

teorema

Vol. XXXVI/3, 2017, pp. 00-00

ISSN: 0210-1602

[BIBLID 0210-1602 (2017) 36:3; pp. 00-00]

Experimental Philosophy, Folk Metaethics and Qualitative Methods

David Moss

RESUMEN

Este artículo describe un nuevo método para investigar empíricamente cuestiones filosóficas, que denomino ‘filosofía cualitativa empírica’, el cual incluye el uso de entrevistas semiestructuradas y análisis cualitativos. Se trata de una extensión y una crítica de la ‘Filosofía experimental’ y su uso de la metodología de encuestas. Aquí me concentro en la aplicación de este método al estudio del relativismo metaético popular. Argumento que las encuestas se enfrentan a serios desafíos para asegurar la comprensión de los sujetos o capturar los detalles de sus puntos de vista, y son simplemente incapaces de detectar algunas posiciones metaéticas plausibles, mientras que los métodos cualitativos tienen varias ventajas.

PALABRAS CLAVE: *metaética, filosofía experimental, metodología filosófica, investigación cualitativa, objetivismo, relativismo, semántica.*

ABSTRACT

This paper describes a novel method of empirically investigating philosophical questions, which I term ‘Qualitative Empirical Philosophy,’ involving the use of semi-structured interviews. This offers an extension and a critique of ‘Experimental Philosophy’ and its use of survey methodology. Here I focus on the application of these methods to the study of folk metaethical relativism. I argue that surveys face severe challenges in ensuring subject comprehension, capturing the details of their views, and are simply incapable of detecting some plausible metaethical positions, whereas qualitative methods have several advantages.

KEY WORDS: *Meta-Ethics, Experimental Philosophy, Philosophical Methodology, Qualitative Research, Objectivism, Relativism, Semantics.*

I. INTRODUCTION

Experimental Philosophy makes use of empirical data to inform philosophical debate [Stich and Tobia (2016)]. As the name suggests,

however, it has more specifically been identified with and inspired by “experimental psychology” [Nadelhoffer and Nahmias (2007), p. 123]. Despite this, the dominant method within experimental philosophy remains ‘the survey method’ [Kauppinen (2007), Kirchin, (unpublished), Ludwig (2016)]. For the purposes of this article I shall be focusing solely on this method, which typically consists in giving questionnaires to participants (often university students), asking fixed questions with closed answers (usually multiple choice, such as “yes” or “no,” or a likert scale indicating agreement or disagreement with a statement), followed by quantitative analysis, and is the method used in all the studies which I consider here.

I will also be focusing exclusively on experimental philosophical work in meta-ethics investigating folk moral objectivism or relativism. It is widely acknowledged that there are multiple programs within experimental philosophy, each with different aims and approaches. For example, “positive accounts” which “follow, at least broadly in the tradition of conceptual analysis”, “negative accounts” which are “supposed to provide evidence against the basic methods of conceptual analysis”, and “doing cognitive science” which is characterised as “revealing... effects and then offering explanations ... in terms of certain underlying cognitive processes” [Knobe (2016), §3.1]. With these distinct, even conflicting, aims in play, offering critique of the methods of experimental philosophy wholesale is likely to misfire. Methods which are very ill-suited to one goal may be quite appropriate to another. Thus it makes more sense to focus on one particular debate.

In the next two sections I will describe some of the reasons why the kinds of surveys employed in this area of research might be expected to be inadequate, and inferior to qualitative methods of investigation and analysis. In the subsequent section I shall outline the qualitative alternative which I propose, before defending against some possible objections.

II. CHALLENGES TO SURVEYS: THE PROBLEM OF COMPREHENSION

Perhaps the largest challenge facing the use of quantitative ‘experimental psychology’ style surveys to investigate meta-ethical questions is that of *comprehension*. Such surveys rely on presenting philosophically untrained respondents with short, fixed questions, the responses to which are supposed to indicate their (implicit) stance on meta-ethical questions. Yet it remains unclear whether respondents actually understand the ques-

tions (and responses which they select) in an appropriate way i.e. in line with the intentions and interpretations of the researchers. If they do not, then this is utterly devastating for the studies which rely on these responses. If subjects do not mean what we think they mean by their responses, then we simply cannot take them to indicate answers to the questions being investigated.

Given this, it might seem misplaced to focus so much attention on using (and choosing our methods so that we can use) quantitative statistical analysis. Quantitative analysis of responses allows us to discern with confidence that, say, precisely 74% of people gave this response, and correlations between this and that response were precisely this much, and this was statistically significant to a certain p-value. But this may be all but meaningless if we do not know what the responses indicate. Indeed, these results may be *worse* than useless, if they tell us about patterns in responses, but we are misled about what those responses actually mean. The choice to use simple surveys, with fixed responses, permits quantitative analysis, but precludes confidence that individuals' responses actually indicate their views about the questions we are interested in.

This establishes that respondents' comprehension is important, but should we actually be concerned that individuals *don't* comprehend the questions they are asked? In the case of research on folk objectivism and relativism, I argue, it is clear that we should. Many worries have been raised about the validity of the questions (and fixed responses) utilised. Beebe notes substantial ambiguity in the responses Goodwin and Darley offer as multiple choice options in their (2008) survey, including "True statement", "An opinion or attitude," "The other person is surely mistaken" [Beebe, (2015), p. 13]. These responses are supposed to indicate objectivism or relativism, but, in fact, are clearly compatible with alternative interpretations. For example, a subjectivist could assent to saying "the other person is surely mistaken" as a first-order moral judgement (according to *their* standards, the person *is* surely mistaken) but not as a meta-ethical statement (according to *the other person's* standards *they* are surely right).¹

Significant evidence that respondents *are* interpreting the question in deviant ways is offered by the fact that Goodwin and Darley found that "objectivist" judgements strongly correlated ($r=.84$) with perceived consensus about a statement [Goodwin and Darley (2010), p. 173]. Worse, Goodwin and Darley, found that by manipulating perceived consensus, they could increase and decrease judgements of objectivism [Goodwin and Darley (2012), p. 254]. This is hard to make sense of if respondents under-

stood the question in the way intended by the researchers, but would fit very well with the possibility that respondents are understanding the question *epistemically* or are influenced by *pragmatic* considerations (for example, if respondents assert that no mistake has been made, in order to convey that individuals can reasonably disagree, in cases where there is widespread disagreement). Similar worries are raised by Beebe and Sackris' counter-intuitive finding that some *factual* statements (e.g. about the numbers of stars in the universe or whether Julius Caesar drank wine before his 21st birthday) are rated as half way between objectivistic and relativistic, and others (e.g. whether exercise usually helps one lose weight) are rated as highly relativistic. [Beebe and Sackris (2016), p. 3] If, as seems plausible, respondents do not understand questions about the status of *factual* statements, then by the same token it seems possible that they do not understand questions about the status of moral statements.

In general, researchers in this area aim to discern whether respondents understand the questions they are being asked through a very small qualitative component, in the form of small 'open comment' boxes where individuals are asked to explain their answers. Responses are then excluded if the individuals' explanations make clear that they interpreted the question aberrantly, for example if, despite instructions, they report that someone who disagrees with them about what is right or wrong must simply be imagining extenuating circumstances [Goodwin and Darley (2008), p. 1347]. However, this methodology is inadequate for confirming comprehension. Such open comments can sometimes make clear when a respondent *misunderstands* a question, but, as a rule, they cannot make clear that a person *does not* misunderstand a question. Thus they are inveterately inclined to promote type II errors, where researchers fail to exclude cases of real misunderstanding. For example, an explanation which appears clearly objectivistic "One must be mistaken, because there is only one moral truth, stealing is always wrong," could simply be expressing a *normative* judgement, or a relativist judgement which appears objectivistic because it is implicitly relativised to a community of human Western student respondents [see Sarkissian et al. (2011), p. 486, for details].²

To their credit, Wright et al. quote a variety of the explanations given by subjects in their studies, and note that many of the explanations given for "relativist" responses are not clearly relativist or seem in tension with relativism [Wright et al. (2013), p. 14]. However, many of the explanations they cite as being clearly relativist or objectivist do not seem unambiguously so. For example, "it's a moral issue," "human life is sacred," [Wright et al. (2013), p. 16] may not express meta-ethical commitments, they may

merely be *first order, normative* statements, or explanations of why something is (subjectively) wrong according to the subject's opinion. Likewise, "[it's] a matter of opinion, not fact," "it would be up to the person" [Wright et al. (2013), p. 14] need not express meta-ethical relativism, but could be the respondents indicating that the matters are not settled and individuals need to make a judgement call. For example, I might say "It is up to researchers to decide whether using surveys is right or wrong". But this would not indicate that I think there is no objective fact of the matter. It might merely indicate that I think there is no settled answer and so researchers will have to decide for themselves how to go forward.

Without follow-up, to press respondents on what they mean by "mistaken" and what they mean by "truth" and so on for other aspects of their explanation, such open comments offer us little reason to believe that the person understands the questions and intends their responses in the right way.

III. CHALLENGES TO SURVEYS: THE PROBLEM OF COMPLEXITY

A second challenge for quantitative surveys is the potential *complexity* of individuals' views. Experimental surveys rely on using a relatively small number of questions, responses to which are taken to indicate objectivism or relativism. They may ask other questions to investigate possible explanations or mediators of individual responses (e.g. about grounds for judgement or perceived consensus), but typically only responses to one question [or two in the case of Goodwin and Darley's (2008) studies] are taken to indicate respondents' objectivism vis-a-vis a particular statement.

Yet individuals' metaethical stances may be much more complex. On some semantic theories, the meaning of moral terms will depend on a whole host of semantic intuitions or platitudes endorsed by speakers, or as Loeb puts it "the intuitions, patterns of thinking and speaking, semantic commitments, and other internal states (conscious or not) of those who employ [moral language] ... semantic dispositions" [Loeb (2008), p. 798]. If this is so, objectivism, relativism, or other metaethical features of moral language may not be discernible simply by looking at responses to one (or two) questions. Individuals may have a host of (potentially conflicting) dispositions with regard to such issues. For example, they may answer "yes" to "must one of these people be mistaken" (an objectivist response) but "no" to "is there a correct answer about this

question” (a relativist response). Goodwin and Darley, in fact, found more than 10% of responses were ‘intermediately objectivist’ in this way, i.e. respondents putatively avowed relativism in response to one question and objectivism in response to another, regarding the same statement [Goodwin and Darley (2008), p. 1353].³

Notably, existing surveys *do* ask subjects to answer their questions with regards to multiple moral statements. Thus they can detect “meta-ethical pluralism” [Wright et al., (2013)], or in Gill’s terms “variability” [Gill (2009)]. But they cannot [with the possible exception of Goodwin and Darley’s (2008) studies described above], reveal inconsistency in subjects’ responses to the *same* moral statement, as hypothesised by Loeb (2008) or indeterminacy in their meta-ethical stances as hypothesised by Gill (2009).

In principle, surveys could simply ask *more* questions to cover *more* responses, but this approach would still face limitations. One such problem is how to actually make sense of what individuals *mean* if we find that individuals give putatively objectivist responses to some questions and putatively relativist responses to others (when asked about the same statement or fact)- and the more questions we ask, each supposedly revealing relativism or objectivism, the more this is a risk we have to contend with. Quantitative analysis is ill-suited to making sense of such sets of responses. What exactly does it mean if we aggregate agreement or disagreement with these various items into an objectivism ‘score’ from 0-3, as Goodwin and Darley do [Goodwin and Darley (2008), p. 1352]? Simply aggregating *opposing* responses to *different* questions (putatively) about objectivism risks obscuring the meaning of individual’s responses (or ignoring the sign that they didn’t understand at least one of the questions).

Other features of these surveys render responses ill-suited to provide insight as to what exactly individuals think about morality. For example, responses are often fixed, with subjects forced to select from multiple choice options or to indicate the extent of their agreement or disagreement (on a 1-7 numbered scale) with a statement they are presented with by the researcher, such as “Since your classmate and Sam have different judgments about this case, at least one of them must be wrong” [Sarkissian et al. (2011), p. 487]. This presumes that subjects’ positions can be adequately captured within the confines of agreement with these *particular* statements. But individuals’ actual responses to these cases may potentially be quite different from the fixed options offered by philosophically informed researchers. For example, Beebe reported unpublished data showing that a non-cognitivist item (that “neither belief is

true or false”) was the most popular response when offered [Beebe (2015), p. 25]. Yet this response is ruled out by most surveys, as it is not offered as an option, and even the item described in this case may be seen to have a *cognitivist* framing (by using the word “belief”).

Notably, in a survey of professional philosophers, questions about meta-ethics received a substantial number of “Other” responses: 17.3% for cognitivism vs non-cognitivism and 35.3% for moral internalism vs externalism, higher than the number selecting either internalism or externalism (unfortunately there was no question about objectivism and relativism). These responses included: “The question is too unclear to answer”, “Accept an intermediate view”, “Accept both”, “Accept another alternative”, etc. [Bourget and Chalmers (2009)]. Indeed 14.8% of *professional philosophers* indicated that they were “Insufficiently familiar with the issue” with regards to moral internalism versus externalism, which casts further doubt on how far we should expect non-philosophers to be able to answer metaethical questions [ibid.].

Another possibility is that respondents’ views may be confused, inconsistent or inscrutable. Respondents might want to say, for example, that murder is always wrong, but that ultimately there are no right or wrong answers and it’s a matter for the individual. Alternatively, they may have no stances on these questions. It is possible that they disagree (sometimes) with individuals about moral questions and tell them they are “wrong,” but they they have no metaethical commitments about whether they are telling them they are wrong because they are objectivists or relativists speaking to someone within the same community or subjectivists or anti-realists merely trying to rhetorically persuade the other to act in a certain way. If so, then presenting respondents with multiple choice responses or a fixed statement for them to agree or disagree with may be forcing their responses into a shape which does not match their view or it may be prompting them to express views where they in fact have none or have only inchoate commitments.

IV. THE QUALITATIVE INTERVIEW METHOD

The previous sections have outlined some of the weaknesses of the survey method. It is now necessary to turn to the advantages of the use of qualitative methods. It is first, however, necessary to describe the precise method which I am going to defend here, since there are a variety of different forms of qualitative research which could be employed, includ-

ing conversation analysis of ordinary moral discussion, [Potter and Wetherall (1987)] ethnography or participant observation [Bryman (2001)] and so on. Though, to a large extent, these might offer a similar set of advantages (and perhaps disadvantages), relative to surveys, to the method I am proposing, it will focus discussion to confine my remarks solely to one method.

The method I propose is using in-depth, semi-structured interviews with non-philosophers.⁴ This would involve asking a series of questions relevant to the metaethical issue in question and about how the folk use moral discourse in general and to probe their judgements about moral questions.

The interviews used are *semi*-structured because they have some pre-defined questions or pre-defined issues to ask about, but allow unstructured questions to arise in the course of the interview, as well as other deviations from a pre-structured ‘script.’ [May (1997), p. 111]. This is a core feature of the method, because it allows multiple follow-up questions and, unlike with surveys, allows them to be tailored to individual’s responses. The interviewer can ask *specific* questions to check individuals’ understanding or to discern the nuances of their view on a particular matter, rather than asking generic questions. It also means that respondents themselves can respond in real-time, asking for clarification, checking their own understanding of questions, and so on. If we have reason to suspect, as I argued above we do, that individuals may struggle with understanding questions or that we may struggle to know what they mean, then this is a decisive advantage.

This method is therefore to be distinguished from the “structured interviews” proposed briefly by Beebe [Beebe (2015), p. 28]. Structured interviews do not allow deviation from a pre-agreed script, demanding exactly the same questions be asked each time [May (1997), p. 110]. It is difficult to see any particular advantage to using structured interviews (essentially a researcher-administered survey) in this context. Asking exactly the same questions facilitates comparison across respondents and can facilitate quantitative analysis, but it comes at the cost of investigating the precise details of what each individual thinks and ensuring subject comprehension. Likewise, there seems little advantage to employing *fully* “unstructured” interviews, which have no pre-defined questions at all [May (1997), p. 112]. Researchers will doubtless have some specific questions they wish to ask, when investigating folk metaethics, across different subjects. Thus even relatively unstructured interviews, which substantially let the flow of the interview be determined by individual responses will likely, technically, be semi-structured in this context.⁵

It is important to distinguish the method I am proposing from existing suggestions. For example, Kauppinen proposes the use of “dialogue and reflection” which he identifies as largely continuous with old-fashioned philosophical practice and Socratic dialogue [Kauppinen (2007), pp. 109-110]. Kauppinen’s proposal is centred around the need to ensure that individual’s intuitions are “robust,” that is, that they are the judgements which respondents hold after proper reflection (e.g. “that infinitely patient and focused respondents would give at the end of a dialogue with a Super-Socrates, who never misleads but engages in maximally skilful midwifery that consists in bringing about conditions” [Kauppinen (2007), p. 110]. For this reason, he proposes aggressive intervention by philosophers to ensure that individuals know the (putative) theoretical shortcomings of their responses e.g. “If they answer negatively, the philosopher could point out that this seems inconsistent with attributing genuine moral convictions to the psychopath” [Kauppinen (2007), p. 109].

This is not a feature of the interviews I propose. If we wish to discern how individuals *actually*, normally, talk and think about morality (which I take as the goal of the empirical metaethics being discussed), then we will wish to discern how they *currently* think and talk morally, not what they *would* ideally say, if we were to guide them through philosophical dialogue, point out reasons why they may be wrong and then see what metaethical view they would then endorse. It is entirely possible that the folk *are* in error, when they speak morally, or that they are otherwise inconsistent or confused. To discern whether this is the case, we don’t want to tell them why their views are wrong, but to elicit from them descriptions or displays of their moral understanding with as light a touch as we can. Of course, this is somewhat in tension with asking follow-up questions to discern understanding, and clarifying what respondents really believe. Asking follow-up questions or drawing distinctions to discern which of two things an individual believes, risks influencing and contaminating their responses, especially if the folk had never thought about these issues at all, before. Nevertheless, the tension is not insuperable, and it is possible to sensitively ask questions in a way that aims to minimise such influence, rather than engaging in philosophical debate with non-philosophers to ensure that individuals have full reflected on all philosophically relevant issues.

For this reason, I propose the use of relatively open questions in interviews, so far as this is possible, which is another key contrast with traditional surveys. For example, one can ask individuals *themselves* to give

examples of moral issues, moral statements and disagreements, and then ask broadly stated questions like “why do you think that?” or “how would you respond if you encountered some people disagreeing about that?” and so on. An advantage of this approach is that it helps avoid framing questions in leading ways and putting words into individuals’ mouths which they otherwise would not have considered.

Asking very open, under-specified questions, seeking to elicit descriptions of their moral thought and talk from subjects may also avoid ‘framing’ morality in particular ways in a broader sense. For example, present surveys in empirical metaethics involve presenting individuals with “statements” [Goodwin and Darley (2008), p. 1343; Beebe et al. (2015), p. 388], and asking subjects to respond to “disagreements” [Goodwin and Darley (2008), p. 1344], “different judgements” [Sarkissian et al. (2011), p. 487] or “an argument” [Nichols (2004), p. 9]. These framings plausibly exert pressure against an understanding of morality in non-cognitive terms (e.g. as expressing an attitude like “boo!”), which it is entirely possible the folk possess.

There are downsides to this approach. Respondents may not know what you are asking, and may say as much. They may respond by describing something irrelevant etc. It is also harder to ensure that the individual explains the precise things that the researcher is interested in. The responses elicited by entirely open questions may simply be impossible to analyse in terms of the meta-ethical questions at hand. For these reasons, the researcher will doubtless need to include *some* more ‘closed’ and potentially leading questions in the course of their interview. There is a continuum between more or less open and closed questioning in an interview and there is room for reasonable disagreement among researchers about precisely how many of each should be included. The details of this should be refined as more studies are run utilising this method and likely a plurality of methods should be utilised to offer ‘triangulating’ evidence [Bryman (2001), p. 274].

Nevertheless, a core advantage of the use of these interviews and qualitative analysis is that, whatever questions are asked, respondents are able to offer answers in their own words and at some length (interviews have no fixed length but may take up to an hour). This also allows respondent validation, as subjects can report whether they agree with your interpretation of their responses [Bryman (2001), p. 273]. Thus even if one says “do you think one person must be mistaken or not” the respondent can always answer “well, it depends what you mean by ‘mistaken’ ... ” and elaborate in detail on their view or challenge the framing of the question.

This provides substantial advantages over the use of the survey method, in terms of *comprehension* (of both the subjects and the researcher) and providing nuance and detail of the *complexity* of individual's views.

The method of *analysis* and interpretation is, of course, also significant. There are a variety of different methods of qualitative analysis available (e.g. grounded theory, [Glaser and Strauss (1967)] discourse analysis [Potter and Wetherall, (1987)], thematic analysis, [Braun and Clarke, (2006)] and generic qualitative analysis [Percy et al. (2015)]. There is no single 'brand' which uniquely best fits the kinds of research being proposed here and space precludes discussing their relative advantages and disadvantages. Typically, these involves some sequence of reviewing a transcript of each interview, highlighting salient parts, 'coding' responses i.e. adding labels to key parts, or writing 'memos' describing themes or points of interest, often in an iterative process where themes or patterns which arise, cause the researcher to go back and revise earlier codes and analysis [Percy et al. (2015)]. The precise details, of course, depend on the research question being investigated.

The core desiderata of the form of qualitative analysis which I propose here, is that the analysis must aim to make sense of what individuals actually mean when they answer questions about morality. This necessarily requires taking account of what they say in the course of a whole interview and interpreting what they say in the light of their other responses. There is no purely formal *method* to undertake this process. It requires the use of the ordinary capacity for interpretation that we standardly employ when speaking to people (about morality or about anything else). For example, if someone says "I'm certain that's true, but that's just my opinion" we need to try to discern what they mean, in part, by looking at the context in which they assert it, how it would fit in with what they have said before and what they say afterwards. Just as in any conversation (e.g. with our colleagues or students) there is no straightforward set of rules or method which tells us how best to discern what the other person means. This relies in large part on the capacities of the researchers, engaged in conversation, to discern what are plausible things that a person might be thinking and trying to convey.

This element of subjectivity to the interpretation of what people mean when they talk about morality and the possibility for misinterpretation suggests a need for great transparency in the interpretation of individual responses, not only with these interviews, but also with the small qualitative components included in traditional surveys. Such transparency should also allow evaluation of the specific questions asked by the in-

interviewers, so that readers can consider the context of and possible influences on responses (e.g. interviewer effects). Direct quotations (with as much context as space will allow) should be included alongside argumentation from the researchers for interpreting them one way or another. This is likely to be where the action is, in terms of controversy over folk relativism and objectivism, since, as I argued above, analysis of what individual's open comments mean is likely to be highly contentious. Statements which researchers deem clearly relativistic or objectivistic may not be so (see my discussion of Wright et al. (2013) above). There is likely room for substantial disagreement about how responses are to be interpreted meta-ethically.⁶ Researchers should therefore devote significant attention to looking at the concrete instances of moral discourse that are put forward as evidence and debating precisely how they should be interpreted.

The method I have proposed *might* be similar to what Nadelhoffer and Nahmias have in mind when they propose an “*Experimental Dialogue and Reflection Model (EDRM)*” in contrast to Kauppinen’s more *informal* model of “dialogue and reflection” which they dub “IDRM” [Nadelhoffer and Nahmias (2007), p. 131]. It is hard to tell, because they do not specify many of the methodological details, simply saying:

we should set up controlled and systematic experiments in order to find out what people’s reflective intuitions, judgments, and beliefs about a given topic really are... we could then code, compare, and analyze participants’ answers in a rigorous and systematic manner that is less subject to the problems associated with IDRM (ibid).

From context, I imagine that they are proposing *conversations* of the kind Kauppinen proposes, but with systematic sampling (not just speaking to acquaintances and students each philosopher happens to know), formal coding, and common questions across interviews. If so, then my method may very well fit what they have in mind. However, my interviews would lack the pursuit of fully “reflective” intuitions that Kauppinen’s model espouses, as I mentioned above. It is also not clear what they have in mind by “controlled and systematic experiments” and “analyze participants’ answers in a rigorous and systematic manner.” If by this they mean asking all participants the same questions without deviation or analysing responses according to a fixed framework to allow simple comparison across individuals then, as I have argued above, I think the cost

would not be worth the benefit. While this method could gather useful data, it risks occluding the details of what individuals actually mean.

V. OBJECTIONS

It is now time to consider some possible objections to what I have said here and some more general defences of the use of the survey method.

A core defence of the use of surveys and quantitative analysis – and perhaps the one which is implicitly motivating most researchers – against what I have said here is that the folk’s survey responses are ‘close enough’. That is to say, there might be *some* confusion on the part of respondents, the survey responses might fail to capture *some* of the nuances and details of responses, and their patterns of responses might be *somewhat* influenced by irrelevant factors that we have not (yet) controlled for, but they represent a good enough approximation of individuals’ stances. Therefore, the argument might go, we can and should use these responses for quantitative analysis.

My arguments in the foregoing sections are, I hope, a response to this. I have argued that the failures of comprehension on display are huge and that even in the open comment responses selected precisely to show that respondents understand, it is not clear that they understand. At some points even the *questions* posed by the researchers appear to betray metaethical confusion. Likewise, the format of the survey questions doesn’t just obscure some of the details, but it rules out wholesale the possibility that the folk might be certain kinds of non-cognitivist or incoherent or indeterminate metaethically. Similarly, the influence of irrelevant factors on individuals’ judgements is not small and easily controlled for, but rather they seem to be strong and robust (e.g. objectivism correlates with perceived consensus). If this is right, then the responses from experimental surveys cannot be relied upon to tell us what we want to know about folk metaethical commitments.

Nevertheless, there are a series of further replies which experimental philosophers might make in defence of the survey method.

One is that the use of statistical methods can itself rule out confusion on the part of subjects. Thus Nadelhoffer and Nahmias, argue, in response to Kauppinen:

you can show that the probability is extremely low that the relevant results obtained because of the irrelevant factors. A statistically significant result

does not indicate that any individual participant's response was not due to inattention or confusion or a mischievous desire to mess up the experiment. But for many types of studies, it does indicate that it is highly unlikely that most participants were inattentive or confused or mischievous – or if they were, they'd all have to 'mess up' in a relatively cohesive way [Nadelhoffer and Nahmias (2007), p. 135].

Unfortunately, this cannot rule out the kinds of confusion that I have argued may apply in the case of experimental philosophical surveys. Subjects *may* be 'messing up in a relatively coherent way.' For example, if they tend to understand the questions about objectivism as about whether there is reasonable disagreement, then they may *en mass* select certain responses, evincing a clear pattern but not the one we think they are indicating.

Another response is to argue that through the use of *further* quantitative surveys, they can rule out the kinds of confusions or influence of irrelevant factors that I suggest are problems. Sarkissian et al. offer a neat attempt to confirm whether subjects are just responding *epistemically* i.e. making judgements not about truth but about what it is reasonable to believe [Sarkissian et al. (2011), p. 497]. The study shows that individuals gave different responses to questions about whether "at least one must be incorrect" (putatively about objectivism) and questions about whether "at least one must not have good reason to believe" (putatively about epistemic justification). [Sarkissian et al. (2011), p. 499]. Unfortunately, this method fails, due to the same problem of *comprehension* that I described earlier. Just as we worried that respondents don't understand questions about whether "one must be incorrect" correctly, so we might fear that they don't understand "one must not have good reason to believe" in the desired way. This seems in doubt given that the study seemed to show that respondents thought that different moral beliefs can each be "correct", but not that different individuals could have "good reason" to believe these (correct) beliefs [ibid]. Thus the method can show that respondents didn't understand two items *in the same way*, but it cannot show that they didn't misunderstand either or both of the questions.

VI. CONCLUSION

I have argued that the survey method has severe limitations for studying folk metaethics and that the use of qualitative analysis and semi-structured interviews carries substantial advantages.

While there has been some acknowledgement that interviews might be a useful method to implement [Beebe, (2015) p. 28, Pohlhaus (2015) p. 8], the response from the philosophical community seems to have been rather tepid. The use of experimental surveys in empirical metaethics has continued apace, but I am aware of no published work using the methods I propose to investigate this area. If my argument here is right, then this is a mistake. There are substantial doubts about whether respondents understand the questions in these surveys and whether we understand what they mean by their responses. If so there is a pressing need to look in more detail at how individuals actually understand questions about metaethics and what they mean when they answer them. I hope to have made the case that qualitative methods are a potentially fruitful method for investigating these questions.

*School of Media, Communication and Sociology
University of Leicester
Bankfield House, 132 New Walk
Leicester, LE1 7JA UK
E-mail: dm301@leicester.ac.uk*

ACKNOWLEDGEMENTS

I would like to thank the anonymous reviewer for many helpful comments and suggestions. I would also like to thank attendees of the 2013 UK Experimental Philosophy Group Workshop for their feedback on an early presentation of these ideas. I would also like to thank Lance Bush for much discussion of these ideas and his collaboration on our ongoing pilot of these methods. Thanks are due to Andres Montealegre for his help with translation.

NOTES

¹ For further critique of the ‘disagreement paradigm’ used by Goodwin and Darley (2008), Nichols et al. (2004), Sarkissian et al. (2011) and Beebe and Sackris (2016), see Bush (2016).

² Skulmowski et al. (2015) employ a survey comprised of multiple open questions, the answers to which were subjected to content analysis, to investigate folk concepts of intentional action. This is a substantial advance over normal survey methodology, but still has at least two major limitations: i) it lacks the opportunity for researchers to follow-up and dialogue with respondents to confirm their understanding; ii) the analysis focuses on taking qualitative content but coding it quantitatively, thus losing some individual details.

³ Notably even higher percentages (up to 27%) endorsed such mixed stances with regards to statements in the factual, aesthetic and social convention domains [private communication with the author]. This seems to suggest that these responses indicate real complexity in subjects' responses (as opposed to mere statistical noise or performance error), though it is unclear precisely what they mean, whether conflict or confusion in people's commitments or deviant interpretations of the question, of the kind described in section II.

⁴ Samples could be drawn from student populations, as in many experimental philosophy studies [such as Goodwin and Darley (2008) and Sarkissian (2011)] or from some broader population using maximum variation sampling [Patton (1990)].

⁵ This should, perhaps, not be surprising, since in-depth semi-structured interviews are the pre-eminent method employed by cultural sociologists investigating how individuals think about morality e.g. Bellah et al. (1985), Lamont (1992), (2002), Wuthnow (1992), and have also received use by psychologists such as Billig et al. (1988) and Damon and Colby (1992).

⁶ It may even be that, in general, all or most instances of moral discourse could plausibly be analysed in accordance with *both* opposing metaethical theories [Sinnott-Armstrong (2009), p. 253]. If so, analysing folk responses as relativist or objectivist will be no easy matter.

REFERENCES

- BEEBE, J.R., and SACKRIS, D. (2016), 'Moral Objectivism Across the Lifespan', *Philosophical Psychology*, (published online).
- BEEBE, J.R., RUNYA, Q., WYSOCKI, T. and ENDARA, M.A. (2015), 'Moral Objectivism in Cross-Cultural Perspective', *Journal of Cognition and Culture*, vol. 15, pp. 386-401.
- BEEBE, J.R. (2015), 'The Empirical Study of Folk Metaethics', *Etyka*, pp. 11-28.
- (2014), 'How Different Kinds of Disagreement Impact Folk Metaethical Judgments', in Jen Cole Wright and Hagop Sarkissian (eds.), *Advances in Experimental Moral Psychology: Affect, Character, and Commitments*, London, Bloomsbury pp. 167-87.
- BELLAH, R. N., MADSEN, R., SULLIVAN, W. M., SWIDLER, A. and TIPTON, S. M. (1985), *Habits of the Heart: Individualism and Commitment in American Life*, London, University of California Press.
- BILLIG, M., CONDOR, S, EDWARDS, D., GANE, M., MIDDLETON D., and RIDLEY A. (1988), *Ideological Dilemmas: A Social Psychology of Everyday Thinking*, London, Sage.
- BOURGET, D. and CHALMERS, D. (2009), 'Philpapers Survey', <accessed online at: http://philpapers.org/surveys/results.pl?affil=Target+faculty&areas0=0&areas_max=1&grain=fine [30/04/16]>.
- BRYMAN, A. (2001), *Social Research Methods*, Oxford, OUP.

- BUNGE A., and SKULMOWSKI, A. (2014), 'Descriptive and Pragmatic Levels of Empirical Ethics: Utilizing the Situated Character of Moral Concepts, Judgement, and Decision-Making', in Luetge C., Rusch, H., and Uhl, M. (eds), *Experimental Ethics: Toward an Empirical Moral Philosophy*, Basingsstoke, Palgrave Macmillan, (ebook).
- BUSH, L. (2016), 'The Disagreement Paradigm', unpublished manuscript.
- COLBY, A., and DAMON, R. (1992), *Some Do Care: Contemporary Lives of Moral Commitment*, New York, The Free Press.
- GILL, M. B. (2009), 'Indeterminacy and Variability in Meta-ethics', *Philosophical Studies*, vol. 145, pp. 215-234.
- GOODWIN, G. P. and DARLEY, J. M. (2008), 'The Psychology of Meta-Ethics: Exploring Objectivism', *Cognition*, vol. 106, pp. 1339-1366.
- (2010), 'The Perceived Objectivity of Ethical Beliefs: Psychological Findings and Implications for Public Policy', *Review of Philosophy and Psychology*, vol. I, pp 161-188.
- (2012), 'Why are Some Moral Beliefs Perceived to be More Objective Than Others?', *Journal of Experimental Social Psychology*, 48, pp. 250-256.
- KAUPPINEN, A. (2007), 'The Rise and Fall of Experimental Philosophy', *Philosophical Explorations: An International Journal for the Philosophy of Mind and Action*, vol. X(2), pp. 95-118
- (2013), 'Ethics and Empirical Psychology – Critical Remarks to Empirically Informed Ethics', in Christen, M., Van Schaik, C., Fischer, J., Huppenbauer, M., Tanner, C. (eds.), *Empirically Informed Ethics: Morality between Facts and Norms*, Springer, pp. 279-305.
- KIRCHIN S. (unpublished), 'What is it Like to be a Philosopher?', (seen in draft).
- KNOBE, J. (2007), 'Experimental Philosophy and Philosophical Significance', *Philosophical Explorations: An International Journal for the Philosophy of Mind and Action*, vol. 10, pp. 119-121.
- (2016), 'Experimental Philosophy is Cognitive Science', in Sytsma, J. and Buckwalter, W. (eds.), *A Companion to Experimental Philosophy*, Blackwell (ebook).
- KNOBE, J., BUCKWALTER, W., NICHOLS, S., ROBBINS, P., SARKISSIAN, H., and SOMMERS, T. (2012), 'Experimental Philosophy', *Annual Review of Psychology*, vol. 63, pp. 81-99.
- KNOBE, J., and NICHOLS, S. (2008), 'An Experimental Philosophy Manifesto', in J. Knobe, S. Nichols (eds.), *Experimental Philosophy*, Oxford, OUP, pp. 3-14.
- LAMONT, M. (1992), *Money, Morals and Manners*, London, Chicago University Press.
- LAMONT, M. (2000), *The Dignity of Working Men: Morality and the Boundaries of Race, Class and Imagination*, London, Harvard University Press.
- LOEB, D. (2008), 'Moral Incoherentism: How to Pull a Metaphysical Rabbit out of a Semantic Hat', in Sinnott-Armstrong, W. (ed.), *Moral Psychology, Vol.2, The Cognitive Science of Morality: Intuition and Diversity*, Cambridge, MA, MIT Press, pp. 355-386.

- LUDWIG, K. (2016), 'Thought Experiments and Experimental Philosophy', in Brown J. (ed.), *Routledge Companion to Thought Experiments*, Routledge, London, (seen in draft). Retrieved from PhilSci Archive (12468).
- MAY, T. (1997), *Social Research: Issues, Methods and Process, 2nd Edition*, Buckingham, Open University Press.
- NADELHOFFER, T. and NAHMIA, E. (2007), 'The Past and Future of Experimental Philosophy', *Philosophical Explorations: An International Journal for the Philosophy of Mind and Action*, vol. X(2), pp. 123-149.
- PERCY, W. H., KOSTERE, K., and KOSTERE, S. (2015), 'Generic Qualitative Research in Psychology', *The Qualitative Report*, 20(2), pp.76-85. Retrieved from <http://www.nova.edu/ssss/QR/QR20/2/percy5.pdf>.
- PATTON, M. Q. (1990), 'Designing Qualitative Studies', in Patton M. Q., *Qualitative Evaluation and Research Methods*, pp. 169-186, London, Sage <accessed online at: <http://legacy.oise.utoronto.ca/research/field-centres/ross/ctd1014/Patton1990.pdf> [30/09/15]>.
- PLAKIAS, A. (2015), 'Metaethics: Traditional and Empirical Approaches', in Sytsma, J. and Buckwalter, W. (eds.), *A Companion to Experimental Philosophy*, Blackwell (ebook).
- POHLHAUS JR., G. (2015), 'Different Voices, Perfect Storms, and Asking Grandma What She Thinks: Situating Experimental Philosophy in Relation to Feminist Philosophy', *Feminist Philosophy Quarterly*, (1), Article 3, pp. 1-23.
- POTTER J. and WEATHERALL, M., (1987), *Discourse and Social Psychology*, London, Sage.
- SARKISSIAN, H., PARK, J., TIEN, D., WRIGHT, J.C., and KNOBE J., (2011), 'Folk Moral Relativism', *Mind & Language*, vol. 26, pp. 482-505.
- SARKISSIAN, H., (2016), 'Aspects of Folk Morality: Objectivism and Relativism', in Sytsma, J. and Buckwalter, W. (eds.), *A Companion to Experimental Philosophy*, Blackwell (ebook).
- SCHOLL, B. (unpublished), 'Two Kinds of Experimental Philosophy (and their methodological dangers)', (seen in draft) <accessed online at: <http://www.xphi-europe.org/publikacje/SchollMethodologyNotes.pdf> [30/04/16]>.
- SINNOTT-ARMSTRONG, W. (2009), 'Mixed-up Metaethics', *Philosophical Issues*, vol. 19, pp. 235-256.
- STICH S. and TOBIA K. P. (2016), 'Experimental Philosophy and the Philosophical Tradition', in Sytsma, J. and Buckwalter, W. (eds.), *A Companion to Experimental Philosophy*, Blackwell (ebook).
- SKULMOWSKI, A., BUNGE, A., COHEN, B. R., KREILKAMP, B. A. K., and TROXLER, N. (2015), 'Investigating Conceptions of Intentional Action by Analyzing Participant Generated Scenarios', *Frontiers in Psychology*, vol. 6, (online only).
- WAINRYB, C., SHAW, L. A., LANGLEY, M., COTTAM, K. and LEWIS, R. (2004), 'Children's Thinking About Diversity of Belief in the Early School Years: Judgments of Relativism, Tolerance, and Disagreeing Persons', *Child Development*, vol. 75, pp. 687-703.

- WRIGHT, J.C., GRANDJEAN, P. T. and MCWHITE, C. B. (2013), 'The Meta-Ethical Grounding of Our Moral Beliefs: Evidence for Meta-Ethical Pluralism', *Philosophical Psychology*, vol. 26, pp. 1-26.
- WUTHNOW, R., (1991), *Acts of Compassion*, Oxford, Princeton University Press.