"Ought Implies Can," Framing Effects, and "Empirical Refutations"

Alicia Kissinger-Knox, School of Psychology

Patrick Aragon, School of Psychology

Moti Mizrahi, School of Arts and Communication

Florida Institute of Technology

Abstract: This paper aims to contribute to the current debate about the status of the "Ought Implies Can" (OIC) principle and the growing body of empirical evidence that undermines it. We report the results of an experimental study which show that people judge that agents ought to perform an action even when they also judge that those agents cannot do it and that such "ought" judgments exhibit an actor-observer effect. Because of this actor-observer effect on "ought" judgments and the Duhem-Quine thesis, talk of an "empirical refutation" of OIC is empirically and methodologically unwarranted. What the empirical fact that people attribute moral obligations to unable agents shows is that OIC is not intuitive, not that OIC has been refuted.

Keywords: ability; moral cognition; moral judgments; moral obligation; moral psychology; ought implies can

1. Introduction

Recently, the principle known as "Ought Implies Can" (henceforth, OIC, for short), according to which a person ought to do something only if she can do it, has received some attention from experimental philosophers. First, in a paper (Mizrahi 2015a) and then a reply to commentators (Mizrahi 2015b), Mizrahi reports the results of two experimental studies which show that people attribute moral obligation to agents even when they also judge that those agents cannot perform the requisite action. (Cf. Kurthy and Lawford-Smith 2015.) Second, Buckwalter and Turri (2015) present evidence that corroborates Mizrahi's findings as well as suggest that attributions of blame are sensitive to considerations of ability (or lack of ability). As Buckwalter and Turri (2015, p. 14) put it, "commonsense morality rejects [OIC and accepts] a 'blame implies can' principle." Finally, Chituc et al. (2016) present further evidence that is consistent with the aforementioned findings. They even go so far as to claim that their results amount to "an empirical refutation" of OIC (Henne et al. 2016).[1]

Despite the fact that the results of their experiments are in agreement, insofar as all of them show that participants judge that agents ought to do something even when they also judge that those agents cannot do it, Mizrahi (2015a, 2015b), Buckwalter and Turri (2015), and Chituc et al. (2016) draw different conclusions from these findings. Mizrahi (2015a, p. 234) sets out to test "the alleged intuitiveness" of OIC. For him, then, the fact that participants attribute moral obligations to agents who are unable to do what they ought to shows that "OIC is not intuitive" (Mizrahi 2015b, p. 251). For "If the truth of OIC is intuitive, such that it is accepted by many philosophers as an axiom, then we would expect people to judge that agents who are unable to perform an action are not morally obligated to perform that action" (Mizrahi 2015a, p. 239).

---

[1] For non-empirical arguments concerning OIC, see Graham (2011). Cf. Littlejohn (2012). On OIC from an epistemic point of view, see Mizrahi (2012).

Since what we find is that people judge the agents who are unable to perform an action are still morally obligated to perform that action, it follows that OIC is not intuitive. The upshot, for Mizrahi (2015b, p. 254), is that "OIC can no longer be taken as axiomatic; it must be argued for without appealing to intuitions."

Unlike Mizrahi, Buckwalter and Turri (2015) set out to test the hypothesis that 'ought' implies 'can', i.e., that OIC is true, not the hypothesis that OIC is intuitive. For this reason, they take their results to be showing that "Commonsense moral cognition rejects the principle that ought implies can" (Buckwalter and Turri 2015, p. 16). They add that "Moral obligations persist with or without ability" (Buckwalter and Turri 2015, p. 16), as opposed to *judgments* about moral obligations and (in)ability, thereby suggesting that they take their results to be showing that it is not the case that 'ought' implies 'can'. Taking their results to show that "ought does not imply can" (Buckwalter and Turri 2015, p. 14), Buckwalter and Turri then wonder whether "blame implies can." They take the results of their experiments to show that, "although ordinary moral cognition does not endorse an 'ought implies can' principle, it may support a closely related principle, namely 'blame implies can'" (Buckwalter and Turri 2015, p. 15).

Like Buckwalter and Turri (2015), Chituc et al. (2016) also set out to test the semantic properties of OIC rather than the hypothesis that OIC is intuitive. In particular, they set out to test the hypothesis that "'ought' analytically or conceptually implies 'can'" (Chituc et al. 2016, p. 21). As Chituc et al. (2016, p. 20) explain, to say that "'ought' analytically or conceptually implies 'can'" (Chituc et al. 2016, p. 21) is to say that "'ought' is supposed to imply 'can' by virtue of the concepts expressed by the words 'ought' and 'can', just as 'bachelor' implies 'male' by virtue of the concepts expressed by the words 'bachelor' and 'male'" (Chituc et al. 2016, p.

20).[2] For this reason, they say that "[OIC] is supposed to follow necessarily from the concepts expressed by the words 'ought' and 'can'" and claim that their "results show that it does not" (Chituc et al. 2016, p. 23). In fact, they even refer to their findings as an "empirical refutation" of OIC, claim that "OIC is not true analytically or conceptually" (Henne et al. 2016, p. 288), and "propose the end of OIC as it is traditionally understood" (Henne et al. 2016, p. 289). Their claim to have refuted OIC has even received some attention in popular media, with an article in *The New York Times*' Sunday Review, entitled "The Data against Kant" (Chituc and Henne 2016). The hypotheses and predictions of the three experimental studies on OIC discussed above are summed up in Table 1.

Table 1. Hypotheses and predictions of the aforementioned experimental studies on OIC.

Study	Hypothesis	Prediction
Mizrahi (2015a)	OIC is intuitive	"If the truth of OIC is intuitive, such that it is accepted by many philosophers as an axiom, then we would expect people to judge that agents who are unable to perform an action are not morally obligated to perform that action" (Mizrahi 2015a, p. 232).
Buckwalter and Turri (2015)	OIC is true	"Our primary question is whether perceptions of ability limit judgments about whether a broad range of moral requirements are present, or if moral requirements are attributed to agents independently of the presence of an ability to fulfill them" (Buckwalter and Turri 2015, p. 2).
Chituc et al. (2016)	OIC is analytic	"If 'ought' analytically or conceptually implies 'can', [...] then participants should deny that the agent 'ought' to do something if they learn that the agent can't do it" (Chituc et al. 2016, p. 21).

---

[2] Even before experimental philosophers started testing OIC empirically, it should have been quite clear that OIC can no longer be "treated as an axiom" (Howard-Snyder 2013). This is because there are at least three candidates for the supposed relation between 'ought' and 'can' (see Howard-Snyder 2013; Mizrahi 2009; Mizrahi 2015a). The first candidate is logical implication, i.e., entailment. The fact that "'ought' *entails* 'can'" is open to counterexamples (see, e.g., Mizrahi 2009 and King 2014), however, has led many to abandon it as the relation between 'ought' and 'can' and propose two other candidates: presupposition (Hare 1951) and conversational implicature (Sinnott-Armstrong 1984). These two candidates for the relation between 'ought' and 'can' are also open to counterexamples (see, e.g., Mizrahi 2009 and King forthcoming), however, and so it is far from clear how 'ought' is supposed to *imply* 'can'. As an anonymous reviewer helpfully pointed out, there are also systems of deontic logic that do not treat OIC as an axiom. See Saka (2000) for modal arguments against OIC. See also Martin (2009). For more on deontic logic and OIC, see Part I of Tessman (2015).

In this paper, we seek to contribute to this ongoing debate about the status of OIC and the growing body of empirical evidence that undermines it by (a) reporting the results of our own experimental survey of judgments about ability and moral obligation, and (b) trying to adjudicate between the different conclusions that can be drawn from such empirical findings. Our results indeed show that people do judge that agents are morally obligated to perform an action even when they also judge that those agents cannot perform the action. Our results also show that people's moral judgments (i.e., judgments about what agents ought to do) exhibit an actor-observer effect,[3] as we might expect from previous experimental studies in moral psychology and experimental philosophy. For this reason, we argue, the empirical fact that people attribute moral obligation to unable agents cannot be taken as "a powerful counterexample to" (Buckwalter and Turri 2015, p. 5) or an "empirical refutation" (Chituc et al. 2016; Henne et al. 2016) of OIC.

2.Study

2a. Methods

Our study was designed to test people's intuitive judgments about cases of moral obligation (ought) and ability (can). In order to adjudicate between the aforementioned conclusions drawn from the empirical evidence on OIC, namely, whether the fact that people attribute moral obligation to agents who are unable to perform the requisite action constitutes an "empirical refutation" of OIC, as Chituc et al. (2016) contend, a "powerful counterexample to OIC," as Buckwalter and Turri (2015, p. 5) contend, or evidence that OIC is not as intuitive as

---

[3] Buckwalter and Turri (2015) also tested for framing effects, but Chituc et al. (2016) did not. We will say more on this in Section 2c.

philosophers commonly think, as Mizrahi (2015b) contends (see Table 1), we manipulated the following variables:

Circumstances (Able/Unable): We gave some participants vignettes in which the agent is able to perform an action and other participants vignettes in which the agent is unable to perform an action.

Agent (Alex/You): We gave some participants vignettes in which they are observers (the agent is "Alex") and other participants vignettes in which they are actors (the agent is "You").

Participants for this study were recruited through Amazon Mechanical Turk and were tested through Qualtrics, an online survey platform. They were compensated $0.15 for approximately two minutes of their time. A total of 1,084 participants were recruited for this study. Of this sample, 3 participants were excluded because they did not consent to the informed agreement and 11 participants were excluded for not completing the questionnaire. This resulted in a sample of 1,070 total participants. Although demographics had no statistical effects, it is worth noting that our sample was extremely diverse, with participants (ages 18-65) claiming at least three different genders (624 men, 433 women, and 6 other), 17 different ethnic backgrounds, and 25 distinct religious affiliations.

After they were given written informed consent, participants were randomly given one vignette in a between-subjects experimental design that read as follows:

You are [/Alex is] on a train. A person with a broken leg who uses crutches to walk boards the train. You remain [/Alex remains] seated even though there are no available seats on the train.

In the vignettes in which the agent is unable to perform the action (i.e., the agent is unable to give up a seat on a train for a disabled person), the second sentence read as follows:

> You are [/Alex is] sleeping and remain[/s] seated even though there are no available seats on the train.

After reading one of the vignettes, participants were randomly assigned to one of the following statements and asked to indicate their agreement or disagreement with them on a standard Likert scale from 1 ("Strongly Disagree") to 7 ("Strongly Agree"):

1. You [/Alex] can give up the seat.

2. You [/Alex] ought to give up the seat.

We have chosen a between-subject design, randomly assigning each participant to a vignette with one of the above statements, in order to avoid interference between the conditions and so make our results more generalizable. After indicating their agreement or disagreement with one of these statements, participants were asked to complete a brief demographics questionnaire.

2b. Results

Likert-type data, like the kind we collected, is ordinal data. This means that the values assigned (1-7) have an order but the distances between values are not equal intervals.[4] Therefore, Likert-type data requires a special type of analysis. To begin with, we looked at the descriptive statistics and graphical representations of the data. This allows us to see the central tendencies in the data and which groups differ the most.[5] Ordinal data cannot be described using averages because

---

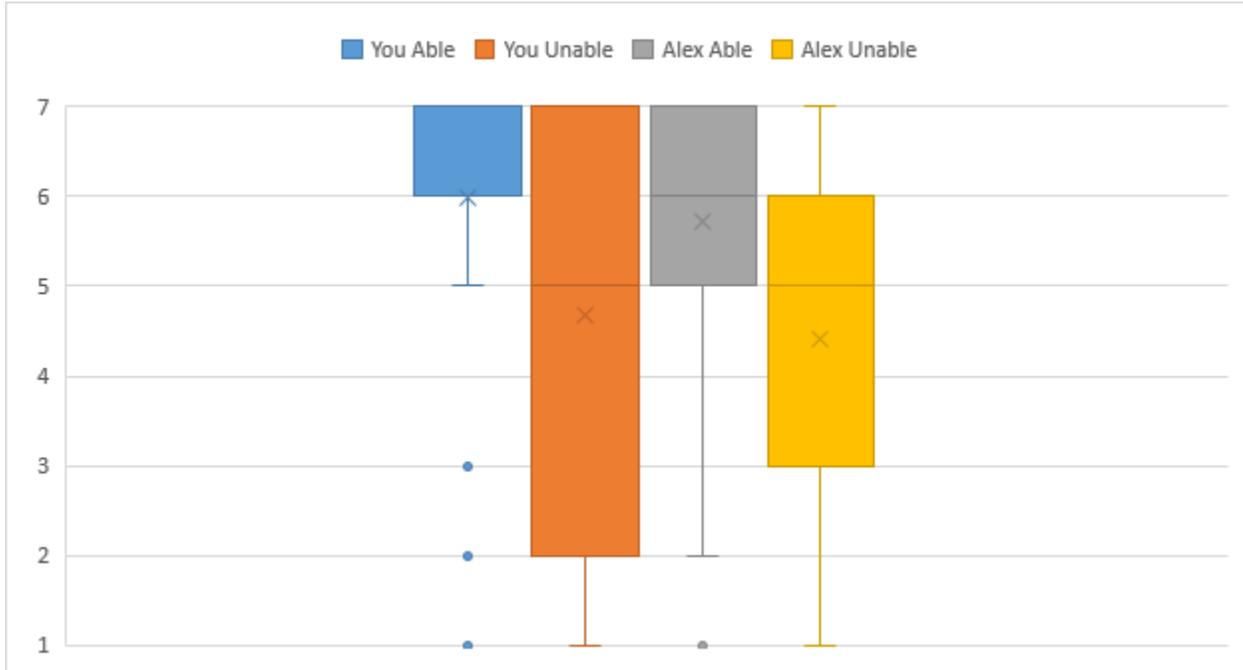[4] See Knapp (1990) for an overview of the controversy about treating ordinal data as interval data.

[5] We have chosen not to dichotomize the data into agree/disagree because we are interested in the strength of agreement as well as the polarity. The strength itself will also tell us which classification judgments would be in. So if the judgment is left as a 7 and not a binary choice, we know it is a strong "ought" judgment and we still know it was agreement with "ought." But if we dichotomize, all we could know is that it was agreement with "ought." For more on the problems with dichotomization, see MacCallum et al. (2002).

there is no meaning for a result that falls between two values. Instead, we look at medians, modes, frequencies, and quartiles. The best graphical representation of this is boxplots.[6]

Figure 1 shows the data for judgments about "can" statements (namely, "You/Alex can give up the seat"). We can see the data are heavily skewed toward "strongly agree" for the vignettes in which the agent (either "You" or "Alex") is able and more toward "somewhat agree" for the vignettes in which the agent (either "You" or "Alex") is unable. Both the median and mode for the vignettes in which the agent ("You") is able are 7 indicating that the most common judgment and the central tendency of the sample are both strong agreement. Vignettes in which the agent ("You") is unable have a median of 5 and a mode of 7; so the central tendency is "somewhat agree" while the most common judgment is "strongly agree." Vignettes in which the agent ("Alex") is able have both a median and mode of 6. This indicates that the central tendency and most common judgment are ones of agreement. For the vignettes in which the agent ("Alex") is unable, the median is 5 and the mode is 6, again highlighting levels of agreement in both statistics.
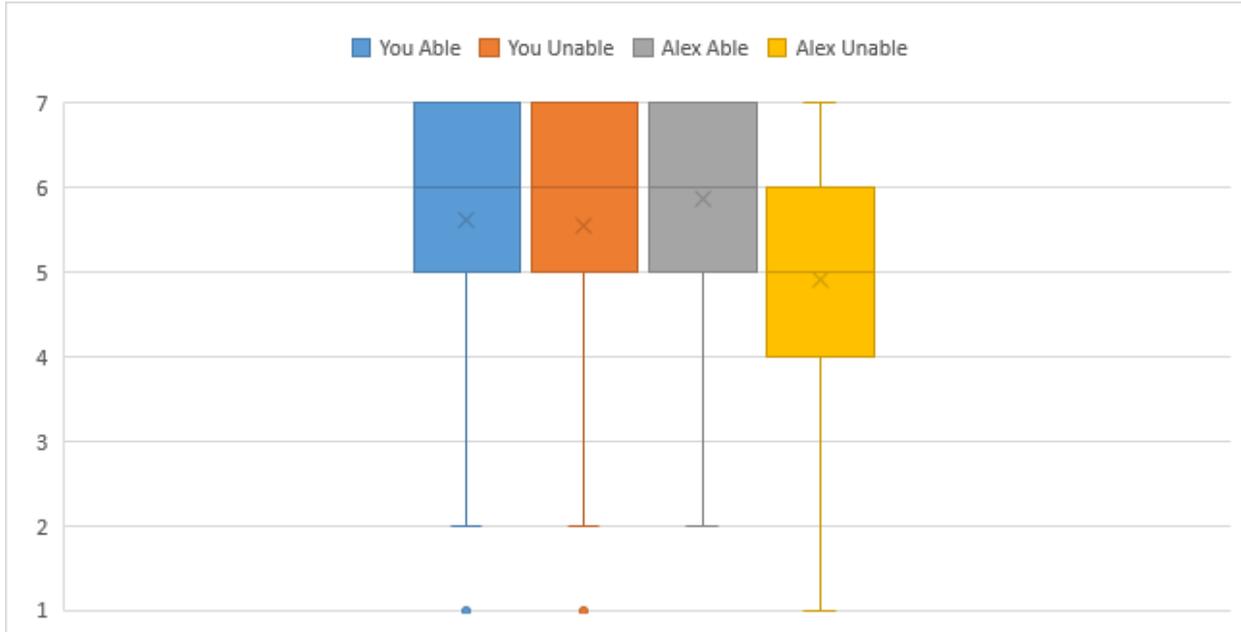
---

[6] We mention all of this in detail in order to explain our statistical choices, which throughout this paper will include less common but more specialized testing. We want to make sure that readers are able to reproduce or replicate our results.

Figure 1. Distribution of agreement with "can" statements.

The data for judgments about "ought" statements (namely, "You/Alex ought to give up the seat") show a somewhat different result (see Figure 2). The results are mainly homogeneous between groups. Overall, participants agree with the "ought" statements (namely, "You/Alex ought to give up the seat"). The central tendencies for vignettes in which the agent ("You") is able, the agent ("You") is unable, and the agent ("Alex") is able are all 6 for agreement. The median for vignettes in which the agent ("Alex") is unable is slightly lower at 5 for "somewhat agree." The most common judgment for vignettes in which the agent (either "You" or "Alex") is unable is 6 for "agree," whereas the most common judgment for vignettes in which the agent (either "You" or "Alex") is able is 7 for "strongly agree."

Figure 2. Distribution of agreement with "ought" statements.

Graphical analyses and descriptive statistics, while useful for understanding the data's

distribution, cannot give us robust information about significant effects. Because means and

variances are meaningless for ordinal data, ANOVA and similar measures of variance are

inappropriate. We chose to perform an Aligned Rank Transform (ART) on the data.[7] This

transform produces statistics from the ordinal data on which regular ANOVA tests can be

performed and analyzed in the usual way. From this method we were able to identify what

factors had significant effects on "can" and "ought" statements.

Agreement with "can" statements is significantly affected by circumstances ($p = 4.11$ x

10^{-05}) but not agent or interaction. Both circumstances ($p = 6.19$ x 10^{-05}) and agent ($p = 0.017$)

have significant effects on agreement with "ought" statements but interaction effects were

---

[7] As common as ANOVA and other interval/ratio tests are in testing Likert data, it is not an appropriate choice and will often lead to erroneous results. For a detailed explanation of ART procedures, see Wobbrock et al.. (2011). We performed the statistical analysis in R using the ARTool package provided by Wobbrock et al.. An overview of the tool and package can be accessed at https://depts.washington.edu/aimgroup/proj/art/.

insignificant. As we mentioned earlier, we did not find demographic factors such as religion, gender, and education to have significant effects on responses.

Knowing that the interaction did not significantly affect agreement with "can" and "ought" statements, we are able to separate the groups and look at single factor ordinal analyses to test other hypotheses. In line with the experimental studies on OIC discussed in Section 1, we would like to see whether participants disagree with the statement "You/Alex ought to give up the seat" when the circumstances make the agent (either "You" or "Alex") unable to give up the seat. To test this, we performed Wilcoxon Rank Sum tests. These tests are appropriate for ordinal data like ours. They test the null hypothesis "true location shift is equal to 0" between two groups in a between-subjects design. When the null hypothesis is rejected, we can say that one group has a shift greater or less than 0 when compared to the other group. The location considered for this test is the central tendency, or median.

In the vignettes in which the agent is "You," there is no shift in location based on ability ($p = 0.117$). In the vignettes in which the agent is "Alex," there is a shift in location based on ability ($p = 2.19 \times 10^{-4}$). Based on a one-tailed test, we determined that the shift is greater than 0 ($p = 1.10 \times 10^{-04}$) indicating that participants tend to agree more with "Alex ought to give up the seat" if Alex is able than if Alex is unable. The median is 6 for all cases except when Alex is unable; in that case, it is 5. So, overall participants still agree that the agent (either "You" or "Alex") ought to do something even when the agent is unable to do it.

We also investigated the actor-observer effect we mentioned above by comparing data where the agent is "Alex" with data where the agent is "You." This was done using Wilcoxon Rank Sum tests as well. In the vignettes where the agent is able, change in agent caused no shift in location ($p = 0.656$). In the vignettes where the agent is unable, however, there was a

significant shift in location (p = 0.011). Using a one-tailed test, we were able to determine that

the shift is greater than 0 (p = 0.005), which indicated that participants are more likely to make

"ought" judgments for "You" than for "Alex" when both are unable.

This actor-observer effect revealed something important about the way people make

moral judgments. That is, it suggests that there is a significant difference between the way people

judge moral statements about others, like "*Alex* ought to give up the seat," and the way they

judge moral statements about themselves, like "*You* ought to give up the seat." As we know from

previous studies in moral psychology and experimental philosophy, people's moral judgments

are affected by biases and factors that are supposed to be irrelevant to the truth or falsity of such

moral judgments. As Buckwalter himself puts it:

> [Research in social psychology and experimental philosophy] has shown that ordinary
>
> people's moral intuitions are influenced by a variety of factors including order effects
>
> […], framing effects […], and environmental variables […]. Since it is widely agreed that
>
> those factors are irrelevant to the truth or falsity of the intuition, these empirical results
>
> cast doubt on the use of intuition as evidence for moral claims (Tobia et al. 2013, p. 630).

For instance, the point of view from which hypothetical scenarios about morality (e.g., the trolley

problem) are presented is supposed to be irrelevant to the truth or falsity of the judgments people

make about those scenarios. So, if it is morally wrong to push the large person off the bridge in

order to save the lives of five workers over the life of one worker on the tracks, then, from a

moral point of view, it should not matter whether the "push" scenario is presented from the point

of view of an actor or an observer. If the point of view does make a difference, as researchers

have found (see, e.g., Nadelhoffer and Feltz 2008), then that is a reason to question the reliability

of such moral judgments. Similarly, our data show that whether a scenario is described from the

point of view of an actor (in our study, these are the vignettes in which the agent is "You") or from the point of view of an observer (in our study, these are the vignettes in which the agent is "Alex") does make a difference to the moral judgments people make about these scenarios. Therefore, the results of our study provide reasons to question the reliability of such intuitive moral judgments, specifically, "ought" judgments, as evidence for or against philosophical theses (like OIC).

To sum up, our main findings are that people judge that agents ought to do something even when they also judge that those agents can't do it and that people's "ought" judgments exhibit an actor-observer effect. That is to say, there is a significant difference between the way people make moral judgments about what moral agents ought to do when those judgments are about someone else ("Alex") than when they are about themselves ("You"). In the next section, we discuss whether these findings amount to a counterexample to OIC, as Buckwalter and Turri (2015) contend, refute OIC, as Chituc et al. (2016) contend, or show that OIC is not intuitive, as Mizrahi (2015b) contends.

2c. Discussion

Although our findings are in line with the results reported by Mizrahi (2015a), Mizrahi (2015b), Buckwalter and Turri (2015), and Chituc et al. (2016), namely, that people's attributions of moral obligations are not sensitive to the ability (or inability) of the moral agent in question, they do raise an important, but hitherto overlooked in the context of the debate over OIC, methodological concern. The worry is about the use of intuitive moral judgments as evidence for and against philosophical theses like OIC. As we know from previous studies in moral psychology and experimental philosophy, people's judgments are affected by factors that are supposed to be irrelevant to the truth or falsity of those judgments. As Sinnott-Armstrong (2006,

p. 354) himself puts it: "Framing effects distort moral beliefs in so many cases that moral believers need confirmation for any particular moral belief."

For instance, several studies have shown that people's moral judgments are subject to order effects (see, e.g., Schwitzgebel and Cushman 2012). Presumably, the order in which hypothetical scenarios about morality (e.g., the trolley problem) are presented is supposed to be irrelevant to the truth or falsity of the judgments people make about those scenarios. So, if it is morally wrong to push the large person off the bridge in order to save the lives of five workers over the life of one worker on the tracks, then, from a moral point of view, it should not matter whether the "push" scenario is presented before or after the "switch" scenario. If the order of presentation does make a difference, as researchers have found (see, e.g., Schwitzgebel and Cushman 2012), then that is a reason to question the reliability of such intuitive moral judgments. Other studies have found that even the judgments of experts (e.g., professional philosophers) are not immune to order effects, framing effects, and other sorts of biases (see, e.g., Schwitzgebel and Cushman 2015). This has led many experimental philosophers to conclude that the evidential status of intuitive moral judgments in response to hypothetical scenarios (or "intuitions") is questionable (see, e.g., Stich and Tobia 2016, pp. 5-21 and Mizrahi 2015c).[8]

Now, like the other experimental studies on OIC mentioned above, our experimental survey asks participants to make intuitive moral judgments; that is, intuitive judgments about what agents ought to (or ought not) do. As is known by now, however, such intuitive moral judgments are affected by all sorts of biases and factors that are irrelevant to the truth or falsity of those judgments. So we should not be surprised if we find that people's "ought" judgments are affected in similar ways. Indeed, as we reported in Section 2b, that is precisely what we have

---

[8] For more on intuitions and experimental philosophy, see Andow (2016a) and (2016b).

found. That is, participants' "ought" judgments show an actor-observer effect. But this means that the evidential status of these intuitive moral judgments is doubtful. And yet that is precisely what Buckwalter and Turri (2015), and Chituc et al. (2016), want to do; that is, they want to use intuitive moral judgments as conclusive evidence against a philosophical thesis. More specifically, Buckwalter and Turri (2015) want to use "ought" judgments as conclusive evidence against or, as they put it, "a powerful counterexample to OIC" (Buckwalter and Turri 2015, p. 5); that is, against the philosophical principle itself, not how people think about what the principle states. Chituc et al. (2016) even go so far as to say that the empirical fact that people attribute moral obligation to agents who are unable to perform the obligatory action constitutes an "empirical refutation" of OIC. That is problematic, however, given the fact that intuitive moral judgments are influenced by biases and other irrelevant factors. As experimental philosophers, Buckwalter, Turri, and Chituc et al. ought to know better.

Indeed, Buckwalter and Turri (2015) have themselves tested whether or not "ought" judgments are affected by perspective, i.e., by whether a vignette is presented from a first-person point of view (where the participant is an actor) or a third-person point of view (where the participant is an observer). Buckwalter and Turri (2015, p. 7) report that "[p]articipants in the Actor conditions were more likely to select [the "Obligated but Unable"] response than participants in Observer conditions." But this only makes their attempt to use intuitive moral judgments as evidence that "ought does not imply can" (Buckwalter and Turri 2015, p. 14) even more puzzling. As experimental philosophers, they are surely aware that such "empirical results cast doubt on the use of intuition as evidence for moral claims" (Tobia et al. 2013, p. 630). In fact, their own empirical results cast doubt on the use of "ought" judgments about unable agents as evidence (conclusive or not) against OIC.

Framing effects, order effects, and other cognitive biases undermine the evidential status of intuitive moral judgments because they suggest that such judgments are unreliable. As Devitt (2015, p. 687) puts it:

> What philosophers *do* think, or at least *should* think, is that intuitions are a source of evidence *because they are reliable*. And if they really are reliable, which does not of course require them to be infallible, then they are indeed a source of evidence. (And, one might add, this is true of any judgment, whether intuitive or not.) (emphasis in original).

Being reliable, then, is a necessary (but perhaps not sufficient) condition for being a trustworthy source of evidence. As mentioned above, however, there is empirical evidence from moral psychology and experimental philosophy suggesting that intuitive moral judgments are biased. In accordance with this empirical evidence, our experimental results suggest that "ought" judgments exhibit an actor-observer bias. Now, biased judgments are not reliable judgments. Being reliable, however, is a necessary condition for being a trustworthy source of evidence. Since intuitive moral judgments fail to meet this necessary condition, given that they are biased, it follows that they cannot serve as evidence for or against philosophical theses, such as OIC. Again, as experimental philosophers, Buckwalter and Turri (2015) and Chituc et al. (2016) are surely familiar with these experimental studies about the unreliability of intuitions, which is why it is puzzling that they try to use intuitive moral judgments to advance so-called "powerful counterexamples" (Buckwalter and Turri 2015, p. 5) and so-called "empirical refutations" (Henne et al. 2016) against OIC.[9]

In addition to the fact that "ought" judgments in particular, like intuitive moral judgments in general, are subject to biases and other factors that are supposed to be irrelevant from a moral

---

[9] For additional arguments to the effect that intuitive judgments are unreliable sources of evidence, see Mizrahi (2015c).

point of view, like an actor-observer effect, there is another, non-empirical, reason why Buckwalter and Turri's (2015) talk of "powerful counterexamples" and Chituc et al.'s (2016) talk of "empirical refutations" are problematic. As philosophers of science since Duhem and Quine have pointed out, there is no such thing as an "empirical refutation." As Quine (1951, p. 38) elegantly puts it, "our statements about the external world face the tribunal of sense experience not individually but only as a corporate body." In philosophy of science, this is known as the Duhem-Quine thesis or confirmation holism. According to the Duhem-Quine thesis, when a theory makes a prediction that is not borne out by experimentation or observation, logic alone does not tell us whether to reject the theory or any number of auxiliary assumptions that were made in order to derive the prediction in the first place. As Okasha (2002, p. 306) puts it:

> According to the doctrine of confirmation holism, also known as the 'Quine-Duhem' thesis, the empirical content of a scientific theory cannot be parcelled out individually among the constituent components of the theory. Thus when a theory makes an empirical prediction which turns out to be false, it will not be automatically obvious where to lay the blame, i.e., which component of the theory to reject. Logic tells us there is an error somewhere in the set of statements which implies the false prediction, but does not tell us where. So there will be various ways of modifying our theory to inactivate the false implication.

Accordingly, when prediction P of theory T is not borne out by experimentation or observation, we cannot simply derive the negation of T as follows:

$T \rightarrow P$

$\sim P$

$$\therefore \sim T$$

This is because, in deriving *P* from *T*, we have made some auxiliary assumptions, and *P* follows

from that set of statements, namely, *T* and $A_1, A_2, A_3, \ldots, A_n$, not from *T* alone. So when *P* turns

out to be false, all we are justified in concluding is that the conjunction $T \& A_1 \& A_2 \& \ldots A_n$ is

false, but not that *T* is false. That is:

$$(T \& A_1 \& A_2 \& A_3 \& \ldots A_n) \rightarrow P$$

$$\sim P$$

$$\therefore \sim(T \& A_1 \& A_2 \& \ldots A_n)$$

To assume that one can simply "refute" a theory empirically is to presuppose some form of naive

falsificationism (Kitcher 1982, p. 44). Since naive falsificationism is false, a "refutation" that

presupposes it is no refutation at all (Mizrahi 2015d).

Accordingly, when Chituc et al. (2016) set out to test OIC empirically, they explicitly

mention some of the auxiliary assumptions they have made in order to derive the prediction that

"*can* and *ought* judgments [should] correlate" (Chituc et al. 2016, p. 22). For instance, they

assume that their sample of participants contains "competent speakers [of English, presumably]"

and that these "competent speakers" are in "good epistemic conditions" (Henne et al. 2016, p.

284).[10] However, they do not explicitly mention two key auxiliary assumptions that they must

assume in order to test OIC experimentally and conclude that OIC is not analytic. These

auxiliary assumptions are the following:

---

[10] As an anonymous reviewer helpfully pointed out, it is important to note here that, like Mizrahi (2015a) and Chituc et al. (2016), we did not ask participants about their native language, whereas Buckwalter and Turri (2015) did ask participants about their native language In all of their experiments, more than 90% of participants have reported that English is their first language. Even though the "official language" of morality is not English, presumably, this is important to note because our survey materials, like those of Mizrahi (2015a), Mizrahi (2015b), Buckwalter and Turri (2015), and Chituc et al. (2016), use English modals for the concepts of moral obligation and ability. We think that it would be interesting to conduct further research in order to find out whether participants' judgments would vary across different languages.

R: The participants' intuitive judgments are not biased or otherwise influenced by factors

that are irrelevant to the truth of intuitive judgments about moral obligation and ability.

J: Making intuitive judgments in response to hypothetical cases about morality (e.g.,

whether an agent is morally obligated to do something) is a reliable method of

discovering the semantic properties (e.g., truth, analyticity, etc.) of philosophical

principles (e.g., OIC).

Their prediction that "*can* and *ought* judgments [should] correlate" (Chituc et al. 2016, p. 22), or

that "competent speakers should deny that an agent ought to do an act when they understand that

the agent cannot do the act" (Henne et al. 2016, p. 284), then, follows from OIC only on the

assumption that *R* and *J* are the case. For if it were not the case that *R*, then the prediction that

"competent speakers should deny that an agent ought to do an act when they understand that the

agent cannot do the act" (Henne et al. 2016, p. 284) would not follow from the hypothesis that

'ought' analytically or conceptually entails 'can', since the intuitive moral judgments that Chituc

et al. (2016) were to look at in this case would be biased or influenced by factors that are

irrelevant to the truth of judgments about moral obligation and ability (i.e., factors other than the

truth or falsity of OIC itself), and hence untrustworthy as evidence that is supposed to amount to

an "empirical refutation" of OIC. Similarly, if it were not the case that *J*, then the prediction that

"competent speakers should deny that an agent ought to do an act when they understand that the

agent cannot do the act" (Henne et al. 2016, p. 284) would not follow from the hypothesis that

'ought' analytically or conceptually entails 'can', since the intuitive moral judgments that Chituc

et al. (2016) were to look at in this case would again be unreliable, and hence untrustworthy as

evidence that is supposed to amount to an "empirical refutation" of OIC.

Accordingly, with the aforementioned assumptions now made explicit, Chituc et al.'s (2016) so-called "empirical refutation" of OIC actually runs as follows:

1. If 'ought' analytically or conceptually entails 'can' and *R* and *J*, then "competent speakers should deny that an agent ought to do an act when they understand that the agent cannot do the act" (Henne et al. 2016, p. 284).

2. "Competent speakers assert that an agent ought to do an act when they understand that the agent cannot do the act" (Henne et al. 2016, p. 284).

3. Therefore, it is not the case that 'ought' analytically or conceptually entails 'can' and *R* and *J*.

For a conjunction to be false, only one of the conjuncts has to be false, and logic alone does not tell us which of the components of (3) is false. It may be the case that "'ought' analytically or conceptually entails 'can'" is false. But it may also be the case that *R* is false or that *J* is false. Logic and Chituc et al.'s (2016) experimental results alone do not tell us which of the components of (3) is to be rejected.

Similar remarks apply to Buckwalter and Turri's (2015, p. 5) talk about "a powerful counterexample to OIC." That is to say, when they set out to test OIC empirically, Buckwalter and Turri (2015) make assumptions without which they could not test OIC empirically. For instance, they assume that their sample of Amazon Mechanical Turk workers is a representative sample. They assume that these Amazon Mechanical Turk workers understand the meaning of words like 'ought', 'obligation', 'unable', etc. They assume that the "moral terminology" (Buckwalter and Turri (2015, p. 16) used in their experiments accurately tracks the concepts of moral obligation, ability, and the like. In other words, Buckwalter and Turri (2015) make pretty much the same assumptions that Chituc et al. (2016) make in order to test OIC empirically.

Crucially, their prediction that "moral requirements are attributed to agents independently of the presence of an ability to fulfill them" (Buckwalter and Turri 2015, 2), follows from OIC only on the assumption that *R* and *J* are the case. For if it were not the case that *R*, then the prediction that "moral requirements are attributed to agents independently of the presence of an ability to fulfill them" (Buckwalter and Turri 2015, 2) would not follow from "the principle that ought implies can" (Buckwalter and Turri 2015, p. 16), since the intuitive moral judgments that Buckwalter and Turri (2015) were to look at in this case would be biased or influenced by factors that are irrelevant to the truth of judgments about moral obligation and ability (i.e., factors other than the truth or falsity of OIC itself), and hence untrustworthy as evidence that is supposed to amount to a "powerful counterexample" to OIC. Similarly, if it were not the case that *J*, then the prediction that "moral requirements are attributed to agents independently of the presence of an ability to fulfill them" (Buckwalter and Turri 2015, 2) would not follow from "the principle that ought implies can" (Buckwalter and Turri 2015, p. 16), since the intuitive moral judgments that Buckwalter and Turri (2015) were to look at in this case would again be unreliable, and hence untrustworthy as evidence that is supposed to amount to a "powerful counterexample" to OIC.

Accordingly, with the aforementioned assumptions now made explicit, Buckwalter and Turri's (2015) attempt to empirically "disprove" OIC by "powerful counterexamples" actually runs as follows:

1. If (OIC & *R* & *J*), then participants would not attribute moral obligation to unable agents.

2. Participants attribute moral obligation to unable agents.

3. Therefore, it is not the case that (OIC & *R* & *J*).

Again, for a conjunction to be false, only one of the conjuncts has to be false, and logic alone does not tell us which of the components of (3) is false. It may be the case that OIC is false. But

it may also be the case that *R* is false or that *J* is false. Logic and Buckwalter and Turri's (2015) experimental results alone do not tell us which of the components of (3) is to be rejected.

In fact, our experimental results provide a reason to think that the problematic conjunct in both Buckwalter and Turri's (2015) and Chituc et al.'s (2016) conclusions, i.e., (3), is *R*. As we have seen, "ought" judgments, like intuitive moral judgments in general, are subject to biases and other factors that are supposed to be irrelevant from a moral point of view (in particular, an actor-observer effect). If such intuitive moral judgments are biased, or influenced by factors that are irrelevant to their truth or falsity, then the reason participants judged that agents ought to perform an action when they can't may have nothing to do with the truth or analyticity of OIC, but rather something to do with prejudices and biases. That is why Buckwalter and Turri's (2015) talk of a "powerful counterexample" to OIC and Chituc et al.'s (2016) talk of an "empirical refutation" of OIC are not only empirically unwarranted but also methodologically problematic.[11]

Of course, just like Mizrahi (2015a), Buckwalter and Turri (2015), and Chituc et al. (2016), in conducting our own experimental survey of intuitive judgments about cases of moral obligation (ought) and ability (can), we have also made some auxiliary assumptions as well, such as assumptions about the trustworthiness of Mechanical Turk workers and the reliability of survey procedures like the ones we have used in our study, which is why we do not claim to have "disproved by counterexamples" or "empirically refuted" anything. Unlike Buckwalter and Turri (2015) and Chituc et al. (2016), however, we did not--and *need not*--assume that intuitive moral judgments are not biased, i.e., we do not assume that *R* is the case, for we do not use the intuitive moral judgments our participants have made in response to our vignettes as evidence (either conclusive or defeasible) against OIC. Rather, we take the fact that our participants judged that

---

[11] For arguments to the effect that *J* is problematic, too, see Mizrahi (2015d).

agents who are unable to perform an action ought to perform it nonetheless as evidence that OIC

is not *intuitive*. For, if OIC were *intuitive*, it would have *seemed* to our participants that an agent

is not morally obligated to do what she cannot do.[12] More explicitly, since intuitions are

intellectual appearances or seemings (i.e., it seems to *S* that *p*), if OIC were intuitive, then it

would have seemed to our participants that an agent ought to perform an action only if the agent

can perform that action. The fact that our participants judged that an agent (either "Alex" or

"You") ought to perform an action even when the agent is unable to perform that action suggests

that it did not seem to them that an agent ought to perform an action only if the agent can

perform that action. In other words, to them, OIC is not intuitive.

The key auxiliary assumptions, along with the hypotheses and predictions of the three

experimental studies on OIC discussed above, are summed up in Table 2.

Table 2. Hypotheses, predictions, auxiliary assumptions, and conclusions of the aforementioned

experimental studies of OIC.

Study	Hypothesis	Key Auxiliary Assumption	Prediction	Conclusion
Mizrahi (2015a)	OIC is intuitive	Intuitions are seemings	"we would expect people to judge that agents who are unable to perform an action are not morally obligated to perform that action" (Mizrahi 2015a, p. 232).	"OIC is not intuitive" (Mizrahi 2015b, p. 251).
Buckwalter and Turri (2015)	OIC is true	*R*; *J*	"Our primary question is whether perceptions of ability limit judgments about whether a broad range of moral requirements are present, or if moral requirements are attributed to agents	"ought does not imply can" (Buckwalter and Turri 2015, p. 14).

---

[12] On intuitions as intellectual appearances or seemings, see Brogaard (2014).

			independently of the presence of an ability to fulfill them" (Buckwalter and Turri 2015, p. 2).	
Chituc et al. (2016)	OIC is analytic	*R*; *J*	"participants should deny that the agent 'ought' to do something if they learn that the agent can't do it" (Chituc et al. 2016, p. 21).	"OIC is not true analytically or conceptually" (Henne et al. 2016, p. 288).

As mentioned above, all three experimental studies on OIC discussed above, as well as our own experimental study, share some auxiliary assumptions, such as assumptions about the reliability of survey procedures using Qualtrics, the trustworthiness of Mechanical Turk workers, the words 'ought' and 'can' and the concepts of moral obligation and ability. Crucially, however, Buckwalter and Turri (2015) and Chituc et al. (2016) make auxiliary assumptions that Mizrahi (2015a) does not. As far as Buckwalter and Turri's (2015) study goes, without assuming that intuitive moral judgments are unbiased and that making intuitive moral judgments in response to hypothetical cases about morality is a reliable method of discovering the truth or falsity of philosophical principles (like OIC), it would not follow that OIC is false from the fact that "[m]oral requirements were attributed [by participants] independently of ability," as Buckwalter and Turri (2015, pp. 2-3) want to argue. As far as Chituc et al.'s (2016) study goes, without assuming that intuitive moral judgments are unbiased and that making intuitive judgments in response to hypothetical cases about morality is a reliable method of discovering the semantic properties (like analyticity) of philosophical principles (like OIC), it would not follow that OIC is not analytic from the fact that "participants readily say that an agent ought to do what they know the agent cannot do," as Henne et al. (2016, p. 286) want to argue. Like Mizrahi (2015a), we do not need to assume *R* and *J* for our results to show that OIC is not intuitive. For we do not

pretend to have discovered anything about the semantic properties (i.e., truth or analyticity) of the meta-ethical principle OIC from intuitive moral judgments, especially since such judgments are biased and appealing to them as evidence is methodologically problematic. Rather, our results simply show that people do not find OIC intuitive. So, like Mizrahi (2015a), we only need to assume that intuitions are seemings.

For these reasons, we think that the appropriate conclusion to draw from the empirical fact that people attribute moral obligations to agents even when those agents are unable to fulfill those obligations is that OIC is not intuitive, not that OIC is false or that OIC has been "empirically refuted." After all, the empirical fact is about people's *intuitive judgments*, whereas OIC is a philosophical thesis about the concepts of moral obligation (ought) and ability (can). The three experimental studies on OIC discussed in this paper, as well as our own study, reveal that people judge that agents ought to do something even when they also judge that those agents cannot do it. From this empirical fact about intuitive moral judgments, Buckwalter and Turri (2015) want to draw the conclusion that OIC is false. But this conclusion follows from that empirical fact only on the assumptions that intuitive judgments are unbiased and that they are reliable evidence for or against philosophical principles, such as OIC. Similarly, from the empirical fact that people judge that agents ought to do something even when they also judge that those agents cannot do it, Chituc et al (2016) want to draw the conclusion that OIC is not analytic. But, again, this conclusion follows from that empirical fact only on the assumptions that intuitive judgments are unbiased and that they are reliable evidence for discovering the semantic properties of philosophical principles, such as OIC. As experimental philosophers have pointed out, however, "empirical results [such as those presented in this paper and the aforementioned experimental studies on OIC] cast doubt on the use of intuition [e.g., intuitive judgments about

what agents ought to do] as evidence for [or against] moral claims" (Tobia et al.. 2013, p. 630).

As experimental philosophers, Buckwalter, Turri, and Chituc et al. ought to know that the

evidential status of intuitive moral judgments in response to hypothetical scenarios is

questionable, in light of results from experimental studies. As philosophers, Buckwalter, Turri,

and Chituc et al. ought to know that there is no such thing as an "empirical refutation," given

confirmation holism. Even though it has not been "refuted," now that there is empirical evidence

that OIC is not intuitive, OIC needs to be argued for without appealing to its alleged intuitiveness

(Mizrahi 2015b, p. 254).

3. Conclusion

In this paper, our aim has been to contribute to the current debate about the status of OIC and the

growing body of empirical evidence that undermines it. In accordance with previous

experimental studies, our results show that people do judge that agents are morally obligated to

perform an action even when they also judge that those agents cannot do it. Our results also show

that people's "ought" judgments (i.e., judgments about what moral agents ought to do) exhibit an

actor-observer effect. This result should not surprise experimental philosophers and moral

psychologists. Recent studies in moral psychology and experimental philosophy have found that

moral judgments are subject to all sorts of biases and are influenced by factors that are supposed

to be irrelevant from a moral point of view. Because of this actor-observer effect on people's

"ought" judgments, and the Duhem-Quine thesis (or confirmation holism), talk of disproving

OIC by "powerful counterexamples," or of an "empirical refutation" of OIC, is empirically

unwarranted and methodologically problematic. Therefore, the correct conclusion to draw from

the empirical fact that people attribute moral obligations to agents even when those agents are

unable to fulfill those moral obligations is that OIC is not intuitive, not that OIC has been

refuted.

Now that there is empirical evidence suggesting that OIC is not intuitive, those who are

not willing to give up on OIC can no longer treat it as an axiom, i.e., as a principle that can be

taken for granted as fundamental without argument. In fact, if deontic logic, like logic in general,

is supposed to be intuitive, but OIC is not intuitive, as the aforementioned empirical evidence

suggests, then there is a good reason to give up on OIC.[13]

Acknowledgments

References

Andow, J. (2016a). Intuitions. *Analysis* 76 (2): 232-246.

Andow, J. (2016b). Reliable but not home free? What framing effects mean for moral intuitions.

Philosophical Psychology 29 (6): 904-911.

Brogaard, B. (2014). Intuitions as intellectual seemings. *Analytic Philosophy* 55 (4): 382-393.

Buckwalter, W. & Turri, J. (2015). Inability and obligation in moral judgment. *PLoS ONE* 10

(8), 1-20.

---

[13] On the OIC problem in deontic logic, see Carmo and Jones (2002).

Carmo, J., and Jones, A. (2002). Deontic logic and contrary-to-duties. In D. M. Gabbay and F. Guenthner (eds.), *Handbook of Philosophical Logic, 2nd Edition, Volume 8* (pp. 265-344). Dordrecht: Springer.

Chituc, V., Henne, P., Sinnott-Armstrong, W., and De Brigard, F. (2016). Blame, not ability, impacts moral "ought" judgments for impossible actions: toward an empirical refutation of "ought" implies "can". *Cognition* 150: 20-25.

Chituc, V. and Henne, P. (2016). The Data against Kant. *The New York Times*. 19 February 2016. Available: http://www.nytimes.com/2016/02/21/opinion/sunday/the-data-against-kant.html?_r=0 . Accessed 20 February 2016.

Devitt, M. (2015). Relying on intuitions: where Cappelen and Deutsch go wrong. *Inquiry* 58 (7-8): 669-699.

Graham, P. A. (2011). 'Ought' and ability. *Philosophical Review* 120 (3): 337-382.

Hare, R. M. (1951). Symposium: Freedom of the will. *Aristotelian Society Supplementary Volume* 25: 201-216.

Henne, P., Chituc, V., De Brigard, F. & Sinnott-Armstrong, W. (2016). An empirical refutation of 'ought' implies 'can'. *Analysis* 76 (3): 283-290.

Howard-Snyder, F. (2013), Ought Implies Can. In Hugh LaFollette (ed.), *The International Encyclopedia of Ethics*. Wiley-Blackwell.

King, A. (2014). Actions that we ought, but can't. *Ratio* 27 (3): 316-327.

King, A. (forthcoming). 'Ought Implies Can': Not so pragmatic after all. *Philosophy and Phenomenological Research*. Available at https://philpapers.org/rec/ALEOIC.

Kitcher, P. (1982). *Abusing Science: The Case Against Creationism*. Cambridge, MA: The MIT Press.

Knapp, T. (1990). Treating ordinal scales as interval scales: an attempt to resolve the controversy. *Nursing Research* 39 (2): 121-123.

Kurthy, M. & Lawford-Smith, H. (2015). A brief note on the ambiguity of 'ought'. Reply to Moti Mizrahi's 'Ought, can and presupposition: an experimental study'. *Methode: Analytic Perspectives* 4 (6): 244-249.

Littlejohn, C. (2012). Does 'ought' still imply 'can'? *Philosophia* 40 (4): 821-828.

MacCallum, R. C., Zhang, S., Preacher, K. J., and Rucker, D. (2002). On the Practice of Dichotomization of Quantitative Variables. *Psychological Methods* 7 (1): 19-40.

Martin, W. (2009). Ought but cannot. *Proceedings of the Aristotelian Society* 109 (1pt2): 103-128.

Mizrahi, M. (2009). 'Ought' does not imply 'can'. *Philosophical Frontiers* 4 (1): 19-35.

Mizrahi, M. (2012). Does 'ought' imply 'can' from an epistemic point of view? *Philosophia* 40 (4): 829-840.

Mizrahi, M. (2015a). Ought, can, and presupposition: an experimental study. *Methode: Analytic Perspectives* 4 (6): 232-243.

Mizrahi, M. (2015b). Ought, can, and presupposition: a reply to Kurthy and Lawford-Smith. *Methode: Analytic Perspectives* 4 (6): 250-256.

Mizrahi, M. (2015c). Three arguments against the expertise defense. *Metaphilosophy* 46 (1): 52-64.

Mizrahi, M. (2015d). Don't believe the hype: why should philosophical theories yield to intuitions? *Teorema* 34 (3): 141-158.

Nadelhoffer, T. and Feltz, A. (2008). The actor-observer bias and moral intuitions: adding fuel to Sinnott-Armstrong's fire. *Neuroethics* 1 (2): 133-144.

Okasha, S. (2002). Underdetermination, holism, and the theory/data distinction. *The Philosophical Quarterly* 52 (208): 303-319.

Quine, W. V. (1951). Two dogmas of empiricism. *The Philosophical Review* 60 (1): 20-43.

Saka, P. (2000). Ought does not imply can. *American Philosophical Quarterly* 37 (2): 93-105.

Schwitzgebel, E. & Cushman, F. (2012). Expertise in moral reasoning? Order effects on moral judgment in professional philosophers and non-philosophers. *Mind and Language* 27 (2): 135-153.

Schwitzgebel, E. & Cushman, F. (2015). Philosophers' biased judgments persist despite training, expertise and reflection. *Cognition* 141: 127-137.

Sinnott-Armstrong, W. (1984). 'Ought' conversationally implies 'can'. *The Philosophical Review* 93: 249-261.

Sinnott-Armstrong, W. (2006). Moral intuitionism meets empirical psychology. In T. Horgan and M. Timmons (eds.), *Metaethics after Moore* (pp. 339-366). Oxford: Clarendon Press.

Stich, S. & Tobia, K. P. (2016). Experimental philosophy and the philosophical tradition. In W. Buckwalter and J. Sytsma (eds.), *A Companion to Experimental Philosophy* (pp. 5-21). Malden, MA: Wiley Blackwell.

Tessman, L. (2015). *Moral Failure: On the Impossible Demands of Morality*. New York: Oxford University Press.

Tobia, K., Buckwalter, W., & Stich, S. (2013). Moral intuitions: are philosophers experts? *Philosophical Psychology* 26 (5): 629-638.

Wobbrock, J.O., Findlater, L., Gergle, D. and Higgins, J.J. (2011). The Aligned Rank Transform for nonparametric factorial analyses using only ANOVA procedures. *Proceedings of the ACM Conference on Human Factors in Computing Systems* (pp. 143-146).