5 or more had a mean of 3 days; with all groups similar in this regard.²⁸

Non-ethicist philosophers showed the strongest relationship between expressed normative attitude and self-reported behavior. Among non-ethicist philosophers who expressed the view that it is bad not to keep in at least monthly contact with one's mother, 92% reported contact within the past 14 days, while among non-ethicist philosophers who rated it as neutral or good 57% reported contact within the past 14 days, a differential of 35%.²⁹ For ethicists, the corresponding percentages were 84% vs. 77% (differential 7%); for non-philosophers 91% vs.

²⁶ $\chi^2 = 0.4$, $p = .82$.

²⁷ ANOVA, $F = 1.2$, $p = .31$. Log-transformed mean: $e^{\text{mean}(\ln(x_i))}$.

²⁸ ANOVAs, $F \leq 1.5$, $p \geq .24$.

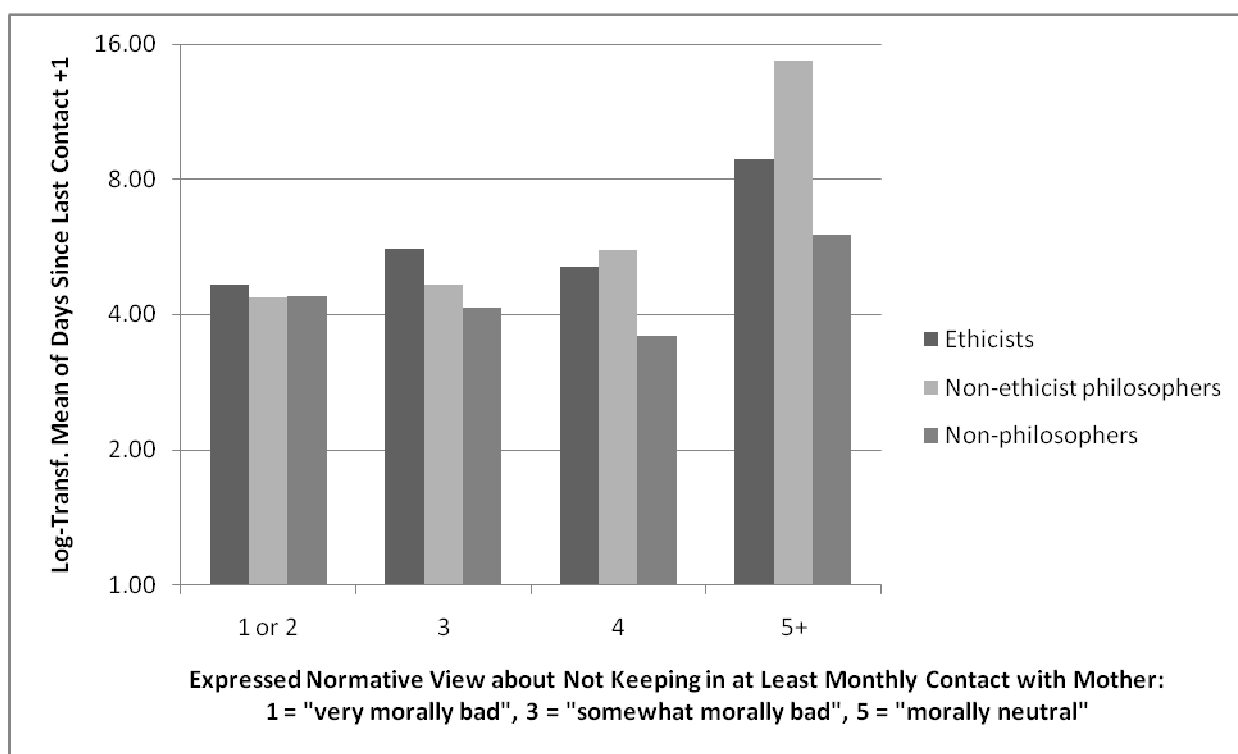
²⁹ $\chi^2 = 20.5$, $p < .001$, excluding respondents reporting deceased mothers.

86% (differential 15%).³⁰ Likewise, non-ethicist philosophers showed a much larger correlation between normative response on the 1-9 scale and days since last reported contact: .45, compared to .13 for ethicists and .09 for non-philosophers.³¹ Figure 3 displays these data graphically.

³⁰ Respectively, $\chi^2 = 0.9$, $p = .35$; $\chi^2 = 4.1$, $p = .04$.

³¹ Pearson's r , $\ln(x+1)$ -transformed, $p < .001$, .17, and .37 respectively; merging all groups, $r = .26$, $p < .001$. Correlations compared using Fisher's z_r conversion: non-ethicist philosophers vs. ethicists, $p = .008$; non-ethicist philosophers vs. non-philosophers, $p = .004$. The differences are more subtle, and not statistically significant, using the vaguer and presumably less valid (though higher N) measure of self-reported rate of monthly contact.

FIGURE 3: Relationship of expressed normative view about “not keeping in at least monthly face-to-face or telephone contact with one’s mother” and log-transformed mean of number of days, plus one, since self-reported last face-to-face or telephone contact. Increasing bar height from left to right reflects correlation between behavior and normative view. Notice that non-ethicist philosophers show the strongest correlation. Note also that non-philosophers show the lowest mean (given approximately equal distribution among the attitudinal groups).



In sum, on this measure the non-philosophers reported the most maternal contact and the non-ethicist philosophers showed the strongest relationship between expressed normative view and self-reported behavior, a pattern of results predicted by none of the simple models offered in the introduction.

7. Vegetarianism.

Question 5 asked respondents to rate “regularly eating the meat of mammals such as beef or pork” on the usual 1-9 normative scale. We asked about the meat of mammals to avoid confusing respondents who see important moral differences between mammals and other animals, but we saw no way to respect everyone’s moral distinctions (e.g., the distinction between factory farmed and humanely raised meat) without invoking further possibly confusing distinctions. (Indeed, similar considerations apply to all our normative questions; we chose to err on the side of simplicity.) Questions 17 and 18 asked respondents to report their own consumption of mammal meat:

17. During about how many meals or snacks per week do you eat the meat of mammals such as beef or pork?

enter number of times per week ____

18. Think back on your last evening meal, not including snacks. Did you eat the meat of a mammal during that meal?

yes

no

don’t recall

We use the term *vegetarianism* to refer to refraining from eating the meat of mammals. By *meat* we mean to include only the meat of mammals.

On this normative question, unlike the previous ones, the groups differed enormously in their responses: 60% of ethicists rated meat-eating on the bad side of the scale, compared to 45% of non-ethicist philosophers and only 19% of non-philosophers.³² (Approximately 5% of respondents in all groups rated it as good.) Differences in qualitative response were not as large but still large enough to be meaningful (perhaps especially since they run opposite what would be predicted from the scaling concerns discussed in Section 3): ethicists' mean 4.0, non-ethicist philosophers' 4.2, and non-philosophers' 4.7.³³

In contrast, when asked about the previous evening meal, the groups' responses were only marginally statistically different and ethicists were not even in the lowest group: 37% of ethicists reported eating meat at their last evening meal, compared to 33% of non-ethicist philosophers and 45% of non-philosophers.³⁴ Self-reported meals per week differed a little more, with a mean of 4.1 self-reported meals per week for ethicists, compared to 4.6 for non-ethicist philosophers and 5.3 for non-philosophers.³⁵ Strict vegetarianism seems morally and psychologically rather different from merely low meat consumption, so we also divided the

³² $\chi^2 = 64.2$, $p < .001$ (collapsing neutral and good).

³³ ANOVA, $F = 17.0$, $p < .001$, Tukey pairwise comparisons finding both philosopher groups different from the non-philosophers but not from each other.

³⁴ $\chi^2 = 5.7$, $p = .06$, excluding two respondents who stated that they didn't recall; comparing just ethicists vs. non-philosophers, $\chi^2 = 2.4$, $p = .12$.

³⁵ ANOVA (square-root transformed), $F = 5.2$, $p = .006$, reducing one response from "500" (marginal comment "Bacon tastes good") to 25.

meals/week responses into zeros and non-zeros. Here we did see a substantial group difference: 27% of ethicists reported eating no mammal meat at all, compared to 20% of non-ethicist philosophers and 13% of non-philosophers.³⁶

We found a bias toward underreporting meat consumption among respondents reporting eating meat at 1-3 meals per week. 21% of respondents who reported eating meat at one meal per week reported eating meat at the previous evening meal. Even if we assume that meat is *only* consumed at evening meals, the number should be closer to 14% (1/7); if we assume, perhaps more plausibly, that approximately half of all meat meals are evening meals, then the number should be closer to 7%. Similarly, 30% of respondents who reported two meat meals per week reported meat at the last evening meal (vs. a target of 14%-29%), as did 51% of respondents who reported three meals/week (vs. a target of 21%-43%). This apparent bias toward socially desirable responding did not differ between the groups: 18 of 54 ethicists in the 1-3 meals/week group reported eating meat the previous night (33%), compared to 18 of 52 non-ethicist philosophers (35%) and 14/40 non-philosophers (35%).³⁷ Two of the 110 respondents who reported 0 meals/week also reported having eaten meat at the previous evening meal (one ethicist, one non-ethicist philosopher).

Expressed normative attitude and self-reported behavior were significantly related for all three groups. Overall, 27% of respondents who rated eating meat as bad reported eating it at the last evening meal, compared to 47% of respondents who rated eating meat as neutral or good. The spread was 52% vs. 27% for ethicists, 37% vs. 28% for non-ethicist philosophers, and 51%

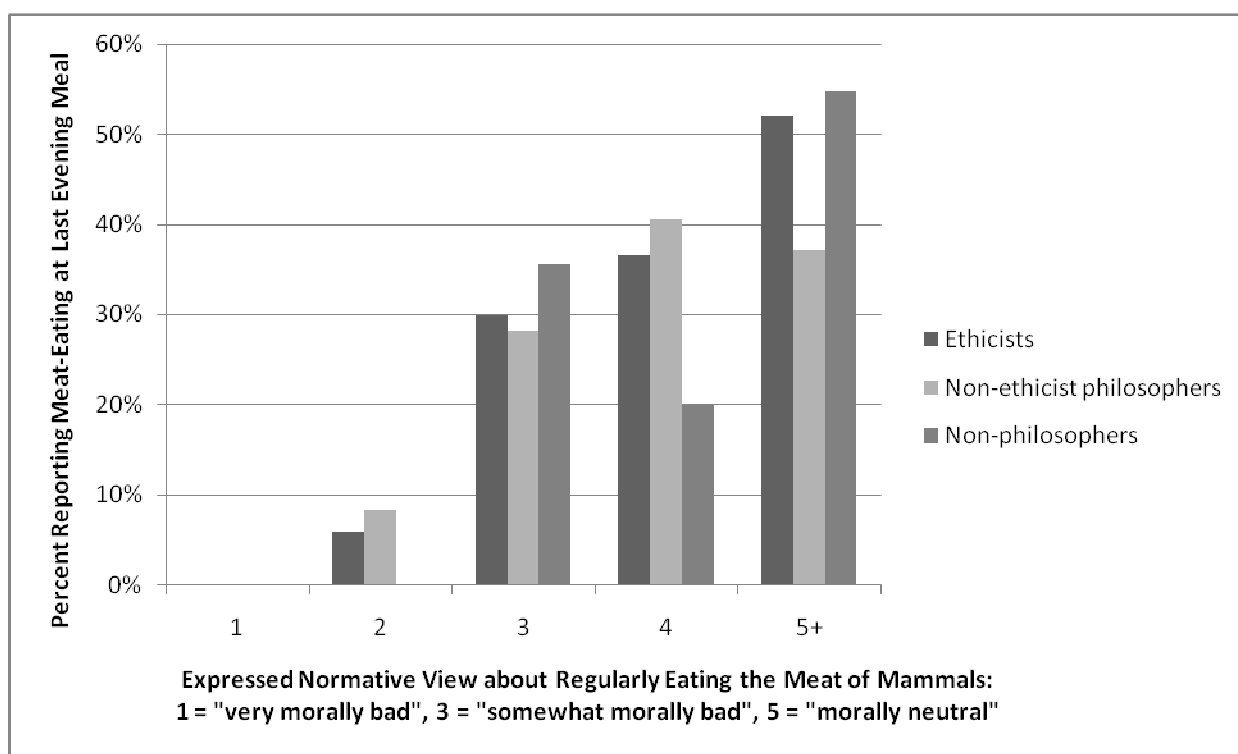
³⁶ $\chi^2 = 9.3$, $p = .01$. Only one respondent, an ethicist, reported between zero and one meals per week (.5, classified as non-vegetarian).

³⁷ $\chi^2 = 0.0$, $p = .98$.

vs. 23% for non-philosophers; and strength of expressed normative attitude predicted low likelihood of reporting meat-eating at the last evening meal as is evident from Figure 4.³⁸

³⁸ Bad vs. not bad: overall $\chi^2 = 24.5$, $p < .001$; ethicists $\chi^2 = 12.0$, $p = .001$; non-ethicist philosophers $\chi^2 = 1.6$, $p = .20$ (but note that a one-tailed p-value might be more appropriate for this test); non-philosophers $\chi^2 = 8.1$, $p = .005$; effect size comparison using Fisher's z_r conversion: r values .26, .09, .22; pairwise, the ethicists vs. non-ethicist philosophers difference was marginally significant, $p = .09$. 1-9 scale: overall Pearson's $r = .24$, $p < .001$; ethicists $r = .28$, non-ethicist philosophers $r = .17$, non-philosophers $r = .24$ (group p 's $\leq .02$, effect sizes not significantly different).

FIGURE 4: Relationship of expressed normative attitude about “regularly eating the meat of mammals such as beef or pork” and reporting having eaten the meat of a mammal at the last evening meal. Increasing bar height from left to right reflects correlation between self-reported behavior and normative view. Note the substantial relationship for all groups. Note also: The “1” column appears empty because none of the twelve respondents giving that answer to the normative question reported having eaten the meat of a mammal at the last evening meal.



In sum, ethicists were much more likely than non-philosophers to rate eating the meat of mammals on the bad side of the scale, while non-ethicist philosophers were intermediate. Ethicists were also substantially more likely to self-report vegetarianism (that is, zero meat meals per week). However, when asked about their last evening meal, ethicists reported eating meat at approximately the same rate as did the other groups. All groups showed at least a moderate relationship between expressed normative view and self-reported behavior.

8. Organ and Blood Donation.

Question 6 asked respondents to rate “regularly donating blood” on our usual 1-9 scale, and Question 7 asked respondents to rate “not having on one’s driver’s license a statement or symbol indicating willingness to be an organ donor in the event of death”. The behavioral self-report questions were:

19. Please look at your driver’s license and indicate whether there is a statement or symbol indicating your willingness to be an organ donor in the event of death.

- yes
- no
- driver’s license not available

20. About how many times per year do you donate blood?

- twice a year or more
- about once a year
- once every few years
- rarely or never
- not eligible to donate blood

21. When was the last time you donated blood?

enter an approximate date (skip this question if you have never donated)

MM / DD / YYYY

Philosophers, especially ethicists, were more likely to rate donating blood as good and not being an organ donor as bad: 84% of ethicists rated donating blood as good, compared to 80% of non-ethicist philosophers and 72% of non-philosophers. And 60% of ethicists rated not having a donor symbol as bad, compared to 56% of non-ethicist philosophers and 42% of non-philosophers.³⁹

Despite this difference in normative view, self-reported behavior was virtually identical among the three groups. 68% of ethicists, 65% of non-ethicist philosophers, and 69% of non-philosophers reported having a donor symbol on their driver's license (excluding six respondents who reported an unavailable driver's license).⁴⁰ Very few from any group reported donating twice a year or more (4% of ethicists, 5% of non-ethicist philosophers, 2% of non-philosophers) or even about once a year (4%, 5%, 5%). Most reported donating rarely or never (42%, 42%, 39%) or being ineligible (38%, 38%, 39%). The percentage reporting having donated in 2008 or

³⁹ $\chi^2 = 8.7$, $p = .01$ (collapsing neutral and bad); $\chi^2 = 13.2$, $p = .001$ (collapsing neutral and good). 5% of respondents rated donating blood as bad and 2% rated not having a donor symbol as good. ANOVA finds differences in mean response for the organ question (ethicists mean 4.0 vs. non-ethicist philosophers 4.1 vs. non-philosophers 4.4, $F = 5.5$, $p = .004$) but not the blood question (6.8 vs. 6.7 vs. 6.9, $F = 0.4$, $p = .65$).

⁴⁰ $\chi^2 = 0.6$, $p = .75$.

2009 (excluding ineligible) was 13% of ethicists, 14% of non-ethicist philosophers, and 10% of non-philosophers.⁴¹

We saw a small social desirability bias in reporting, insofar as such a bias can be inferred from differences between the fudgeable general question and the presumably less fudgeable recent-instance question. Respondents reporting blood donation twice a year or more reported a mean most recent donation of 2008.5, those reporting about once a year reported a mean of 2007.4, and those reporting once every few years reported a mean of 2004 – compared to expected means of about 2009.0, 2008.7, and 2006.⁴² The numbers were too small for meaningful comparison of groups' bias toward socially desirable responding.

The non-ethicist philosophers showed the strongest relationship between expressed normative view and self-reported behavior on the organ donation question but trended toward showing the weakest relationship on the blood donation questions.⁴³ Figure 5 visually displays

⁴¹ $\chi^2 = 0.8$, $p = .67$.

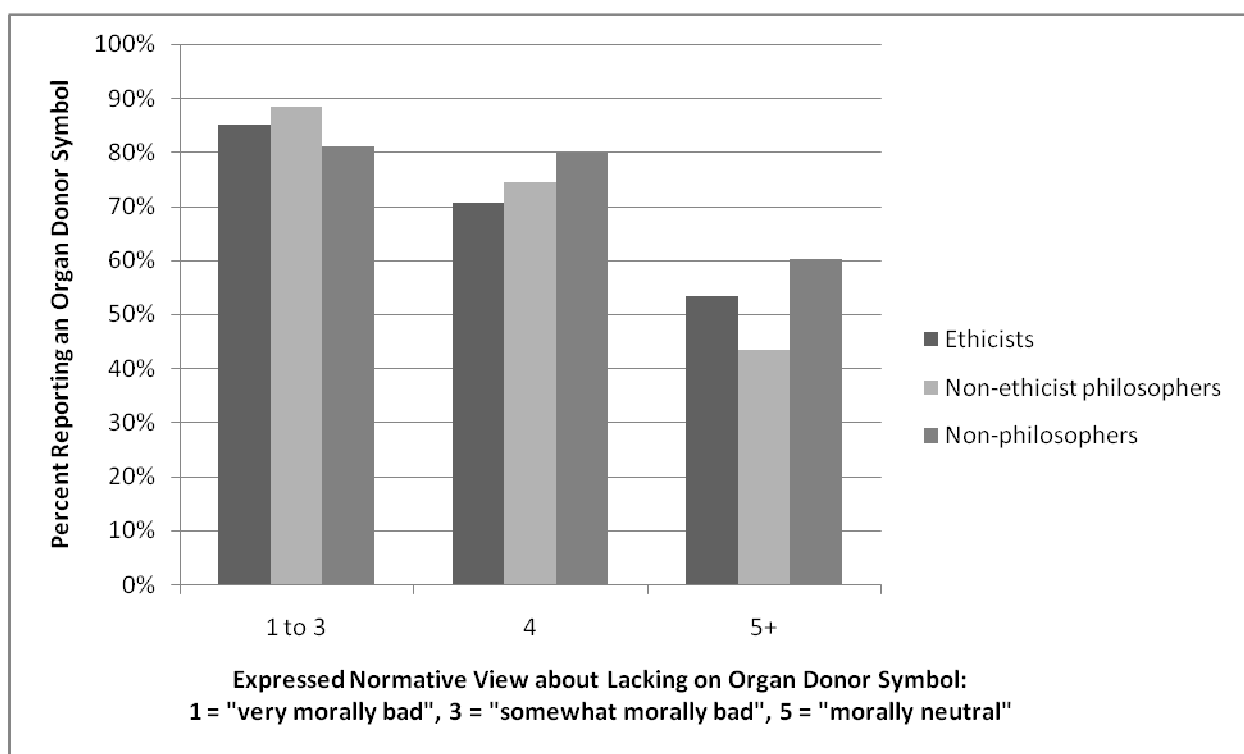
⁴² The last two comparisons are statistically significant by the one-sample t-test: $t = 4.7$, $p < .001$; $t = 2.0$, $p = .046$.

⁴³ Among non-ethicist philosophers, 82% of those who rated it as bad not to have a donor symbol reported having a donor symbol, compared to 44% who rated it as neutral or good. The split was 79% vs. 53% for non-ethicist philosophers and 81% vs. 60% for non-philosophers; $\chi^2 = 13.2$, $p < .001$; $\chi^2 = 31.7$, $p < .001$; $\chi^2 = 7.5$, $p = .006$. Effect size comparison using Fisher's z_r conversion: r values .27, .41, .22; pairwise, non-ethicist philosophers vs. non-philosophers $p = .048$, non-ethicist philosophers vs. ethicists $p = .13$. T-tests comparing mean normative view between donors and non-donors yield similar results with greater statistical significance for the effect size comparison: non-ethicists philosophers mean 3.8 vs. 4.6 ($t = 6.5$, $p < .001$) compared

the organ donation results. Due to the small number of regular blood donors, the blood donation results do not admit of similar graphical presentation.

to 3.8 vs. 4.4 for ethicists ($t = 4.0, p < .001$) and 4.3 vs. 4.6 for non-philosophers ($t = 1.4, p = .15$); effect size comparison using Fisher's z_r conversion, non-ethicist philosophers vs. ethicists $p = .13$, vs. non-philosophers $.005$. 16% of ethicists who rated blood donation as good reported donating in 2008 or 2009, vs. 0% of ethicists who did not rate blood donation as good; the split was 16% vs. 4% for non-ethicist philosophers and 14% vs. 0% for non-philosophers; overall $\chi^2 = 10.7, p = .001$; by group, ethicists $\chi^2 = 3.6, p = .06$; non-ethicist philosophers $\chi^2 = 2.3, p = .13$; non-philosophers $\chi^2 = 4.8, p = .03$. Among eligible ethicists, 36% of those who rated blood donation as good reported donating at least once every few years, compared to 15% who rated it as neutral or bad ($\chi^2 = 3.3, p = .07$); the split was 36% vs. 21% for non-ethicist philosophers ($\chi^2 = 2.1, p = .15$) and 47% vs. 10% for non-philosophers ($\chi^2 = 12.6, p < .001$). Pairwise Fisher z_r conversion correlation comparison finds a marginally significant trend toward a larger attitude-behavior relationship among the non-philosophers than the non-ethicist philosophers ($r = .36$ vs. $.13, p = .09$); however as usual we treat the generalization question more cautiously than the specific instance question. T-tests comparing mean normative rating depending on self-reported donation at least once every few years (excluding ineligible) show similar results: ethicists mean 7.4 vs. 6.7 ($t = 2.4, p = .02$); non-ethicist philosophers 6.9 vs. 6.6 ($t = 1.1, p = .27$); non-philosophers 7.4 vs. 6.5 ($t = 3.4, p = .001$); effect size comparison using Fisher's z_r conversion, non-philosophers vs. non-ethicist philosophers $r = .30$ vs. $.10, p = .12$.

FIGURE 5: Relationship of expressed normative view about “not having on one’s driver’s license a statement or symbol indicating willingness to be an organ donor in the event of death” and reporting having an organ donor statement or symbol on driver’s license. Decreasing bar height from left to right reflects correlation between self-reported behavior and normative view. Note that non-ethicist philosophers show the strongest trend.



9. Email Responsiveness.

Question 8 asked respondents to rate “not consistently responding to student emails” on the usual 1-9 scale from “very morally bad” to “very morally good”. Question 22 was:

22. About what percentage of student emails do you respond to?

enter a percentage _____

We also measured behavior directly by sending respondents up to three emails designed to look as though they were from students, spread across the course of almost a year. The emails concerned office hours, declaring a major, and a request about a future class. They came from Yahoo and Gmail accounts, were tested against spam filters, and always addressed the professor by name; some of the emails contained specific information pertinent to the professor, such as the name of course the professor was scheduled to teach in the upcoming term (drawn from the schedule of courses at the professor’s university). Any non-automated reply was coded as a “response”. We replaced identifying data with unique tags in a way that prevented us from being able to determine the responses of particular individuals; we were interested only in group trends. A more detailed discussion of our procedure and results is available in Rust and Schwitzgebel (in preparation).

The groups did not differ in expressed normative attitude. 83% of ethicists rated non-responsiveness to student emails on the morally bad side of the scale, compared to 83% of non-

ethicist philosophers and 85% of non-philosophers (1% rated it good), with means of 3.4, 3.4, and 3.1, near the “somewhat morally bad” label of the scale.⁴⁴

Nor did the groups differ detectably in measured behavior. Among survey respondents who received all three emails, ethicists replied to a mean of 1.9 of our messages, non-ethicist philosophers to a mean of 1.8, and non-philosophers to a mean of 1.9.⁴⁵ Survey respondents’ reply rates to each email considered individually were also within statistical chance.⁴⁶

Unsurprisingly, survey non-respondents were less likely to reply to our email messages. However the difference was not large: Among survey recipients who received all three emails, the mean number of replies was 1.6 for the survey non-respondents vs. 1.8 for the survey respondents.⁴⁷ We find this modest difference in the predicted direction reassuring on the issue of survey non-response bias. Still, we expect that normative attitudes toward email

⁴⁴ $\chi^2 = 0.4$ (collapsing neutral and good), $p = .83$. Although the means are statistically different (ANOVA, $F = 5.2$, $p = .006$), because of the scaling issues discussed above we do not interpret differences of less than 0.5 points as meaningful when non-philosophers are in the most extreme group.

⁴⁵ Among respondents, 97 ethicists, 110 non-ethicist philosophers, and 94 non-philosophers received all three emails. ANOVA, $F = 0.9$, $p = .43$.

⁴⁶ First email: 63% for each group ($\chi^2 = 0.0$, $p = .99$). Second email: 61% for ethicists, 55% for non-ethicist philosophers, 56% for non-philosophers ($\chi^2 = 1.0$, $p = .60$). Third email: 69% vs. 62% vs. 66% ($\chi^2 = 2.0$, $p = .38$). Although the trend favored ethicists on emails 2 and 3, this trend did not approach statistical significance even on an expanded sample (see Rust and Schwitzgebel in preparation).

⁴⁷ $t = -2.8$, $p = .006$ (215 non-respondents vs. 301 respondents).

responsiveness and opinions about one's own rate of responsiveness may be a bit higher among survey respondents than in the target population as a whole.

The groups also did not detectably differ in self-reported email responsiveness. About half of the respondents from all three groups claimed to respond to 100% of student emails (49% of ethicists, 48% of non-ethicist philosophers, and 55% of non-philosophers); and 84% of respondents claimed to respond to at least 95% of student emails (82% of ethicists, 83% of non-ethicist philosophers, and 87% of non-philosophers).⁴⁸ We are inclined to regard these self-reported response rates as unrealistically high. There were 122 respondents who claimed 100% responsiveness to student emails, who received all three of our email messages, and who also took the electronic version of our survey (thus confirming the validity of their recorded email address). Among these respondents, only 31% did in fact reply to all three emails (39% replied to two, 25% to one, and 4% to none). One caveat, however, is this: In subsequent discussion, some respondents suggested that they interpreted "student" to mean something like "student whose name I recognize"; they would presumably not have recognized the fictional names we used in our emails.

The relationship between self-reported email responsiveness and responsiveness to our three email messages was very weak: a correlation of only .14.⁴⁹ Self-described email responsiveness thus predicted only 2% of the variance in measured behavior. The correlation was highest for the non-philosophers ($r = .26$, predicting 7% of the variance), intermediate and

⁴⁸ $\chi^2 = 2.0$, $p = .37$; $\chi^2 = 1.7$, $p = .42$.

⁴⁹ Pearson's r , $p = .001$, with self-reported responsiveness transformed by $-(\log(101-x))$.

marginally significant for the ethicists ($r = .13$, predicting 2% of the variance), and essentially non-existent for the non-ethicist philosophers ($r = .06$, predicting 0% of the variance).⁵⁰

We classified the following pattern of responses to our email questions as showing a suspicious tilt toward socially desirable responding: a self-reported rate of 100% responsiveness to student emails and failure to respond to at least one of the email messages we sent or a self-reported rate of at least 95% responsiveness and responsiveness to fewer than 50% of the emails we sent. The groups did not differ by this measure: 38% of ethicist respondents fell into this pattern, compared to 41% of non-ethicist philosophers and 40% of philosophers.⁵¹

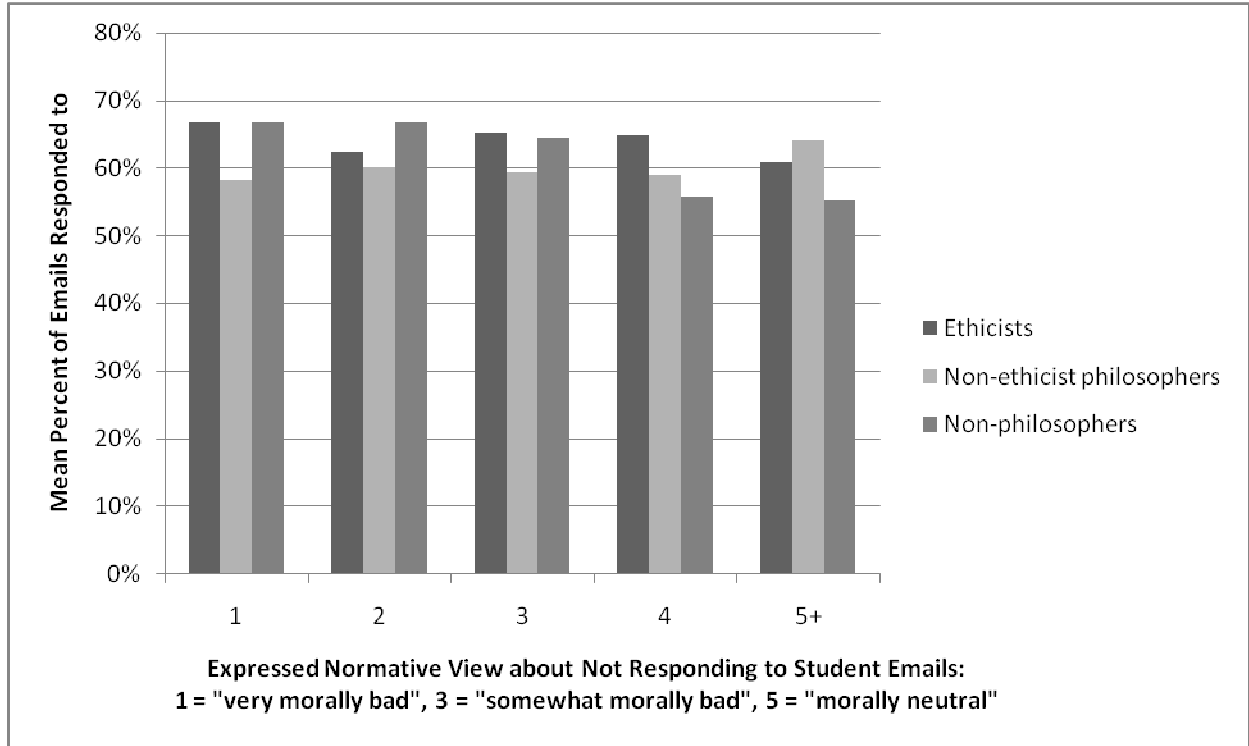
Expressed normative attitude was also largely unresponsive of email responsiveness. The average reply rate to our emails for professors who rated not consistently responding to student emails on the bad side of the scale was 63%; for those who rated it as neutral (or good) the response rate was 61%. For ethicists, the spread was 65% vs. 61%, for non-ethicist philosophers 59% vs. 64% (actually trending in the unpredicted direction), and for non-philosophers 64% vs. 55%. All of these differences are within statistical chance.⁵² Figure 6 displays these data visually.

⁵⁰ Pearson's r , p values of .001, .08, and .37 respectively. Effect size comparison by Fisher's z_r conversion finds non-philosophers vs. non-ethicist philosophers marginally significant at $p = .06$.

⁵¹ $\chi^2 = 0.3$, $p = .88$.

⁵² Merged groups, $t = 0.5$, $p = .60$; ethicists $t = 0.6$, $p = .54$; non-ethicist philosophers $t = -0.8$, $p = .40$; non-philosophers $t = 1.3$, $p = .21$.

Figure 6: Relationship of expressed normative attitude about “not consistently responding to student emails” and mean percentage responsiveness to emails we sent that were designed to appear as though from students. Decreasing bar height from left to right reflects correlation between self-reported behavior and expressed normative attitude.



Expressed normative attitude was moderately correlated with *self-reported* response rate, especially among ethicists: For respondents as a whole the correlation was .26. For ethicists it was .38, for non-ethicist philosophers .24, and for non-philosophers .13.⁵³

In sum, the groups showed no difference in their normative attitude about responsiveness to student emails, in their self-described email responsiveness, or in their measured responsiveness; expressed attitude about the morality of not consistently responding to student emails was not detectably related to measured email responsiveness for any of the groups; and self-reported email responsiveness was only very weakly related to measured responsiveness.

10. Charitable Donation.

The last two questions of Part I concerned charitable donation. Question 9 asked respondents to rate “donating 10% of one’s income to charity” on the usual nine-point moral scale. Question 10 was:

About what percentage of income should the typical professor donate to charity?
(Enter 0 if you think it’s not the case that the typical professor should donate to charity.)

enter a percentage from 0 to 100: _____

In Part II we asked:

⁵³ Pearson’s r , overall, $p < .001$; by group, $p < .001$, $p = .001$, $p = .099$, respectively; effect size comparison by Fisher’s z_r conversion, ethicists vs. non-philosophers $p = .01$, ethicists vs. non-ethicist philosophers $p = .13$.

23. In 2008, about what percentage of your annual income did you donate to charity?

enter a percentage from 0 to 100 (“0” if you never donated): _____

As discussed in Section 2, half of the recipients received a \$10 charity incentive. This allowed us to examine group differences in responsiveness to the charity incentive as a measure of charitable behavior. As we will explain, this manipulation also produced data about response stability and non-response bias.

Ethicists expressed the strongest normative views favoring charity: 89% of ethicists rated donating 10% of one’s income to charity as morally good, compared to 85% of non-ethicist philosophers and 73% of non-philosophers (3% of respondents rated it bad).⁵⁴ The mean response on the 1-9 scale was 7.5 for ethicists, 7.4 for non-ethicist philosophers, and 7.1 for non-philosophers.⁵⁵ In response to the question about how much the typical professor should donate, only 9% of ethicists entered “0”, compared to 24% of non-ethicist philosophers and 25% of non-philosophers.⁵⁶ Among the non-zeros, ethicists’ log-transformed mean answer was 5.9% of income, compared to 4.8% for the other two groups.⁵⁷

⁵⁴ $\chi^2 = 17.0$, $p < .001$ (collapsing neutral and bad). One respondent who rated donating 10% bad said that the typical professor should donate 15% and thus appears to have rated 10% bad because 10% is too little.

⁵⁵ ANOVA, $F = 4.3$, $p = .01$, Tukey pairwise comparisons finding ethicists higher than non-philosophers. Although the difference in means is not large, it is in the direction opposite of that predicted by the apparent scaling differences so we interpret it as meaningful.

⁵⁶ $\chi^2 = 18.2$, $p < .001$ (rounding to zero one ethicist’s response of .001%).

⁵⁷ ANOVA, $F = 3.6$, $p = .03$, Tukey pairwise comparisons only marginally significant.

Non-ethicist philosophers reported having donated the least to charity in 2008. 10% reported having donated nothing, compared to 4% of ethicists and 6% of non-philosophers. Excluding the 0's, non-ethicist philosophers' (log-transformed) mean self-reported donation rate was 2.6%, compared to 3.7% for ethicists and 3.6% for non-philosophers.⁵⁸

However, in our one direct measure of charitable behavior, non-ethicist philosophers behaved the most charitably. Only non-ethicist philosophers showed a significantly higher response rate to the charity version of our study than to the non-charity version: 67% responsiveness with the charity incentive vs. 59% without. For non-philosophers the spread was 55% vs. 52%. Ethicists showed no difference whatsoever, 59% for both versions.⁵⁹

These small differentials in response rate suggest either that the charity incentive was almost entirely unmotivating or, alternatively, that it led some respondents to reconceptualize responsiveness as warranted only if the incentive seemed worth it, weakening whatever prior motivations they may have had. (For what it's worth, \$10 for 10 minutes of work would

⁵⁸ On the zeros: $\chi^2 = 5.9$, $p = .052$ (again rounding to zero one ethicist's response of .001%). On the mean: ANOVA, $F = 5.5$, $p = .004$, Tukey pairwise comparisons finding significant differences between non-ethicist philosophers and both other groups. Variance in charitable donation was similar among the groups: log-transformed SD for all groups 1.0, Bartlett's test, $p = .87$. The zeros and percentages can reasonably be merged using a square-root transform (excluding one outlying non-ethicist philosopher who reported donating 500% of 2008 income): ANOVA, $F = 10.4$, $p < .001$, Tukey pairwise finding significant differences between the non-ethicist philosophers and both other groups.

⁵⁹ Non-ethicists $\chi^2 = 2.8$, $p = .097$, a one-tailed p value of .049; non-philosophers $\chi^2 = 0.2$, $p = .64$; ethicists $\chi^2 = 0.0$, $p = 1.0$.

translate to an annual rate of \$120,000 a year at 40 hours per week for 50 weeks, well over professors' median salary of about \$80,000 in 2009.) These hypotheses can to some extent be distinguished by examining responses to the charity questions in the survey. Suppose that offering the charity incentive led recipients of the charity version of the survey to conceptualize responsiveness as quid-pro-quo and thus led them to respond not from the motives that generate responsiveness to the unincentivized version of the survey but rather only if \$10 to charity for ten minutes' time seemed to them a good deal. If that were the case, then we should expect psychological differences between respondents to the two versions of the survey: Respondents to the charity version should be those among the charity-version recipient pool who value charity relatively highly (finding the quo worth the quid), while respondents to the non-charity version should have a broader and more representative spread of attitudes toward charity (i.e., lower overall). This prediction is not born out. Respondents to the charity version trended only slightly, and not statistically significantly, toward higher mean moral ratings regarding donating 10% of one's income to charity: 7.4, compared to 7.3 for respondents to the version without the charity incentive; and both groups of respondents reported having donated about the same amount to charity in 2008, (log-transformed) mean 3.3% vs. 3.2%, again well within statistical chance.⁶⁰ This lack of difference in responses to the two versions, combined with the tiny difference in overall response rates, thus suggests that recipients were generally little moved by the charity incentive. On the bright side, it also suggests a certain amount of *stability* in response: Respondents' answers to our charity questions don't appear to have been much influenced by the presence or absence of the charity incentive.

⁶⁰ $t = 0.9, p = .36; t = 0.2, p = .83.$

Non-philosophers showed the strongest relationship between expressed normative attitude about how much the typical professor should give and how much they themselves reported giving in 2008, and ethicists showed the weakest attitude-behavior relationship: The correlation was .62 for non-philosophers, .46 for non-ethicist philosophers, and .33 for ethicists.⁶¹ However, there was no correlation whatsoever between expressed normative view on the nine-point scale and self-reported charitable donation in 2008 ($r = -.07$ or $.00$, depending on calculation method).⁶²

In sum, ethicists expressed the strongest normative attitudes in favor of charitable donation. However, ethicists and non-philosophers reported having donated about the same amount of income in 2008. Although non-ethicist philosophers reported lower rates of charitable donation in 2008 than did the other two groups, in our one direct measure of behavior – responsiveness to the charity incentive included with half the surveys – it was only the non-ethicist philosophers who appeared motivated to behave charitably. Ethicists showed the weakest relationship between expressed normative view and self-reported behavior, non-philosophers the strongest.

⁶¹ Pearson's r , all p 's $< .001$ (log-transformed and excluding 0's); effect-size comparison by Fisher's z_r conversion, ethicists vs. non-philosophers $p < .001$, ethicists vs. non-ethicist philosophers $p = .20$, non-ethicist philosophers vs. non-philosophers $p = .06$. The Pearson correlations, square-root transformed and including the 0's (but excluding the outlier) are .24 for ethicists, .41 for non-ethicist philosophers, and .51 for non-philosophers.

⁶² Log-transformed Pearson's correlation (excluding zeros) $-.07$, $p = .13$; non-log-transformed correlation (including zeros but excluding one outlier) exactly $.00$, $p = 1.00$.

11. Survey Response Honesty.

Although the charity questions were the final questions in Part I of the survey, we added one more normative question near the end of the survey. Question 24 was “Using the same scale as in the first part of this questionnaire, please indicate the degree to which reporting dishonestly to survey questions such as the ones presented here is morally good or morally bad”. This question was followed by the usual 1-9 scale. We then asked:

25. Were you dishonest in any of your answers to previous questions?

yes

no

don't recall

This concluded Part II of the survey. We also directly measured survey honesty or accuracy by combining several measures of suspicious responding, described below.

The groups did not appear to differ in normative view, with 89% of ethicists rating survey dishonesty on the bad side of the scale, compared to 87% of non-ethicist philosophers and 86% of non-philosophers.⁶³ Very few respondents self-reported dishonesty: 2% of ethicists, 4% of non-ethicist philosophers, and 2% of non-philosophers (not a statistically detectable difference).⁶⁴ In the comments field, several respondents indicated surprise that we would ask explicitly for reports of survey dishonesty, but since we were also measuring honesty or accuracy directly we thought it best to give respondents a chance to self-report it.

⁶³ $\chi^2 = 0.6$, $p = .74$ (collapsing neutral and good; 3% rated dishonesty as good). Means 3.2, 3.4, and 2.9; ANOVA, $F = 7.0$, $p = .001$, but as usual we don't interpret as meaningful the slightly more extreme mean of the non-philosophers, given scaling concerns.

⁶⁴ Fisher's exact test, $p = .47$.

The most straightforward measure of honesty or accuracy was a comparison of self-reported APA membership with presence on the membership list, discussed in Section 4 above. As mentioned there, we found a weak trend for ethicists to be more likely to overreport membership than non-ethicists. In contrast, non-philosophers trended toward being most likely to have demonstrably overreported voting, though the numbers were small.⁶⁵ For the other questions, suspicious responding had to be statistically inferred (see note 66 for details), but on no measure did we find a significant difference between the groups.

As a final comparison, we aggregated suspicious responding across all measures, awarding respondents one point for each suspicious response.⁶⁶ We did this only for philosophers, due to denominator differences in society membership honesty and reported meat meals per week. Ethicists did not differ from non-ethicist philosophers by this measure: 56% of both groups gave at least one suspicious response – which, we emphasize, is quite consistent

⁶⁵ See note 19.

⁶⁶ One suspicious response point was awarded for overreporting APA membership; for demonstrably overreporting voting, either in 2008 or overall; for reporting maternal contact once a month and not within 30 days, 2-4 times a month and not within 15 days, 5 times a month and not within 10 days; for reporting 3 or fewer meat meals per week and meat the previous night; for reporting at least biennial blood donation and none in 2009, annual donation and none since 2008, or donation every few years and none since 2006; for reporting responding to 100% emails and not replying to all of ours or reporting at least 95% responsiveness and replying to fewer than 50% of ours. More conservative measures of suspicious responding also show no difference between the groups.

with actual scrupulous honesty for any individual survey respondent – and 15% of both groups gave two or more suspicious responses.⁶⁷

Somewhat oddly, perhaps, philosophers who gave at least two suspicious responses rated survey dishonesty morally worse on our 1-9 scale than did philosophers who gave at most one suspicious response: mean 2.8 vs. 3.2 for ethicists, 2.9 vs. 3.5 for non-ethicists.⁶⁸ Non-philosophers showed no such difference, trending instead toward attitude-behavior consistency.⁶⁹

In sum, the rate of survey dishonesty, or bias toward socially desirable responding, appears to be about the same for all three groups, though philosophers showed a negative relationship between expressed attitude and measured behavior.

12. Summary Statistics of Moral Behavior.

Overall, there appears to be no trend toward ethicists behaving better across measures. To confirm this more formally we created three summary statistics of measured and self-reported moral behavior. We are cautious about the value of such summary statistics and certainly do not regard them as valid overall indices of people's moral character. However, aggregating our data across the several measures offers the chance to pick up small but consistent trends that might not show statistically in the measures considered individually.

⁶⁷ $\chi^2 = 0.2$, $p = .89$; $\chi^2 = 0.0$, $p = .95$.

⁶⁸ Ethicists $t = -2.4$, $p = .02$; non-ethicist philosophers also $t = -2.7$, $p = .01$; merged $t = -3.6$, $p < .001$.

⁶⁹ Due to differences in denominator, non-philosophers' suspicious responding scores are not straightforwardly comparable to philosophers'. Only 17 non-philosophers showed at least two suspicious responses: mean 3.2, vs. 2.8 for the remainder ($t = 1.2$, $p = .23$).

Our first summary statistic gave recipients one point each for behavior in accord with the normative view of the majority of philosophy respondents, using one cross-group measure of each type of behavior. Specifically, we awarded one point for self-reported societal membership, one point for self-reported voting frequency of at least once per year, one point for reporting contact with one's mother once per month or more, one point for reporting not having eaten mammal meat at the previous evening meal, one point for being an organ donor, one point for reporting having donated blood since 2005, and one point for responding to at least half of the emails we sent.⁷⁰ Out of seven possible points, then, the mean was 4.4 for ethicists, 4.3 for non-ethicist philosophers, and 4.5 for non-philosophers – well within statistical chance.⁷¹ All the groups also showed the same variance in scores, contrary to what one might predict if one thought (as we have sometimes heard) that even if ethicists behave on average the same as non-

⁷⁰ We choose self-reported membership because measured membership is unavailable for non-philosophers. We choose self-reported vote since many respondents would have to be excluded if we used only measured vote rate, and we detected no pattern of bias in self-reported voting. The donating blood measure disadvantages those ineligible, but that should not affect the group comparison since the rate of self-reported ineligibility is virtually identical among the groups. Patterns of suspicious responding are not included in the measure because of numerator differences between philosophers and non-philosophers. Charitable behavior is excluded because the self-report results went the opposite direction from the measured-behavior results, problematizing the interpretation of both. Respondents missing any component (e.g., reporting “driver’s license not available”) are excluded from the analysis.

⁷¹ ANOVA, $F = 1.6$, $p = .21$.

ethicists, they split bimodally into those who behave substantially better than average and those who behave substantially worse.⁷²

A second summary statistic, limited to philosophers, adds in one more possible point for giving no suspicious survey responses and substitutes measured APA membership for self-reported membership. This statistic finds, again, similar totals for ethicists and other philosophers, ethicists with a mean score of 4.7 out of 8 and other philosophers with a mean of 4.6.⁷³

A third statistic modifies the first by excluding the two most controversial behaviors: societal membership and vegetarianism. The means for this statistic were also within chance, though there was a statistically marginal spread between the groups, with ethicists in the middle: 3.0, 2.8, and 3.1 respectively (out of 5).⁷⁴

Across the measures, however, it does seem that ethicists were the most moralizing – that is, the least likely to describe the nine behaviors as morally neutral. Ethicists rated as “morally neutral” an average of 1.8 of the nine scaled normative questions in Part I, compared to an

⁷² SD = 1.2 for all groups, Bartlett’s test, $p = .61$.

⁷³ $t = 0.5$, $p = .59$, SD = 1.4 for both groups.

⁷⁴ ANOVA, $F = 2.7$, $p = .07$, SD = 0.9, 1.0, 1.0.

average of 2.1 for non-ethicist philosophers and 2.7 for non-philosophers.⁷⁵ The difference is significant even if the vegetarianism question is excluded: 1.4 vs. 1.5 vs. 1.9.⁷⁶

Finally, we aggregated our measures of the attitude-behavior relationship by combining correlation scores across the measures using the tools of statistical meta-analysis (Rosenthal and Rosnow 1984/2008). The differences were well within chance: Using qualitative measures of normative attitude (good vs. not good or bad vs. not bad), the aggregated attitude-behavior correlation was .16 for ethicists and .15 for the other two groups. Using the 1-9 scale of normative attitude, the aggregated attitude-behavior correlation was .20 for ethicists, .24 for non-ethicist philosophers, and .16 for non-philosophers.⁷⁷

⁷⁵ ANOVA, $F = 12.8$, $p < .001$, including only respondents who answered all nine normative questions; pairwise Tukey comparisons finds ethicists and non-ethicist philosophers both different from non-philosophers.

⁷⁶ ANOVA, $F = 4.8$, $p = .009$, Tukey pairwise comparisons finds ethicists different from non-philosophers.

⁷⁷ All r 's were converted to Fisher's z_r , then averaged, then converted back to r . Pairwise comparison of the aggregate r 's, using Fisher's z_r conversion and geometric mean N 's, found all p 's $\geq .48$. For between-group comparability, we excluded the honesty measure and used self-reported society membership. To include the zeros in the charity measure, we used the square-root transformation (excluding the 500% outlier) rather than the logarithmic transformation. For blood donation, we used self-reported donation rates of at least every few years (excluding ineligible) as the behavioral measure. Otherwise the behavioral measures were those with which we concluded the corresponding sections.

13. Age and Gender.

We had gender information for most survey recipients, either discovered in voter registration records or inferred from name and photograph on their academic website; and for about half we had birth year information from voter registration records. We had been concerned about gender and age as potential confounds, especially gender, given that women are underrepresented among philosophers compared to their representation in academia as a whole and given recent work suggesting that they often have different philosophical views (Buckwalter and Stich 2010).

Among survey recipients for whom we had birth year information, there was no statistically detectable age difference among the groups.⁷⁸ Respondents trended toward being younger than non-respondents (mean birth year 1955 vs. 1953), though this difference did not achieve statistical significance.⁷⁹ Age confounds thus seem unlikely to be playing much of a role in survey responses. Only in two respects did older respondents detectably differ in their survey responses or measured behavior: They were less likely to endorse vegetarianism and they voted more frequently.⁸⁰

⁷⁸ Mean birth year 1953 for all three groups of recipients (ANOVA, $F = 0.1$, $p = .92$).

⁷⁹ Overall birth year difference $t = 1.5$, $p = .13$. Mean birth year was 1955 for ethicist and non-ethicist philosopher respondents and 1953 for non-philosopher respondents (ANOVA, $F = 1.4$, $p = .26$).

⁸⁰ Mean birth year of those rating meat eating as bad 1958 vs. 1952 for not bad; $t = 4.1$, $p < .001$. Pearson correlation of age and measured vote rate for all recipients (square-root transformed), $r = .18$, $p < .001$. Pearson correlation of age and self-reported vote rate (square-root transformed), $r = .19$, $p = .001$. As a correction for multiple comparisons we set alpha at .01

Women were more likely to respond to the survey than were men (65% vs. 57%), and they were less likely to be philosophers, especially non-ethicist philosophers (28% of ethicists were women, 17% of non-ethicist philosophers, and 35% of non-philosophers).⁸¹ Thus, gender could potentially have played a confounding role. Fortunately, however, the only issue on which we found statistically significant gender differences was vegetarianism.⁸² Among female respondents, 55% rated eating meat as bad, compared to 37% of men.⁸³ Since women were less likely to be philosophers, this gender difference is even more evident when respondents are

for statistical significance and .05 for marginality. Older respondents were marginally more likely to rate low maternal contact as bad (mean birth year for bad 1954, for not bad 1957; $t = -2.2$, $p = .03$), marginally more likely to rate it bad not consistently to respond to student emails (mean birth year for bad 1954, for not bad 1959; $t = -2.5$, $p = .015$); and marginally less likely to rate it good to give 10% to charity (mean birth year for good 1956, for not good 1951; $t = 2.6$, $p = .011$). They also showed a statistically marginal trend toward reporting having given larger proportions of their income to charity in 2008: Pearson correlation of age and log-transformed charitable rate, $r = .14$, $p = .02$.

⁸¹ $\chi^2 = 5.3$, $p = .02$; $\chi^2 = 29.4$, $p < .001$.

⁸² Again setting alpha at .01 for a significant difference and .05 for a marginal trend as a correction for multiple comparisons. Three marginal trends emerged: 93% of women rated voting as good, compared to 86% of men ($\chi^2 = 5.3$, $p = .02$); 66% rated not keeping in monthly contact with one's mother as bad, compared to 76% of men ($\chi^2 = 6.1$, $p = .013$); and 75% of women reported being organ donors, compared to 66% of men ($\chi^2 = 4.6$, $p = .03$). See the subsequent text for analysis of the gender differences on vegetarianism.

⁸³ $\chi^2 = 28.3$, $p < .001$.

divided by group: The gender split was 78%-50% among ethicists, 70%-39% among non-ethicist philosophers, and 24%-14% among non-philosophers.⁸⁴ Women also reported eating much less meat: In each analysis group, women reported eating about half as many meat meals per week as men: 2.6 vs. 4.9 for ethicists, 2.5 vs. 5.1 for non-ethicist philosophers, and 3.9 vs. 6.4 for non-philosophers.⁸⁵ The splits by last evening meal were not quite as large, however: 27% vs. 44% for ethicists, 32% vs. 33% for non-ethicist philosophers, 39% vs. 51% for non-philosophers.⁸⁶

Thus, the normative question about vegetarianism showed large differences by group, gender, and age. These factors combined to generate very large cumulative differences: Altogether, 81% of female philosophy respondents born 1960 or later rated eating the meat of mammals as morally bad, compared to 7% of male non-philosophers born before 1960.⁸⁷ However, this huge difference in expressed normative attitude does not appear to be accompanied by a correspondingly large shift in actual behavior. Women and philosophers, as described above, were only marginally less likely to report having eaten meat at their previous evening meal; and by age trend was even weaker, with an average birth year of 1954 for those reporting having eaten meat at the previous evening meal and 1955 for those reporting not

⁸⁴ $\chi^2 = 13.0, p < .001; \chi^2 = 11.9, p = .001; \chi^2 = 25, p = .11.$

⁸⁵ All three t 's $\geq 3.6, p$'s $< .001.$

⁸⁶ $\chi^2 = 4.7, p = .03; \chi^2 = 0.0, p = .94; \chi^2 = 1.9, p = .17; \text{overall } 33\% \text{ vs. } 41\%, \chi^2 = 2.8, p = .096.$

⁸⁷ The numbers meeting all three criteria are small, so exact percentages should be treated cautiously, though the difference is extreme enough to be highly statistically significant: 22/27 vs. 3/41, $\chi^2 = 38.5, p < .001.$

having eating meat at the previous evening meal.⁸⁸ Among female philosophers born 1960 or later, 38% reported having eaten meat at the previous evening meal – compared to 37% for all remaining respondents combined.⁸⁹

13. Abstractness of Ethical Interest and Normative Ethical Theory.

In informal conversation, philosophers have often suggested to us that metaethicists – who study the most abstract questions about the nature of moral norms, such as whether there are moral truths at all – behave on average worse than do normative ethicists (who study particular broad normative theories like consequentialism, deontology, or virtue ethics) and applied ethicists (who study issues like abortion, environmentalism, animal rights, or global poverty). Others have suggested that applied ethicists behave morally better than do the other two groups. Also in informal conversation, many philosophers have suggested to us that deontologists (who hold, roughly, that to be moral is to follow a certain set of rules) tend to behave morally worse than consequentialists (who hold, roughly, that morality is about producing good consequences) and/or virtue ethicists (who hold, roughly, that to be moral is to possess a suite of virtues like kindness and courage). Our empirical data support none of these contentions.

Philosophy recipients received two questions, set aside in Part III, that we did not pose to non-philosophy recipients.

⁸⁸ $t = 0.5$, $p = .59$.

⁸⁹ $\chi^2 = 0.0$, $p = .89$, including only respondents for whom gender and birth year information were available. Again, due to the small number of young female philosophers (26), the exact percentage should be treated cautiously: 95% CI for difference -18% to 21%.

26. If an ethics-related area is among your specializations which of the following best reflects the level of abstraction at which you tend to consider ethical issues?

(check all that apply)

- metaethics
- normative ethics
- applied ethics
- no ethics-related area among my specializations

27. What normative ethical view do you find broadly most appealing? (check one)

- deontological
- consequentialist
- virtue ethical
- skeptical
- no settled position

Among ethicist respondents, 35% expressed a focus in metaethics, 64% in normative ethics, and 47% in applied ethics (these categories obviously overlapping). Among philosopher respondents as a whole, 23% reported attraction to deontology, 14% to consequentialism, 29% to virtue ethics, and 4% to skepticism, while 31% reported no settled position.⁹⁰

⁹⁰ These results are rather different from those to a similar question in the PhilPapers survey by David Bourget and David Chalmers in November, 2009 (URL <http://philpapers.org/surveys/>). Differences in subject pool and in response options might explain some of the difference in results.

We detected no statistically significant differences in either expressed moral attitude or moral behavior among ethicists who pursue their trade at different levels of abstraction. However, due to the relatively small sample size, our power to statistically detect differences in moral behavior is somewhat limited.⁹¹

Looking at our data by normative ethical view – including only philosophers who selected either deontology, consequentialism, or virtue ethics from our list – we found virtue ethicists tending to express more neutral moral views on APA membership, organ donation, and blood donation.⁹² Consonantly, they were least likely to report being organ donors and

⁹¹ N = 98 metaethicists, 156 normative ethicists, and 109 applied ethicists. Given the number of comparisons, we set the alpha level for statistical significance at .01. Among the attitudinal questions, the lowest χ^2 p value we found was .04, with 84% of applied ethicists and 93% of other ethicists rating 10% charitable donation as good. Among the behavioral measures, the lowest p value was also .04, with 47% of metaethicists reporting having eaten the meat of a mammal at their previous evening meal, compared to 32% of other ethicists. On the broadest measure of overall behavior – the second measure described in Section 13, including meat eating, measured APA membership, and measured suspicious responding – the metaethicists scored an average of 4.6, the normative ethicists an average of 4.7, and the applied ethicists an average of 4.7, out of 8.

⁹² Rating APA membership as good: virtue ethicists 59% vs. 80% of deontologists and 74% of consequentialists ($\chi^2 = 8.8$, $p = .012$); rating not having an organ donor symbol as bad: virtue ethicists 48% vs. 65% for both other groups ($\chi^2 = 8.8$, $p = .03$); rating blood donation as good 81% vs. 94% and 90% ($\chi^2 = 6.6$, $p = .04$). Because of the consistency of the trend we

marginally least likely to belong to the APA.⁹³ There were no other detectable differences by normative theory. Despite the association of vegetarianism with consequentialism (e.g., Singer 1975/2002), we did not find consequentialists any more likely to rate eating meat as bad or any less likely to report having eating meat at the previous evening meal.⁹⁴

14. Conclusion.

The four simple models offered in the introduction made different predictions about the data, given the background framing assumptions. We will now compare the predictions with the results. Given the complexity of human behavior and the imperfections in our measures, we will not hold it against a model if a few results run contrary to its predictions, as long as the contrary results appear to be outliers from a general trend.

interpret the p values as significant despite the multiple comparisons. All χ^2 's in this note and the next are 3x2.

⁹³ Measured APA membership: virtue ethicists 58%, deontologists 65%, consequentialists 74% ($\chi^2 = 5.1$, $p = .08$); organ donation: 58% vs. 78% and 70% ($\chi^2 = 6.7$, $p = .04$); but blood donation since 2005 did not split in the direction predicted by expressed normative view: 14% vs. 20% and 12% ($\chi^2 = 1.5$, $p = .47$). On the broadest summary measure (measure 2), deontologists scored 4.9, consequentialists 4.6, and virtue ethicists 4.4 – within statistical chance (ANOVA, $F = 2.2$, $p = .11$).

⁹⁴ Rating meat eating as morally bad: consequentialists 58% vs. 53% for both other groups ($\chi^2 = 0.5$, $p = .79$); reporting meat at least meal: consequentialists 38% vs. deontologists 34% and virtue ethicists 38% ($\chi^2 = 0.3$, $p = .87$).

The booster model predicts that ethicists will show the best moral behavior, the truest moral attitudes, and the strongest attitude-behavior relationship, on the assumption that they tend to engage in more philosophical moral reflection than do the comparison groups and are otherwise socially similar in dimensions relevant to the moral behavior under investigation. This model, we think, is very difficult to sustain in light of our data on moral behavior. We now have a number of direct measures of moral behavior. In previous research, we have found ethicists, including political philosophers, to vote at about the same rate as other professors (Schwitzgebel and Rust 2010), to reply emails apparently written by students at about the same rate as other professors (Rust and Schwitzgebel in preparation), to behave about as courteously at philosophy conferences (though environmental ethicists did appear to litter less; Schwitzgebel, Rust, Huang, Moore, and Coates forthcoming), to be rated morally about the same overall by their peers (Schwitzgebel and Rust 2009), and to misappropriate library books no less often (perhaps more often, Schwitzgebel 2009). In the current research, we found ethicists about as likely to pay membership dues to the APA, to be about as accurate or honest in their survey responses, and to be no more responsive than the other groups to the charity incentive (perhaps less responsive). Turning to self-report measures, all groups reported similarly high rates of having an organ donor symbol on their driver's license and similarly low rates of blood donation. Non-philosophers reported keeping best in contact with their mothers; non-philosophers and ethicists reported similar overall rates of charitable donation, somewhat higher than non-philosophers'; and non-philosophers were marginally most likely to report having eating meat at the previous evening meal. We regard these results as an approximate tie among the groups – as is also suggested by our three summary measures.

The remaining three models all predict that ethicists will behave about the same as otherwise socially similar non-ethicists (except for the toxic rationalization sub-model, which predicts worse behavior). In this respect, all three models match the majority of the data described above. However, the three models make different predictions about normative attitudes and attitude-behavior consistency.

The epiphenomenal model predicts no differences in attitude or attitude-behavior consistency. This model stands somewhat in tension with our data, but perhaps not irreconcilably. Ethicists' attitudes were similar to non-ethicists' on the majority of issues; but on several issues – organ and blood donation, charitable donation, and especially vegetarianism – they exhibited more demanding moral views than did the other groups. This is a consistent enough tendency for us to think there is something to it: Ethicists' moral views are not identical to those of non-ethicists, perhaps tending toward the more demanding. Of course this claim, if true, says nothing about causal direction or the role of philosophical moral reflection in generating such views. A revised epiphenomenalism might attribute these differences in moral view to antecedent differences between the types of people who enter ethics vs. other areas of academia, differences arising from causes other than philosophical moral reflection (thus rejecting one of our framing background assumptions). The epiphenomenalist model similarly only partly fits with our data on attitude-behavior consistency. Although on the majority of measures all groups showed similarly low-to-moderate relationships between expressed normative attitude and measured or self-reported behavior (and they also showed very similar attitude-behavior relationships in our two aggregate measures), there were significant between-group differences in attitude-behavior consistency on four of the normative issues under investigation: Ethicists showed the strongest attitude-behavior relationship on the issue of voting,

non-ethicist philosophers showed the strongest relationship on the issues of organ donation and staying in contact with one's mother, and non-ethicists showed the strongest relationship on the issue of charitable donation. We see no unified interpretation of these results; probably multiple factors are at work (including possibly shortcomings in the survey that play out differently for the different groups).

The inert rationalization model predicts no differences in moral behavior (thus, "inert" rationalization, as opposed to "toxic"), but it does predict that ethicists will show a stronger attitude-behavior relationship due to an increased tendency to rationalize their behavior post-hoc by adopting self-flattering normative attitudes. Although the moral behavior data fit fairly well with this model, as noted above, the majority of our data on the attitude-behavior relationship do not. Nonetheless, it's perhaps worth noting that the voting data fit the inert rationalization model beautifully: Ethicists voting participation is virtually identical to other groups', both in mean and in variance, but ethicists show a substantially stronger relationship between behavior and attitude than do the other groups. The simplest interpretation of this pattern of results is inert rationalization: Professors' voting behavior is determined by factors having nothing to do with philosophical reflection and is thus the same among all the groups (given that they are adequately matched in other respects); philosophical moral reflection serves only to shift respondent's attitudes into line with those pre-existing behavioral tendencies, ethicists showing the greatest attitudinal shift (and thus the highest correlation between attitude and behavior) and non-philosophers showing the smallest shift.

The motivationally ineffectual discovery model predicts that ethicists and non-ethicists will behave similarly, that ethicists will have normative views closer to the truth, and that ethicists will probably show lower norm-behavior consistency (to the extent that discovering

moral truths frees them from self-serving rationalizations that artificially inflate attitude-behavior consistency in non-ethicists). The data summarized above mostly confirm the prediction about moral behavior and mostly disconfirm the prediction about attitude-behavior consistency, except on the issue of charitable donation. Whether ethicists also have normative views closer to the truth we can conditionally assess. Ethicists were more likely than the other groups to praise blood donation, to demand relatively high rates of charitable donation, and to condemn meat eating and not having an organ donation symbol. *If* these attitudes are closer to the moral truth, then that speaks in favor of this aspect of the motivationally ineffectual discovery model.

We are inclined toward a hybrid view according to which the psychological story is complicated. No simple view fits our data very well, and given the many limitations of our studies – including the background assumption that ethicists are similar to the comparison groups in all relevant dimensions apart from a greater tendency toward philosophical moral reflection – we think it premature to draw strong conclusions. We do think it is probably warranted at this point, however, pending further evidence, to reject the view that ethicists behave, on average, morally better than do non-ethicists. This result appears to leave us with three broad options, each empirically explorable with further research: (1.) Philosophical reflection simply has no influence on moral behavior. (2.) Philosophical reflection has both positive and negative effects on moral behavior (maybe the booster is sometimes right and the toxic rationalization view sometimes right), approximately cancelling out on average. Or (3.) One of the background assumptions of our research is false: Perhaps (3a.) ethicists would tend to behave morally worse than non-ethicists except that their philosophical moral reflection works to bring them up to average (e.g., maybe they tend to have deficient moral emotional reactions that they compensate for intellectually), or perhaps (3b.) ethicists in fact are no more likely to engage in philosophical

moral reflection on issues related to their outward behavior than are socially similar non-ethicists. A next step would be to test 3a and 3b and to try to pry apart 1 from 2.

A priori and phenomenologically, we do not find Option 1 very attractive. However, empirically, it seems very much in play. It remains to be shown that even a lifetime's worth of philosophical moral reflection has any influence upon one's real-world moral behavior.⁹⁵

⁹⁵ For help with survey design, many thanks to Ruth Chao, David Crow, Martin Johnson, and the U.C. Riverside Survey Research Center. For useful comments and discussion, thanks to audiences at University of Sydney, the American Philosophical Association, CalTech, the Australasian Association of Philosophy, University of Nevada Las Vegas, Oxford University, University of Edinburgh, University of Miami, University of Arizona, University of Osnabrück, and The Splintered Mind blog, and especially to P.J. Ivanhoe, Jean Kazez, Joshua Knobe, Edouard Machery, Michael Ridge, and ****.

References

- Aristotle (4th c. BCE/1962). *Nicomachean Ethics*. Trans. M. Ostwald. New York: Macmillan.
- Baier, Annette (1985). *Postures of the Mind*. Minneapolis: University of Minnesota.
- Bernstein, Robert, Anita Chadha, and Robert Mountjoy (2001). Overreporting voting: Why it happens and why it matters. *Public Opinion Quarterly*, 65, 22-44.
- Buckwalter, Wesley, and Stephen Stich (2010). Gender and philosophical intuition. Unpublished manuscript.
- Hadot, Pierre (1995). *Philosophy as a Way of Life*. Trans. A. Davidson. Malden, MA: Blackwell.
- Holbrook, Allyson L., and Jon A. Krosnick (2010). Social desirability bias in voter turnout reports: Tests using the item count technique. *Public Opinion Quarterly*, 74, 37-67.
- Ivanhoe, Philip J. (1993/2000). *Confucian Moral Self-Cultivation*, 2nd ed. Indianapolis: Hackett.
- Kant, Immanuel (1785/2002). *Groundwork for the Metaphysics of Morals*. Trans. A. Zweig. Ed. T.E. Hill & A. Zweig. Oxford: Oxford.
- Knobe, Joshua, and Brian Leiter (2007). The case for Nietzschean moral psychology. In B. Leiter & N. Sinhababu (eds.), *Nietzsche and Morality*. Oxford: Oxford.
- Mill, John Stuart (1859/2003). *On Liberty*. In M. Warnock (ed.), *Utilitarianism and On Liberty*, 2nd ed. Malden, MA: Blackwell.
- Moeller, Hans-Georg (2009). *The Moral Fool*. New York: Columbia.
- Posner, Richard A. (1999). *The Problematics of Moral and Legal Theory*. Cambridge, MA: Harvard.
- Rosenthal, Robert, and Ralph L. Rosnow (1984/2008). *Essentials of Behavioral Research*, 3rd ed. Boston: McGraw Hill.

- Rust, Joshua, and Eric Schwitzgebel (in preparation). Ethicists' and non-ethicists' responsiveness to student emails: Relationships among expressed normative attitude, self-described behavior, and experimentally observed behavior.
- Schwitzgebel, Eric (2009). Do ethicists steal more books? *Philosophical Psychology*, 22, 711-725.
- Schwitzgebel, Eric, and Joshua Rust (2009). The moral behaviour of ethicists: Peer opinion. *Mind*, 118, 1043-1059.
- Schwitzgebel, Eric, and Joshua Rust (2010). Do ethicists and political philosophers vote more often than other professors? *Review of Philosophy and Psychology*, 1, 189-199.
- Schwitzgebel, Eric, Joshua Rust, Linus T. Huang, Alan Moore, and Justin Coates (forthcoming). Ethicists' courtesy at philosophy conferences. *Philosophical Psychology*.
- Singer, Peter (1972). Famine, affluence, and morality. *Philosophy and Public Affairs*, 1, 229-243.
- Singer, Peter (1975/2002). *Animal Liberation*, 3rd ed. New York: HarperCollins.
- Williams, Bernard (1985). *Ethics and the Limits of Philosophy*. Cambridge, MA: Harvard.