

The moral behavior of ethics professors: Relationships among self-reported behavior, expressed normative attitude, and directly observed behavior

Eric Schwitzgebel and Joshua Rust

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 expressed values? These questions have never been systematically studied. 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 the self-reports. Ethicists expressed somewhat more stringent normative attitudes on some issues, such as vegetarianism and charitable donation. However, on no issue did ethicists show unequivocally 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. We discuss implications for several models of the relationship between philosophical reflection and real-world moral behavior.

Keywords: Applied Ethics; Attitude-Behavior Consistency; Charity; Ethics; Experimental Philosophy; Moral Behavior; Moral Psychology; Vegetarianism; Voting

Eric Schwitzgebel is Professor of Philosophy at the University of California at Riverside.

Joshua Rust is Assistant Professor of Philosophy at Stetson University.

Correspondence to: Eric Schwitzgebel, Department of Philosophy, University of California, Riverside, CA 92521-0201, United States. Email: eschwitz@ucr.edu

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. However, we have found in conversation and informal polling that most philosophers have views on these two questions—views diverse but nonetheless often regarded as obvious; views that either valorize or devalue philosophical moral reflection, or at least certain types of philosophical moral reflection. And philosophers *should* have views: the two issues are central to understanding the potential social and personal value of our discipline.

Historically, many philosophers 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 Hadot, 1987/1995, and Ivanhoe, 2000). Kant sees moral philosophy as arming us with reason against the corrupting force of our wishes and inclinations (1785/2002, p. 206). 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). Singer aims to induce vegetarianism and charitable donation in his readers (e.g., 1972, 2002).

Other philosophers have expressed doubts about the practical effects of philosophical moral reflection, at least as normally practiced. Williams (1985) and Baier (1985) argue that philosophical ethics can undercut our intuitive sense of what’s right. Moeller (2009) argues that philosophical moralizing tends to promote retributive violence. Knobe and Leiter (2007) and Posner (1999) argue that philosophical moral reasoning is largely post hoc rationalization of antecedent opinions, 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 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.

Some prominent twentieth century psychological theories of moral development emphasized the importance of philosophical or philosophical-style moral reflection in shaping moral attitudes and behavior (especially Kohlberg, 1984). Recent psychological research has tended to emphasize, in contrast, the power of irrational, emotional, and non-conscious influences on moral judgment, at the expense of the power of explicit reasoning (Cushman, Young, & Hauser, 2006; Haidt, 2001; Mikhail, 2011). However, the relationship between extended philosophical reflection and real-world behavior has not been systematically explored, nor has the relationship between philosophical moral reflection and attitude-behavior consistency. (Existing research on cognitive elaboration and self-report of reasons suggests that ordinary

reflection in non-philosophical contexts has variable effects on attitude-behavior consistency; Fabrigar, Wegener, & MacDonald, 2010; Wilson, Dunn, Kraft, & Lisle, 1989.)

The possibilities are manifold, the relationships complex, and systematic empirical data almost entirely absent. To help fix ideas, consider as reference points the following four simple hypotheses about the relationship between philosophical moral reflection, moral attitudes, 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 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 would be objectionably sexist). Such discoveries have a significant positive overall 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.

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 his or her behavior rather than his or her behavior shifting to match his or her attitudes. The philosophically 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 attitudes 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.

Inert discovery view: as in the booster view, philosophical moral reflection tends to lead to the discovery of moral truths. However, such discoveries have no material 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, despite the attitude change, one would continue to eat factory-farmed meat at virtually the same rate as one would have done absent any philosophical reflection on the matter.

Epiphenomenalist view: philosophical moral reflection is virtually powerless to change moral behavior or moral attitudes, either for better or for worse, or to improve their correlation. Here's one way it might work: if you can defend your behavior, or some pre-existing attitude, by creating an appealing justification of that behavior or attitude without making any material changes to your attitudes, you do so; otherwise, you simply acknowledge the lack of justification, or call the attitude an "intuition," or deflect with a joke, and go on believing and behaving the same as before. (One might think of such superficial defenses as "rationalizations," but if so, they differ from rationalizations characteristic of what we're calling the "rationalization view," which involve materially shifting one's attitudes to cohere with one's behavior or behavioral impulses.)

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, 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 prompt 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 other professors in all dimensions relevant to moral attitude and 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 inert rationalization, inert discovery, and epiphenomenalist 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 close enough to moral truths; we recognize that the metaethical issues are complex), the booster and inert discovery views predict that ethicists will be better in touch with those truths than are non-ethicists, as a result of their increased 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 attractive misconduct.

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 inert 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 don't do justice to the nuances of Aristotle, Mencius, Kant, Williams, etc. We present them only as starting points for reflection in an area where a dearth of data leaves the empirical options relatively unconstrained.

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.

At a Pacific Division meeting of the American Philosophical Association (APA), 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 & 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 & Rust, 2010). In other studies, we have found ethicists neither more nor less likely than other professors to answer student emails (Rust & Schwitzgebel, forthcoming), to pay their conference registration fees (Schwitzgebel, forthcoming), 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, & Coates, 2012). 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. We chose these five states because we had voting participation data for those states from a previous study (Schwitzgebel & Rust, 2010). We classified philosophy recipients as ethicists or non-ethicists based on research interests stated on recipients' academic homepages (see Schwitzgebel & Rust, 2010, for details of the sampling and classification procedure). After up to five communications, including a hard copy sent to electronic nonrespondents, we received completed surveys from 198 ethicists (59% of ethicist recipients), 208 non-ethicist philosophers (63% of recipients), and 167 non-philosophers (53% of recipients). Although responses were tagged with ID numbers to enable comparisons of behavioral data and survey response data, no identifying information about individuals was stored with the survey responses.

The survey had 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. 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 directing respondents to a comment field. We read all such clarifications and critiques, but did not adjust responses in light of them; the clarifications and critiques brought to light no systematic problems in the data. We chose to solicit responses on a scale of moral goodness rather than, say, a scale of rightness, permissibility, or obligation, because

the badness-goodness scale seemed to us conceptually the simplest and the most symmetrical between the positive and negative sides.

The second section 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. This section 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.¹

Philosophy respondents' questionnaires contained a third section consisting of two questions. The first asked the respondent to report the level of abstraction at which he or she tends to consider ethical issues (metaethics, normative ethics, applied ethics, or no ethics-related area among his or her specializations). The second asked the respondent to report what normative ethical view he or she finds broadly most appealing (deontological, consequentialist, virtue ethical, skeptical, or 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 \$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, as explained in section 10 below.

A full facsimile of the survey is available online as 'Supplementary material'.

3. Theft of \$1000

As mentioned above, the survey began by asking respondents to rate various actions on a nine-point scale from "very morally bad" (1) to "very morally good" (9). The first of these actions 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: 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).² This might reflect a real difference in opinion about the moral badness of theft, but 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, 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) versus rating it as morally neutral (5) versus rating it as morally good (6–9). We will also disregard differences of less than 0.5 in mean rating when the non-philosophers are in the most extreme group.

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. 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.³ Neither did the groups differ much in their mean responses on the scale: 6.2 for ethicists versus 6.3 for non-ethicist philosophers and 6.6 for non-philosophers.⁴ Only three respondents rated membership on the bad side of the scale (two ethicists, one non-ethicist philosopher).

Part 2 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 2 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.⁵

Before distributing the surveys, we had gathered data on philosophy recipients’ membership in the APA 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 review board.) 68% of survey respondents appeared on the membership list (67% of ethicists, 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 due to annual updating), 44% of ethicist non-members falsely reported membership, as did 39% of non-ethicist non-members (not statistically different rates⁶)—seemingly reflecting a bias toward “socially desirable responding,” that is, toward giving answers that would appear to cast the respondent

in a good light (Paulhus, 2002). In section 11, we will discuss the general issue of response accuracy or honesty and whether it varies among the groups.

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 nonrespondents: 68% membership versus 52%.⁷ *Nonresponse bias* is the tendency for survey respondents to differ from survey nonrespondents in a dimension of interest (Groves, Dillman, Eltinge, & Little, 2002). 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% appeared on the APA membership list, compared to 52% of ethicists who did not rate membership on the morally good side of the scale; for non-ethicist philosophers, the corresponding spread was 73% versus 61%.⁸ This difference in relationship between behavior and expressed attitude did not approach statistical significance.⁹ 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.

In sum: regarding societal membership, we found no statistically significant differences between ethicists and non-ethicist philosophers in normative attitude, self-reported behavior, directly measured behavior, accuracy of response (which was

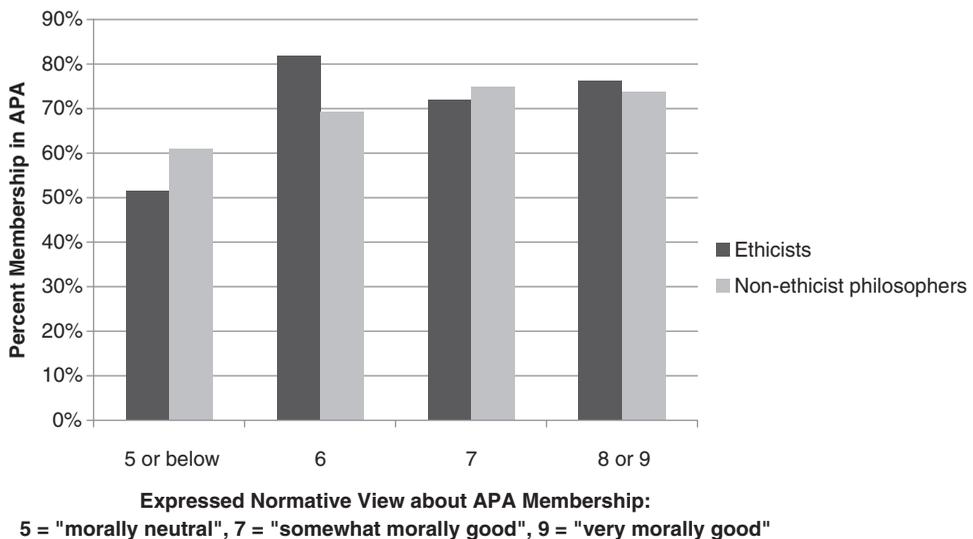


Figure 1. Relationship of expressed normative view about “Regularly paying membership dues to support one’s main academic disciplinary society (the APA, the MLA, etc., as Appropriate)” and appearance on the American Philosophical Association’s Membership List. Increasing bar height from left to right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

low for APA non-members), or attitude-behavior consistency. 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 2, 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.¹⁰ (Only one respondent, a non-ethicist philosopher, rated voting as bad; the remaining 12% rated it 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 versus 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 3.¹¹

Non-Ethicist philosophers were somewhat less likely than either ethicists or non-philosophers to report being a U.S. citizen eligible to vote (84% versus 91% and 93% respectively).¹² 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.¹³ Virtually all respondents who reported being eligible to vote in the U.S. also reported voting in the 2008 U.S. general Presidential election: an incredible 98%. 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 & 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.¹⁴ 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).¹⁵

Survey respondents voted somewhat more frequently than did nonrespondents: 1.17 versus 1.03 votes per year.¹⁶ 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 residence in a different locale. We find the relatively small difference in vote rate between respondents and nonrespondents encouraging on the issue of nonresponse bias.

Despite a tendency for Americans in general to substantially overreport voting (Holbrook & 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 impressive 96% participation rate overall (compared to 62% participation among eligible U.S. citizens generally).¹⁷ Though the numbers are somewhat small, respondents for whom we had full voting records since the year 2000 tended to slightly underreport voting, perhaps due to difficulty remembering local elections and runoffs.¹⁸

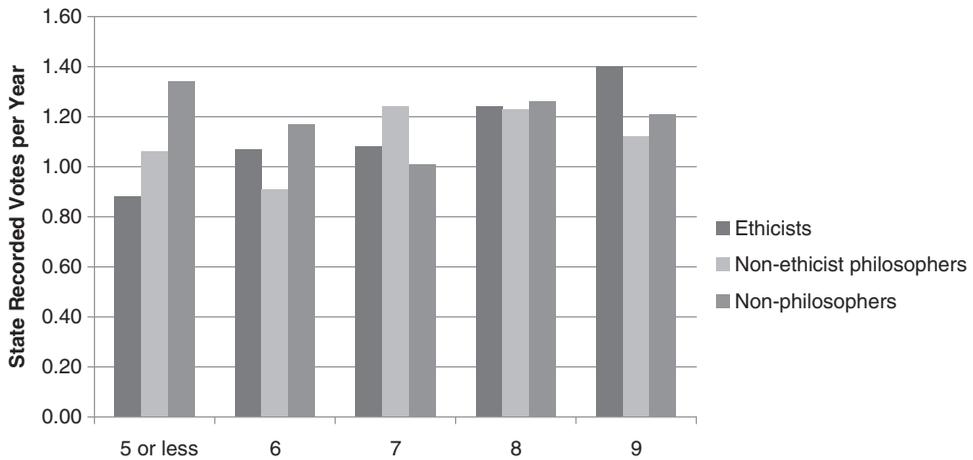
Among the 88% of respondents who expressed the view that voting is morally good, all three groups voted at virtually the same rate according to state records: 1.19 votes per year for ethicists, compared to 1.17 votes per year for both the other groups.¹⁹ The 12% who rated voting as morally neutral trended only slightly toward voting less often: 1.11 votes per year on average (with numbers too small for meaningful statistical comparison among the groups).²⁰ However, ethicists showed a significant relationship between expressed normative attitude on the 1–9 scale and state-recorded voting frequency, with a correlation of 0.36—compared to only 0.14 for the non-ethicist philosophers and 0.01 for the non-philosophers.²¹ 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.

In sum, the groups differed little in overall normative view and had virtually identical voting behavior overall, but only ethicists showed a statistically significant correlation between their measured voting rates and expressed normative attitude about voting.

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.)



Expressed Normative View about Regularly Voting in Public Elections:
 5 = "morally neutral", 7 = "somewhat morally good", 9 = "very morally good"

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.

- 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 then, second, a more specific question that we thought 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.²² (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.²³

Non-Philosophers reported the most frequent contact: 83% of non-philosophers reported contact at least twice a month, compared to 70% of ethicists and 74% of non-ethicist philosophers.²⁴ If we accept the norm of at least monthly contact, non-philosophers trended toward reporting 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.²⁵ On the specific question about days since contact, differences were more subtle, and not statistically significant, but in the same direction.²⁶ 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 geometric mean plus one of 12 days (36% reporting more than 15 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 professor groups similar in this regard.²⁷

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, 94% reported contact within the past 15 days, while among non-ethicist philosophers who rated it as neutral or good, 60% reported contact within the past 15 days, a difference of 34%.²⁸ For ethicists, the corresponding percentages were 87% versus 77% (difference 10%); for non-philosophers, 91% versus 80% (difference 11%).²⁹ Likewise, non-ethicist philosophers showed a much larger correlation between normative response on the 1–9 scale and days since last reported contact: 0.45, compared to 0.13 for ethicists and 0.09 for non-philosophers.³⁰ Figure 3 displays these data graphically.

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 without inviting further confusion. (Indeed, similar considerations apply

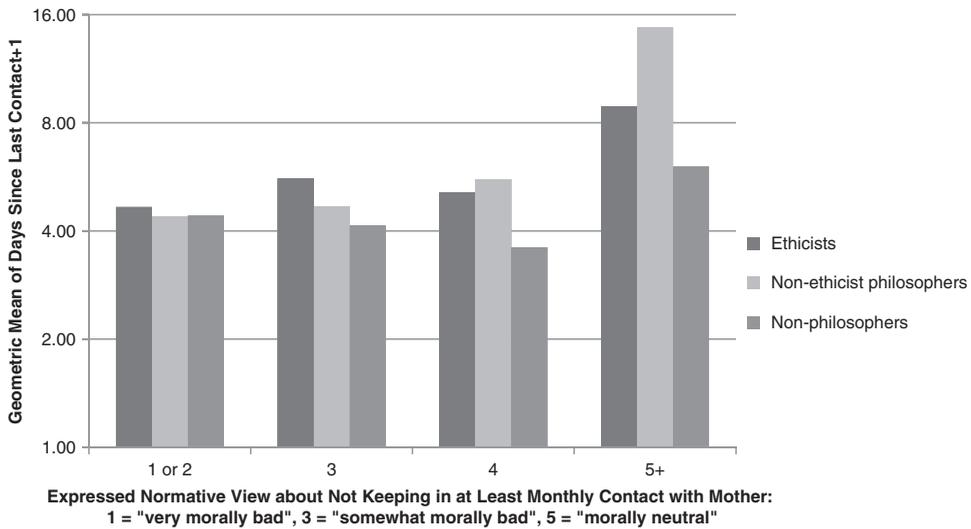


Figure 3. Relationship of expressed normative view about “Not keeping in at least monthly face-to-face or telephone contact with one’s mother” and geometric mean of number of days, plus one, since self-reported last face-to-face or telephone contact. Increasing bar height from left to right reflects 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).

to all our normative questions; we chose to err on the side of simplicity.) Questions 17 and 18 asked respondents to report their own meat consumption:

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.³¹ (Approximately 4% of respondents in all groups rated it as good.) Differences in qualitative response were not as striking, but still meaningful: ethicists’ mean 4.0, non-ethicist philosophers’ 4.2, and non-philosophers’ 4.7.³²

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.³³ 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.³⁴ Strict vegetarianism seems morally and psychologically rather different from merely low meat consumption, so we also divided the 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 14% of non-philosophers.³⁵

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, 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 (versus a target of 14%–29%), as did 51% of respondents who reported three meals/week (versus a target of 21%–43%). This apparent bias toward socially desirable responding did not differ between the groups: 33% of ethicists in the 1–3 meals/week group reported eating meat the previous night, compared to 35% of non-ethicist philosophers and 35% of non-philosophers.³⁶ Two of the 110 respondents who reported zero meals/week also reported having eaten meat at the previous evening meal (one ethicist, one non-ethicist philosopher).³⁷

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%–27% for ethicists, 37%–28% for non-ethicist philosophers, and 51%–23% for non-philosophers, and strength of expressed normative attitude predicted a low likelihood of reporting meat-eating at the last evening meal, as shown in Figure 4.³⁸

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. 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

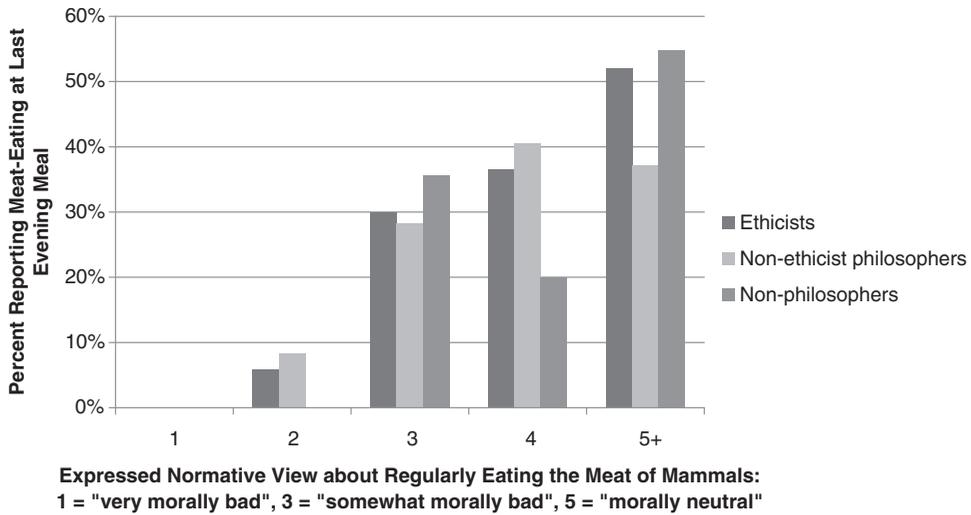


Figure 4. Relationship of expressed normative attitude about “Regularly eating the meat of mammals such as beef or pork” and reporting having eaten the meat of a mammal at the last evening meal. Increasing bar height from left to right reflects correlation between self-reported behavior and normative view. Notice 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.

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.

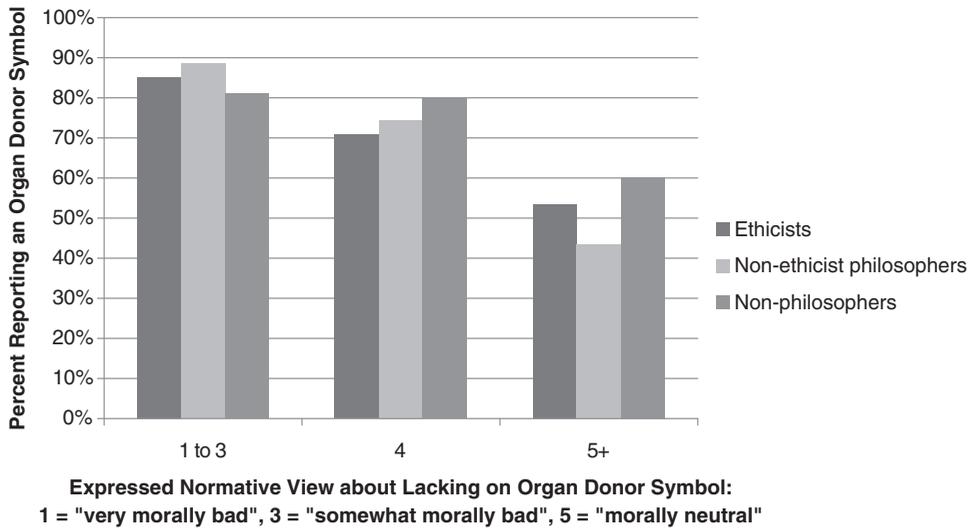


Figure 5. Relationship of expressed normative view about “Not having on one’s driver’s license a statement or symbol indicating willingness to be an organ donor in the event of death” and reporting having an organ donor statement or symbol on driver’s license. Decreasing bar height from left to right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

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 Eligible and ineligible donors did not detectably differ in normative view.⁴⁰

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.⁴¹ 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 2009 (excluding ineligible) was 13% of ethicists, 14% of non-ethicist philosophers, and 10% of non-philosophers.⁴²

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.⁴³ 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.⁴⁴ Figure 5 displays the organ donation results. Due to the small number of

regular blood donors, the blood donation results do not admit of similar graphical presentation.

9. Email Responsiveness

Question 8 asked respondents to rate “not consistently responding to student emails” on the usual 1–9 scale. 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 a 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. A more detailed discussion of our procedure and results is available in (Rust & Schwitzgebel, forthcoming).

The groups did not differ in expressed normative attitude. 83% of ethicists rated nonresponsiveness 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.⁴⁵

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 1.9.⁴⁶ Survey respondents’ reply rates to each email considered individually were also within statistical chance.

Unsurprisingly, survey nonrespondents 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 1.8 for the survey respondents versus 1.6 for the survey nonrespondents.⁴⁷ We find this modest difference in the predicted direction reassuring on the issue of survey nonresponse bias.

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 weak: a correlation of only 0.14.⁴⁹ Self-Described email responsiveness thus predicted only 2% of the variance in measured behavior. The correlation was highest for the non-philosophers ($r=0.26$, predicting 7% of the variance), intermediate and marginally significant for the ethicists ($r=0.13$, predicting 2% of the variance), and essentially non-existent for the non-ethicist philosophers ($r=0.06$, 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 non-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%–61%, for non-ethicist philosophers 59%–64% (thus trending in the unpredicted direction), and for non-philosophers 64%–55%. All these differences are within statistical chance.⁵² See Figure 6.

Expressed normative attitude was moderately correlated with *self-reported* response rate, especially among ethicists: for respondents as a whole the correlation was 0.26. For ethicists it was 0.38, for non-ethicist philosophers 0.24, and for non-philosophers 0.13.⁵³

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 (though it was moderately related to self-described responsiveness, especially for ethicists). Self-reported email responsiveness was only very weakly related to measured responsiveness.

10. Charitable Donation

The last two questions of part 1 concerned charitable donation. Question 9 asked respondents to rate “donating 10% of one’s income to charity” on the usual

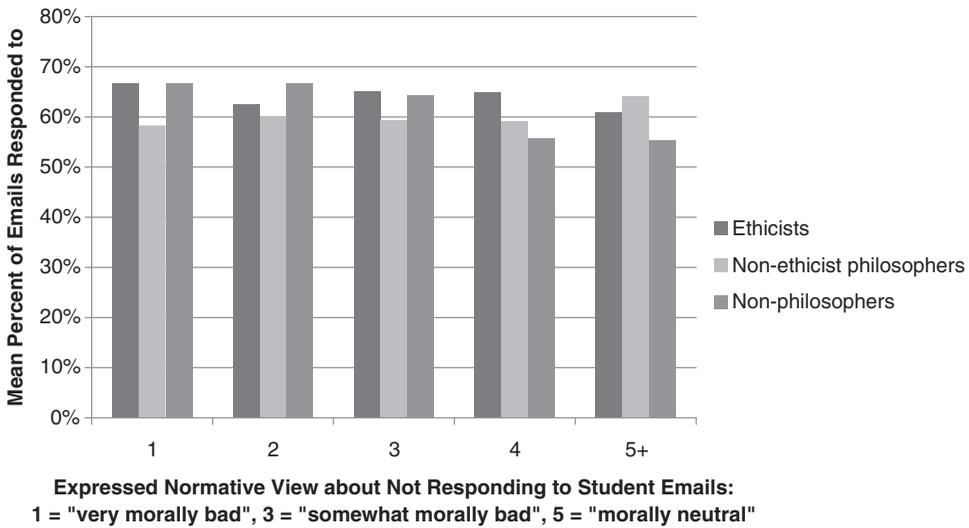


Figure 6. Relationship of expressed normative attitude about “Not consistently responding to student emails” and mean percentage responsiveness to emails we sent that were designed to appear as though from students. Decreasing bar height from left to right reflects correlation between observed behavior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

nine-point moral scale. Question 10 was:

10. 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 2, we asked:

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 nonresponse 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’ geometric mean answer was 5.9% of income, compared to 4.8% for the other two groups.⁵⁷

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 zeros, non-ethicist philosophers' geometric 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 versus 59% without. For non-philosophers, the spread was 55% versus 52%. Ethicists showed no difference whatsoever, 59% for both versions.⁵⁹

These small differences 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 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, geometric mean 3.3% versus 3.2%, 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.

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 relationship: the correlation was 0.62 for non-philosophers, 0.46 for non-ethicist philosophers, and 0.33 for ethicists.⁶¹ However, there was no correlation whatsoever between expressed

normative view on the nine-point scale and self-reported charitable donation in 2008.⁶² We examined relationships among charity choice (Oxfam America, World Wildlife Fund, CARE, Make-a-Wish, Doctors Without Borders, American Red Cross) and other measures, but found no striking relationships, perhaps due to limited sample size.⁶³

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.

11. Survey Response Honesty

Although the charity questions were the final questions in part 1 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 responding 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 2 of the survey.

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.⁶⁴ Very few respondents self-reported dishonesty: 2% of ethicists, 4% of non-ethicist philosophers, and 2% of non-philosophers (not a statistically detectable difference).⁶⁵ 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.

To test overall survey honesty, 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 entirely consistent 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 versus 3.2 for ethicists, 2.9 versus 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 ethicists compared to other philosophers, and both groups 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 measures offers the chance to detect small but consistent trends that might not show statistically in the measures considered individually.

Our first summary statistic gave recipients one point each for behavior in accord with the expressed normative view of the majority of philosophy respondents, using one measure of each type of behavior. Specifically, we awarded one point each for self-reported societal membership, self-reported voting frequency of at least once per year, self-reported contact with one's mother once per month or more, self-reported non-consumption of mammal meat at the previous evening meal, self-reported status as an organ donor, self-reported blood donation since 2005, and experimentally observed responsiveness 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-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 intermediate: 3.0 for ethicists, 2.8 for non-ethicist philosophers, and 3.1 for non-philosophers (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 1, compared to an 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 versus 1.5 versus 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 & Rosnow, 2008). The differences were well within chance: using qualitative measures of normative attitude (good versus not good, or bad versus not bad), the aggregated attitude-behavior correlation was 0.16 for ethicists and 0.15 for the other two groups. Using the 1–9 scale of normative attitude, the aggregated attitude-behavior correlation was 0.20 for ethicists, 0.24 for non-ethicist philosophers, and 0.16 for non-philosophers.⁷⁷

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 & Stich, unpublished manuscript).

Among survey recipients for whom we had birth year information, there was no statistically detectable age difference among the groups or between respondents and nonrespondents.⁷⁸ 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.⁷⁹

Women were more likely to respond to the survey than were men (65% versus 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).⁸⁰ Thus, gender could potentially have played a confounding role. Fortunately, the only issue on which we found statistically significant gender differences was vegetarianism.⁸¹ Among female respondents, 55% rated eating meat as bad, compared to 37% of men.⁸² This gender difference is even more evident when respondents are divided by group: the gender split for rating eating meat as bad was 78%–50% among ethicists, 70%–39% among non-ethicist philosophers, and 24%–14% among non-philosophers.⁸³ The differences in self-reported meat eating at the previous evening meal were not correspondingly large: 27% versus 44% for ethicists,

32% versus 33% for non-ethicist philosophers, and 39% versus 51% for non-philosophers.⁸⁴

Thus, the normative question about vegetarianism showed large differences by recipient 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 1960.⁸⁵ However, this huge difference in expressed normative attitude does not appear to be accompanied by a correspondingly large shift in behavior. Women, as described above, and philosophers, as described in section 7, were only marginally less likely to report having eaten meat at their previous evening meal. By age, the trend was also very weak, with an average birth year of 1954 for those reporting having eaten meat at the previous evening meal and 1955 for those reporting not having eaten meat.⁸⁶ 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.⁸⁷

14. 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).⁸⁸ Our empirical data support none of these contentions.

Philosophy recipients received two questions, set aside in part 3, that we did not pose to non-philosophy recipients.

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. Among philosopher respondents as a whole, 23% reported attraction to deontology, 14% to consequentialism, 28% to virtue ethics, and 4% to skepticism, while 31% reported no settled position.⁸⁹

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.⁹⁰

Looking at our data by normative ethical view—excluding skeptics and those stating no settled position—we found philosophers favoring virtue ethics tending to express more neutral moral views on APA membership, organ donation, and blood donation.⁹¹ Consonantly, they were least likely to report being organ donors and marginally least likely to belong to the APA.⁹² There were no other detectable differences by normative theory.

15. 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.

The booster model predicts that ethicists will show better moral behavior, truer moral attitudes, and more attitude-behavior consistency than do the comparison groups, 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 issues under investigation. This set of predictions appears to conflict with our data. In previous research, we have found ethicists, including political philosophers, to vote at about the same rate as other professors (Schwitzgebel & Rust, 2010), to reply to emails apparently written by students at about the same rate as other professors (Rust & Schwitzgebel, forthcoming), to behave about as courteously at philosophy conferences (though environmental ethicists did appear to litter less; Schwitzgebel et al., 2012), to skip paying their conference registration fees about as often (Schwitzgebel, forthcoming), to be rated morally about the same overall by their peers (Schwitzgebel & 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 annual 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 licenses 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-ethicist philosophers'; and non-philosophers were marginally most likely to report having eaten 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. If the booster is right that philosophical moral reflection tends to improve moral behavior, this isn't evident from overall trends in the moral behavior of professional ethicists.

The remaining three models all predict that ethicists will behave about the same as otherwise socially similar non-ethicists (except for the toxic rationalization submodel, which predicts worse behavior). In this respect, all three models match the majority of the behavioral 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. 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 tendency was consistent enough for us to think that there is something behind it. Of course, even if ethicists do tend to embrace more demanding moral views, it's an open question what the role of philosophical moral reflection is in generating such views. Similarly, the epiphenomenal model 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.

The inert rationalization model predicts no differences in practical 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, the voting data fit the inert rationalization model perfectly: 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. Schwitzgebel and Cushman's (2012) finding that philosophers are

highly prone to order effects in their responses to hypothetical moral scenarios also provides evidence that philosophical reflection may sometimes enhance post hoc rationalization more than it enhances sound reasoning.

The inert 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 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 only 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 inert 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 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—it is premature to draw strong conclusions. It is probably warranted at this point, however, 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 material 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.

Acknowledgements

The authors wish to thank our survey respondents for their time; for advice on survey design, Ruth Chao, David Crow, and Martin Johnson; for help in data collection,

David Crow and the U.C. Riverside Survey Research Center; for valuable discussion and feedback, readers of the Splintered Mind blog and audiences at the Society for Philosophy and Psychology, the Loyola Marymount Department of Economics, Cal State Long Beach, University of Sydney, Edouard Machery's CalTech seminar, University of Nevada Las Vegas, Oxford University, University of Edinburgh (especially Michael Ridge), University of Miami, the University of Arizona undergraduate philosophy conference, University of Osnabrück, Australian National University, U.C. Riverside Cognitive Lunch, and U.C. Riverside Comparative Literature; and for partial funding, the U.C. Riverside Academic Senate.

Notes

- [1] Overall, only 25 respondents (4%) who began the survey failed to complete it. Most questions had response rates from 98%–100% excluding dropouts.
- [2] Philosophers 70% versus non-philosophers 92%, $\chi^2 = 31.2$, $p < 0.001$. Ethicists 74% versus non-ethicist philosophers 65%, $\chi^2 = 3.8$, $p = 0.052$.
- [3] $\chi^2 = 0.9$, $p = 0.63$ (collapsing neutral and bad).
- [4] ANOVA, $F = 5.2$, $p = 0.006$, but see the scaling discussion in section 3.
- [5] $\chi^2 = 5.6$, $p = 0.06$ (excluding 3 respondents answering “don't recall”); merging the two philosophy groups, $\chi^2 = 5.6$, $p = 0.02$.
- [6] $\chi^2 = 0.3$, $p = 0.56$.
- [7] $\chi^2 = 18.7$, $p < 0.001$.
- [8] Ethicists: $\chi^2 = 11.1$, $p = 0.001$; non-ethicist philosophers: $\chi^2 = 3.0$, $p = 0.08$.
- [9] We compare strength of 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 & Rosnow, 2008): ethicists $r = 0.24$ versus non-ethicist philosophers $r = 0.12$, $Z = 1.23$, $p = 0.22$. For most of the data presented in this article, multiple regression analysis for group-by-attitude interaction effects yields similar p values. However, we prefer the simplicity and interpretability of the Rosenthal-Rosnow approach.
- [10] $\chi^2 = 2.4$, $p = 0.30$ (collapsing neutral and bad).
- [11] ANOVA, $F = 16.0$, $p < 0.001$. Tukey pairwise comparisons find non-philosophers different from both ethicists and non-ethicist philosophers.
- [12] $\chi^2 = 8.2$, $p = 0.02$ (see also Schwitzgebel & Rust, 2010).
- [13] ANOVA, $F = 0.7$, $p = 0.51$. All vote rate data are square-root transformed before analysis for a better approximation to normality.
- [14] ANOVA, $F = 0.3$, $p = 0.76$, excluding recipients whose first state-recorded vote was in 2008 and recipients without a good name match in the voter rolls (see Schwitzgebel & Rust, 2010, for methodological details).
- [15] $SD = 0.22$ for ethicists, 0.22 for non-ethicist philosophers, and 0.21 for non-philosophers (Levene's test, $p = 0.83$).
- [16] $t = 4.0$, $p < 0.001$.
- [17] U.S. participation rates from http://elections.gmu.edu/Turnout_2008G.html. Only three ethicists, five non-ethicist philosophers, and two non-philosophers appeared to falsely report voting in the 2008 U.S. general election.
- [18] Only 88 respondents had full voting records since the year 2000, partly because the California database did not include local elections. These 88 respondents voted an average of 10.6 times according to state records but self-reported having voted only 9.2 times (paired-t, $t = -3.1$, $p = 0.003$). Among these 88, 8/30 ethicists overreported voting versus 7/24 non-ethicist philosophers and 15/34 non-philosophers ($\chi^2 = 2.5$, $p = 0.28$).

- [19] ANOVA, $F=0.1$, $p=0.93$.
- [20] $t=1.0$, $p=0.34$.
- [21] Pearson's r ; $p < 0.001$, 0.18, and 0.91 respectively. Correlations compared using Fisher's z_r conversion: ethicists versus non-ethicist philosophers, $Z=1.59$, $p=0.11$; ethicists versus non-philosophers, $Z=2.42$, $p=0.02$.
- [22] $\chi^2=0.5$, $p=0.79$ (collapsing neutral and good).
- [23] ANOVA, $F=4.3$, $p=0.01$, but see the scaling discussion in section 3.
- [24] $\chi^2=8.0$, $p=0.02$.
- [25] $\chi^2=5.7$, $p=0.06$.
- [26] Excluding respondents reporting deceased mothers, 8% of non-philosophers versus 10% of ethicists and 11% of non-ethicist philosophers reported more than 30 days since last contact ($\chi^2=0.4$, $p=0.82$). Geometric mean of reported days plus one since last contact: 5.0 for non-philosophers versus 5.9 for ethicists and 6.7 for non-ethicist philosophers (ANOVA, $F=1.2$, $p=0.31$). (Geometric mean used because of skew.)
- [27] ANOVAs, $F \leq 1.5$, $p \geq 0.24$.
- [28] $\chi^2=22.0$, $p < 0.001$, excluding respondents reporting deceased mothers.
- [29] Respectively, $\chi^2=1.7$, $p=0.20$; $\chi^2=2.6$, $p=0.104$.
- [30] Pearson's r , $\ln(x+1)$ -transformed, $p < 0.001$, 0.17, and 0.37 respectively; merging all groups, $r=0.26$, $p < 0.001$. Correlations compared using Fisher's z_r conversion: non-ethicist philosophers versus ethicists, $Z=2.76$, $p=0.006$; non-ethicist philosophers versus non-philosophers, $Z=2.94$, $p=0.003$.
- [31] $\chi^2=64.2$, $p < 0.001$ (collapsing neutral and good).
- [32] ANOVA, $F=17.0$, $p < 0.001$, Tukey pairwise comparisons finding both philosopher groups different from the non-philosophers.
- [33] $\chi^2=5.7$, $p=0.06$, excluding two respondents who answered "don't recall."
- [34] ANOVA (square-root transformed), $F=5.2$, $p=0.006$, reducing one response from "500" to 25.
- [35] $\chi^2=8.6$, $p=0.01$. Only one respondent, an ethicist, reported between zero and one meals per week (0.5, classified as non-vegetarian).
- [36] $\chi^2=0.0$, $p=0.98$.
- [37] Compare Haddad and Tanzman's (2003) finding that about two-thirds of self-described vegetarians, in a representative U.S. sample, report having eaten at least ten grams of meat, poultry, or fish when queried in specific about two full days' worth of food consumption.
- [38] Bad versus not bad: overall $\chi^2=21.5$, $p < 0.001$; ethicists $\chi^2=12.0$, $p=0.001$; non-ethicist philosophers $\chi^2=1.6$, $p=0.20$; non-philosophers $\chi^2=8.1$, $p=0.005$; effect size comparison using Fisher's z_r conversion: r values 0.26, 0.09, 0.22; pairwise, the ethicists versus non-ethicist philosophers difference was marginally significant, $Z=1.69$, $p=0.09$. 1–9 scale: overall Pearson's $r=0.24$, $p < 0.001$; ethicists $r=0.28$, non-ethicist philosophers $r=0.17$, non-philosophers $r=0.24$ (group p 's ≤ 0.02 , effect sizes not significantly different).
- [39] $\chi^2=8.7$, $p=0.01$ (collapsing neutral and bad); $\chi^2=13.2$, $p=0.001$ (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 versus non-ethicist philosophers 4.1 versus non-philosophers 4.4, $F=5.5$, $p=0.004$) but not the blood question (6.8 versus 6.7 versus 6.9, $F=0.4$, $p=0.65$).
- [40] 78% of eligible rated blood donation as good versus 82% of ineligible ($\chi^2=1.2$, $p=0.27$). Eligible mean 6.8 versus ineligible mean 6.9 ($t=0.9$, $p=0.40$).
- [41] $\chi^2=0.6$, $p=0.75$, excluding six respondents who reported an unavailable driver's license.
- [42] $\chi^2=0.8$, $p=0.67$.
- [43] Using March 10, 2009 as average survey completion date and target average last donation dates of 100 days ago for self-reported donation twice a year or more, 200 days ago for self-reported donation about once a year, and 500 days ago for self-reported donation once every

few years, 53/86 respondents reported last donation older than the target date (62% versus 50%, $Z = 2.2$, $p = 0.03$).

- [44] 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 ($\chi^2 = 31.7$, $p < 0.001$). The split was 79%–53% for ethicists and 81%–60% for non-philosophers ($\chi^2 = 13.2$, $p < 0.001$; $\chi^2 = 7.5$, $p = 0.006$). Effect size comparison using Fisher's z_r conversion: r values 0.27, 0.41, 0.22; non-ethicist philosophers versus non-philosophers, $Z = 1.97$, $p = 0.048$. T -tests comparing mean normative view between donors and non-donors yield similar results. 16% of ethicists who rated blood donation as good reported donating in 2008 or 2009, versus 0% of ethicists who did not rate blood donation as good; the split was 16%–4% for non-ethicist philosophers and 14%–0% for non-philosophers ($\chi^2 = 3.6$, $p = 0.06$; $\chi^2 = 2.3$, $p = 0.13$; $\chi^2 = 4.8$, $p = 0.03$).
- [45] $\chi^2 = 0.4$ (collapsing neutral and good), $p = 0.83$. ANOVA, $F = 5.2$, $p = 0.006$, but see section 3 on scaling.
- [46] Among respondents, 97 ethicists, 110 non-ethicist philosophers, and 94 non-philosophers received all three emails. ANOVA, $F = 0.8$, $p = 0.43$.
- [47] $t = 2.8$, $p = 0.005$.
- [48] $\chi^2 = 2.0$, $p = 0.37$; $\chi^2 = 1.7$, $p = 0.42$.
- [49] Pearson's r , $p = 0.001$, with self-reported responsiveness transformed by $-(\log(101 - x))$.
- [50] Pearson's r , p values of 0.001, 0.08, and 0.37 respectively. Effect size comparison by Fisher's z_r conversion finds non-philosophers versus non-ethicist philosophers marginally significant at $Z = 1.90$, $p = 0.06$.
- [51] $\chi^2 = 0.3$, $p = 0.88$.
- [52] Merged groups, $t = 0.5$, $p = 0.61$; ethicists $t = 0.6$, $p = 0.52$; non-ethicist philosophers $t = -0.8$, $p = 0.44$; non-philosophers $t = 1.2$, $p = 0.22$.
- [53] Pearson's r , overall, $p < 0.001$; by group, $p < 0.001$, $p = 0.099$, respectively (self-reported rate transformed by $-(\log(101 - x))$); effect size comparison by Fisher's z_r conversion, ethicists versus non-philosophers $Z = 2.46$, $p = 0.01$, ethicists versus non-ethicist philosophers, $Z = 1.50$, $p = 0.13$.
- [54] $\chi^2 = 17.0$, $p < 0.001$ (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 too little.
- [55] ANOVA, $F = 4.3$, $p = 0.01$, Tukey pairwise comparisons finding ethicists higher than non-philosophers.
- [56] $\chi^2 = 18.2$, $p < 0.001$ (rounding to zero one ethicist's response of 0.001%).
- [57] ANOVA, $F = 3.6$, $p = 0.03$, Tukey pairwise comparisons marginally significant.
- [58] On the zeros: $\chi^2 = 5.9$, $p = 0.052$. On the mean: ANOVA, $F = 5.5$, $p = 0.004$, Tukey pairwise comparisons finding significant differences between non-ethicist philosophers and both other groups. Variance in charitable donation was similar among the groups: ln-transformed SD the same for all groups at 1.0, Levene's test, $p = 0.75$. Philosophers' salaries are perhaps somewhat below average among the traditional letters and sciences from which most of our comparison respondents were drawn (see Oklahoma State University, 2010, including 115 research institutions).
- [59] Non-ethicists $Z = 1.7$ (one-tailed), $p = 0.048$; non-philosophers $Z = 0.5$, $p = 0.32$; ethicists $Z = 0.0$, $p = 0.50$.
- [60] $t = 0.9$, $p = 0.36$; $t = 0.2$, $p = 0.83$.
- [61] Pearson's r , all p 's < 0.001 (ln-transformed and excluding zeros); effect-size comparison by Fisher's z_r conversion, ethicists versus non-philosophers, $Z = 2.94$, $p = 0.003$, non-ethicist philosophers versus non-philosophers, $Z = 1.69$, $p = 0.09$. Only 76% of those answering 0% to the normative question responded to the 2008 donation question, and the group numbers were small and unbalanced, problematizing analysis. For what it's worth, the geometric mean self-reported 2008 donation among normative zeros was: 4.5% (12 ethicists) versus

- 2.0% (33 non-ethicist philosophers) and 2.2% (28 non-philosophers); ANOVA, $F = 1.9$, $p = 0.15$.
- [62] In-Transformed Pearson's correlation (excluding zeros) -0.07 , $p = 0.13$.
- [63] For example, ethicists did not detectably differ from the comparison groups in their choices among the six charities, at an alpha level of 0.01 for multiple post hoc comparisons. WWF donors rated regularly eating the meat of mammals worse than did respondents selecting other charities (3.7 versus 4.4 on the attitude scale; $N = 38$ versus 262, $t = 2.3$, $p = 0.03$; equal variances not assumed), but WWF selection did not detectably relate to self-reported behavior; nor did Red Cross selection ($N = 24$) detectably relate to the blood donation questions.
- [64] $\chi^2 = 0.6$, $p = 0.74$ (collapsing neutral and good; 3% rated dishonesty as good). Means 3.2, 3.4, and 2.9; ANOVA, $F = 7.0$, $p = 0.001$, but see section 3 on scaling.
- [65] Fisher's exact test, $p = 0.47$.
- [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 semiannual blood donation and none in 2009, annual donation and none since 2008, or donation every few years and none since 2006; and 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.
- [67] $\chi^2 = 0.2$, $p = 0.89$; $\chi^2 = 0.0$, $p = 0.95$.
- [68] Ethicists $t = -2.4$, $p = 0.02$; non-ethicist philosophers $t = -2.7$, $p = 0.01$; merged $t = -3.6$, $p < 0.001$.
- [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 normative response 3.2, versus 2.8 for the remainder ($t = 1.2$, $p = 0.23$).
- [70] Patterns of suspicious responding are excluded due to denominator 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, $F = 1.6$, $p = 0.21$.
- [72] SD = 1.1 for ethicists and non-philosophers, 1.2 for non-ethicist philosophers; Levene's test, $p = 0.75$.
- [73] $t = 0.5$, $p = 0.59$, SD = 1.4 for both groups.
- [74] ANOVA, $F = 2.7$, $p = 0.07$, SD = 0.9, 1.0, 1.0.
- [75] ANOVA, $F = 12.8$, $p < 0.001$, 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, $F = 4.8$, $p = 0.009$, 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 ≥ 0.48 . 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 a square-root transformation (excluding a 500% outlier). 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.

- [78] Respondents' mean birth year 1955 versus nonrespondents' 1953 ($t = 1.5, p = 0.12$). Mean birth year 1953 for all three groups of recipients (ANOVA, $F = 0.1, p = 0.92$). Among respondents, ethicists' mean birth year was 1955 versus 1956 for non-ethicist philosophers, and 1953 for non-philosophers (ANOVA, $F = 1.3, p = 0.26$).
- [79] Mean birth year of those rating meat eating as bad 1958 versus 1952 for not bad; $t = 4.1, p < 0.001$. Pearson correlation of age and measured vote rate for all recipients (square-root transformed), $r = 0.18, p < 0.001$. As a correction for multiple comparisons we set alpha at 0.01 for statistical significance.
- [80] $\chi^2 = 5.3, p = 0.02$; $\chi^2 = 29.4, p < 0.001$.
- [81] Again setting alpha at 0.01 for a significant difference.
- [82] $\chi^2 = 15.3, p < 0.001$.
- [83] $\chi^2 = 13.0, p < 0.001$; $\chi^2 = 11.9, p = 0.001$; $\chi^2 = 2.5, p = 0.11$.
- [84] $\chi^2 = 4.7, p = 0.03$; $\chi^2 = 0.1, p = 0.94$; $\chi^2 = 1.9, p = 0.17$; overall 33% versus 41%, $\chi^2 = 2.8, p = 0.096$.
- [85] 22/27 versus 3/41, $\chi^2 = 38.5, p < 0.001$.
- [86] $t = 0.5, p = 0.60$.
- [87] $\chi^2 = 0.0, p = 0.87$. Due to the small number of young female philosophers (26), the exact percentage should be treated cautiously: 95% CI for difference -18% to 21% .
- [88] On other possible psychological differences between philosophers with different normative ethical theories, see Bartels and Pizarro (2011) and Greene (2007).
- [89] Contrast, however, the 2009 Bourget and Chalmers PhilPapers survey: <http://philpapers.org/surveys/>.
- [90] Including only ethicist respondents, $N = 70$ metaethicists, 126 normative ethicists, and 93 applied ethicists. We set the alpha level for statistical significance at 0.01. Among the attitudinal questions, the lowest p value we found was 0.04, 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 0.04, 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 12—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.
- [91] Rating APA membership as good: virtue ethicists 59% versus 80% of deontologists and 74% of consequentialists ($\chi^2 = 8.8, p = 0.012$); rating not having an organ donor symbol as bad: virtue ethicists 48% versus 65% for both other groups ($\chi^2 = 8.8, p = 0.03$); rating blood donation as good 81% versus 94% and 90% ($\chi^2 = 6.6, p = 0.04$). Because of the consistency of the trend, we interpret the p values as significant despite the multiple comparisons.
- [92] Measured APA membership: virtue ethicists 58%, deontologists 65%, consequentialists 74% ($\chi^2 = 5.1, p = 0.08$); organ donation: 58% versus 78% and 70% ($\chi^2 = 6.7, p = 0.04$); but blood donation since 2005 did not split in the direction predicted by expressed normative view: 14% versus 20% and 12% ($\chi^2 = 1.5, p = 0.47$). On the broadest summary measure (measure 2), deontologists scored 4.9, consequentialists 4.6, and virtue ethicists 4.4—within statistical chance (ANOVA, $F = 2.2, p = 0.11$).

References

- Aristotle (4th c. BCE/1962). *Nicomachean ethics* (M. Ostwald, Trans.). New York, NY: Macmillan.
- Baier, A. (1985). *Postures of the mind*. Minneapolis, MN: University of Minnesota.
- Bartels, D.M., & Pizarro, D.A. (2011). The mismeasure of morals: Antisocial personality traits predict utilitarian responses to moral dilemmas. *Cognition*, 121, 154–161.
- Buckwalter, W., & Stich, S. (unpublished manuscript). Gender and philosophical intuition.

- Cushman, F., Young, L., & Hauser, M. (2006). The role of conscious reasoning and intuition in moral judgment: Testing three principles of harm. *Psychological Science*, 17, 1082–1089.
- Fabrigar, L.R., Wegener, D.T., & MacDonald, T.K. (2010). Distinguishing between prediction and influence: Multiple processes underlying attitude-behavior consistency. In C.R. Agnew, D.E. Carlston, W.G. Graziano, & J.R. Kelly (Eds.), *Then a miracle occurs* (pp. 162–185). Oxford: Oxford University Press.
- Greene, J.D. (2007). The secret joke of Kant's soul. In W. Sinnott-Armstrong (Ed.), *Moral psychology* (Vol. 3, pp. 35–79). Cambridge, MA: MIT Press.
- Groves, R. M., Dillman, D. A., Eltinge, J. L., & Little, R.J.A. (Eds.). (2002). *Survey nonresponse*. New York, NY: John Wiley.
- Haddad, E.H., & Tanzman, J.S. (2003). What do vegetarians in the United States eat? *American Journal of Clinical Nutrition*, 78, 626S–632S.
- Hadot, P. (1995). *Philosophy as a way of life* (A. Davidson, Trans.). Malden, MA: Blackwell. (Original work published 1987)
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, 108, 814–834.
- Holbrook, A.L., & Krosnick, J.A. (2010). Social desirability bias in voter turnout reports: Tests using the item count technique. *Public Opinion Quarterly*, 74, 37–67.
- Ivanhoe, P.J. (2000). *Confucian moral self-cultivation* (5th ed.). Indianapolis: Hackett.
- Kant, I. (2002). *Groundwork for the metaphysics of morals* (A. Zweig, Trans.). Oxford: Oxford University Press. (Original work published 1785)
- Knobe, J., & Leiter, B. (2007). The case for Nietzschean moral psychology. In B. Leiter & N. Sinhababu (Eds.), *Nietzsche and morality* (pp. 83–109). Oxford: Oxford University Press.
- Kohlberg, L. (1984). *The psychology of moral development*. Cambridge, MA: Harper & Row.
- Mikhail, J. (2011). *Elements of moral cognition*. Cambridge: Cambridge University Press.
- Mill, J.S. (2003). On liberty. In M. Warnock (Ed.), *Utilitarianism and on liberty* (2nd ed., pp. 88–180). Malden, MA: Blackwell. (Original work published 1859)
- Moeller, H. (2009). *The moral fool*. New York, NY: Columbia.
- Oklahoma State University (2010). *2009–2010 Faculty salary survey by discipline*. Stillwater, OK: Office of Institutional Research & Information Management, Oklahoma State University.
- Paulhus, D.L. (2002). Socially desirable responding: The evolution of a construct. In H.I. Brown, D.I. Jackson, & D.E. Wiley (Eds.), *The role of constructs in psychological and educational measurement* (pp. 49–69). Mahwah, NJ: Lawrence Erlbaum.
- Posner, R.A. (1999). *The problematics of moral and legal theory*. Cambridge, MA: Harvard University Press.
- Rosenthal, R., & Rosnow, R.L. (2008). *Essentials of behavioral research* (3rd ed.). Boston, MA: McGraw Hill.
- Rust, J., & Schwitzgebel, E. (forthcoming). Ethicists' and non-ethicists' responsiveness to student emails: Relationships among expressed normative attitude, self-described behavior, and empirically observed behavior. *Metaphilosophy*.
- Schwitzgebel, E. (2009). Do ethicists steal more books? *Philosophical Psychology*, 22, 711–725.
- Schwitzgebel, E. (forthcoming). Are ethicists any more likely to pay their registration fees at professional meetings? *Economics & Philosophy*.
- Schwitzgebel, E., & Cushman, F. (2012). Expertise in moral reasoning? Order effects on moral judgment in professional philosophers and non-philosophers. *Mind & Language*, 27, 135–153.
- Schwitzgebel, E., & Rust, J. (2009). The moral behaviour of ethics professors: Peer opinion. *Mind*, 118, 1043–1059.
- Schwitzgebel, E., & Rust, J. (2010). Do ethicists and political philosophers vote more often than other professors? *Review of Philosophy and Psychology*, 1, 189–199.
- Schwitzgebel, E., Rust, J., Huang, L., Moore, A., & Coates, J. (2012). Ethicists' courtesy at philosophy conferences. *Philosophical Psychology*, 25, 331–340.

Singer, P. (1972). Famine, affluence, and morality. *Philosophy and Public Affairs*, 1, 229–243.

Singer, P. (2002). *Animal liberation* (3rd ed.). New York, NY: HarperCollins.

Williams, B. (1985). *Ethics and the limits of philosophy*. Cambridge, MA: Harvard University Press.

Wilson, T.D., Dunn, D.S., Kraft, D., & Lisle, D.J. (1989). Introspection, attitude change, and attitude-behavior consistency: The disruptive effects of explaining why we feel the way we do. *Advances in Experimental Social Psychology*, 22, 287–343.