

Swarthmore College

## Works

---

Philosophy Faculty Works

Philosophy

---

3-1-2011

### A Puzzle About Responsibility A Problem And Its Contextualist Solution

Peter Baumann

*Swarthmore College*, pbauman1@swarthmore.edu

Follow this and additional works at: <https://works.swarthmore.edu/fac-philosophy>



Part of the [Philosophy Commons](#)

[Let us know how access to these works benefits you](#)

---

#### Recommended Citation

Peter Baumann. (2011). "A Puzzle About Responsibility A Problem And Its Contextualist Solution". *Erkenntnis*. Volume 74, Issue 2. 207-224. DOI: 10.1007/s10670-010-9259-6  
<https://works.swarthmore.edu/fac-philosophy/18>

This work is brought to you for free and open access by . It has been accepted for inclusion in Philosophy Faculty Works by an authorized administrator of Works. For more information, please contact [myworks@swarthmore.edu](mailto:myworks@swarthmore.edu).

## A Puzzle about Responsibility. A Problem and its Contextualist Solution

*Peter Baumann*

*Erkenntnis 74, 2011, 207-224*

### *Abstract*

This paper presents a puzzle about moral responsibility. The problem is based upon the indeterminacy of relevant reference classes as applied to action. After discussing and rejecting a very tempting response I propose moral contextualism instead, that is, the idea that the truth value of judgments of the form "S is morally responsible for x" depends on and varies with the context of the attributor who makes that judgment. Even if this reply should not do all the expected work it is a first step.

There is a serious problem about moral responsibility – so serious that it deserves to be called a “puzzle”: In some cases there seems to be an indeterminacy of moral responsibility. Before I can present the puzzle in detail (II) I will have to set the stage (I). After presenting the puzzle I will discuss a prima facie very attractive response and explain why it does not work (III). I will conclude with my own, contextualist, response to the problem (IV). The main focus, though, will lie on the problem rather than on the details of the proposed solution (it lies, at best, on the first steps towards such a solution).

### *I. Setting the Stage: Judging Cases*

Consider the following case:

(1) A is driving in his car at 20 mph on deserted Z-street. Suddenly a child appears out of the blue and runs onto the street. A cannot avoid hitting and killing the child.

Something terrible has happened. But did A act in a morally reprehensible way? A was not drunk and was, let us assume, driving even more carefully than one should normally expect under these circumstances. Only very rarely does somebody, especially a child, get to deserted Z-street. There is a sense in which A suffers from bad luck: It was only by chance that his behavior led to disastrous consequences. It is a moot point if this kind of bad luck has a moral side, too, but if it has, then one would have to see it as a form of moral bad luck (see for the basic idea: Nagel 1979; Williams 1981).

Now consider the following case:

(2) A is driving in his car at 20 mph in the densely populated X-neighborhood.

Suddenly a child appears out of the blue and runs onto the street. A cannot avoid hitting and killing the child.

Again, something terrible has happened. A was not drunk. However, there are many kindergartens and schools in this neighborhood, including warning signs. There is a high chance in this neighborhood that children suddenly run onto a street. What is true of Z-street is thus not true of the X-neighborhood. A knows all this. We want to say something like the following: He should have driven more carefully. A has made a moral mistake and everyone would criticize him accordingly. A cannot claim to suffer from bad luck,

especially moral bad luck. Case (2) is a case of morally bad behavior whereas (1) isn't and just involves bad luck (perhaps even moral bad luck).

Some short general remarks are necessary before we get closer to our puzzle. - There is a very plausible general principle that I have implicitly used in the evaluation of cases (1) and (2):

(RESP) If an agent is responsible for some consequence of their action, then there is a sufficiently high (conditional) probability that the consequence will happen, given the action.

This is still a pretty rough approximation of the correct principle but for the purposes of this paper we can do with the formulation just given. It is very plausible, as the consideration of cases can show. Suppose that there is an extremely small probability that I electrocute your hamster by turning the light on. If I turn the light on and do, unfortunately, thereby electrocute your hamster, then I am not responsible for the sad outcome (given normal circumstances) and don't deserve blame. Another case: Suppose there is a great chance that treating you to dinner will make you happy. I treat you to dinner which makes you happy. I am thus responsible for that and deserve praise.

However, what about a variation of case (1) where the driver is speeding recklessly? Isn't he responsible then and doesn't this show that (RESP) doesn't formulate a necessary condition of responsibility? I don't think so. Why should we say that the driver is speeding recklessly and driving too fast – if not because we are (implicitly) assuming that there is a

sufficiently high probability of hitting someone, given the speed and the circumstances (see also below on the role of utilities and disutilities)? If we're certain that nobody will ever get to that street, then there is no reason to qualify the speed as „reckless“ with respect to potential pedestrian traffic. But if that is so, then our modified case of the recklessly speeding driver is similar to case (2) rather than to case (1); thus, (RESP) is o.k. with the attribution of responsibility and blame to the speedy, reckless driver and no objection has been made against (RESP) and its statement of a necessary condition of responsibility.<sup>1</sup>

It should be stressed that (RESP) and similar principles only state a necessary condition of responsibility and not a sufficient condition.<sup>2</sup> There are other factors which co-determine our judgments of responsibility: for instance, whether the agent intended the outcome or merely foresaw it but did not intend it. We do, e.g., not ascribe responsibility if the agent did not intend the outcome and was also unable to foresee it.<sup>3</sup> An agent is responsible for some consequence of their action only if they believed or could or should have believed that there is a sufficiently high chance that the consequence might happen. In the following, I will (for the sake of simplicity) disregard cases in which the agent did not or could not or was not expected to have a belief about the relevant probabilities.<sup>4</sup>

---

1 Another question: Don't we attribute responsibility even in cases of (moral) bad luck, that is, also in cases like (1)? Perhaps, but this kind of responsibility is different from the one discussed here. Thanks to a referee for pressing me both on this point and the one above.

2 So, the following reverse principle is false:  
 (Reverse RESP) If there is a sufficiently high (conditional) probability that a consequence will happen, given an action, then the agent is responsible for the consequence of their action.  
 On why (Reverse RESP) is not true, see also the following remarks.

3 The intention is, of course, crucial for the description of the act: Saying that Mafioso M was just doing his job underdescribes M's intentional killing of another person.

4 Perhaps it is knowledge which is required or justified belief or something else; similar things will hold then, *mutatis mutandis*. For more on this, see below.

Another factor is the utility or disutility of the outcome: The more (less) is to gain from an action, the higher (lower) the acceptable risk, *ceteris paribus* (see also the remarks about stakes and thresholds in the next paragraph). I should also stress that I am not claiming here that all responsibility is responsibility for outcomes. If it should turn out that one can also be responsible for non-outcomes, then one simply has to take what I am dealing with here as one (though very important) kind of responsibility. Anyway, our principle (RESP) explains the difference between our judgments of cases (1) and (2): Only in (2) but not in (1) was there a sufficiently high chance of a bad outcome.<sup>5</sup>

Where the (admittedly fuzzy) borderline between chances that are and chances that are not sufficiently high lies, may depend on and vary with the circumstances. Both the circumstances of the subject under discussion and the circumstances of the attributor who is making a judgment are relevant here. Imagine a case where someone crosses a street when the risk of getting hit by a car is .1. Suppose further that he is crossing the road "just

---

5 Here is an additional, graded principle (leaving aside, again, further precisifications like, e.g., *ceteris paribus* clauses):

(RESP\*) The higher the (conditional) probability that a certain consequence will happen, given a certain action, the greater the responsibility of an agent for the consequence of their action (should it happen).

A referee presented me with the following objection. Suppose A has a gun with one bullet in its six chambers, knows this but doesn't know which chamber is "loaded", aims at a person, pulls the trigger and shoots and kills that person. Suppose further that B has a gun with five bullets in it, knows this but doesn't know which chamber is empty, aims at someone, pulls the trigger and shoots and kills that person. According to (RESP\*), B is more responsible for murder than A. But aren't they both responsible, and to the same degree? No. To see more clearly why not consider the following (parallel) cases. Mary knows that there is a 1-in-a-Million chance that her porch will collapse if she has her neighbors over to sit on it with her; she invites them over, they come over, sit with her on the porch and, alas, the porch collapses. Chuck knows that there is a 50% chance that his porch will collapse if he invites his neighbors over to sit on it with him; he invites them over, they come over, sit with him on the porch and – no big surprise – the porch collapses. It seems simply false to say that both Mary and Chuck are responsible for what happened on their porch and that they are responsible to the same degree. Since the shooting cases above are structurally identical, the same same things hold in those cases. (RESP\*) makes the correct prediction. - For the sake of simplicity, I will stick with the first, non-graded version of the principle. Nothing essential hinges on this.

for kicks". If he is hit by a car, we would hold him at least partly responsible for that. However, if he is crossing the road in order to save a drowning child, then we would not hold him responsible for being hit by the car and rather think of bad luck. The threshold which determines whether the probability is sufficiently high or not is at least partly determined by what is at stake in the situation (according to the subject or the attributor). The greater (lower) the utility of the intended or foreseeable outcome of the action, the higher (lower) the threshold value is and the harder (easier) it is to blame the agent (*ceteris paribus*).<sup>6</sup>

More interesting than the subject's circumstances are, perhaps, the attributor's circumstances here: How strict are the standards of responsibility of those who make judgments about the cases? How careful ought one to be in such situations, according to different attributors? The answer to these questions will (partly) determine whether and how much responsibility will have to be attributed to an agent. It is this latter point – the dependence on the attributor's circumstances – which contextualists like to emphasize (more on contextualism below). I am not going to elaborate on the (quite controversial) attributor-relativity of standards of responsibility in general or construct a contextualist argument that parallels common contextualist arguments about knowledge. I rather want to look at a different aspect which I will motivate through our problem. This problem gives us very good reasons to take contextualism seriously in ethics.

---

6 Thanks to Tony Milligan who pressed me on this point.

## *II. The Puzzle*

Back to our two cases (1) and (2). It seems relatively unproblematic how to judge them in isolation. But what if (1) = (2)?<sup>7</sup> What if our two descriptions refer to the same case? The deserted Z-street might well be located in the populated X-neighbourhood. In that case we have a problem if not a puzzle: How could our agent A both deserve and not deserve blame? It is hard to see how such a contradiction could be true. The root of this puzzle is the so-called problem of the relevant reference class. What is that?<sup>8</sup>

Consider this illustration. Max is a carpenter. He learns that, according to reliable statistics for carpenters, the life expectancy for carpenters is 71 years. Does he thus have a good reason to assume that he has got good chances to see his 71st birthday? Well, it seems so because Max belongs to the class of carpenters – that class is a relevant reference class here. But suppose Max also smokes cigarettes and the statistics for smokers says that their life expectancy is 68 years. What should Max make of these two statistics? What life expectancy does he have? 71 years or 68 years or some other number of years? Suppose further that Max lives in a big city, likes to ride horses and is also a bit overweight (and so on). He also belongs to these reference classes; the life expectancy for these classes might also be different.<sup>9</sup> If we don't have a unique relevant reference class, then we might get

---

7 For purely expository reasons, I did not include this identity in the initial description of (1) and (2) above. However, nothing substantial depends on this “surprise” strategy of mine. The problem does not go away even given that the driver is aware of the fact that (1) = (2). He (like us) still has to choose between thinking about the event in terms of Z-street or in terms of the X-neighborhood. This should become even clearer below. Hence, it is also not necessary to “update” the probabilities for (1) or (2) upon mentioning that (1) = (2).

8 For a short mention of the problem see my 2008.

9 To avoid misunderstandings: This example is about probabilities to live to a certain age and as such not confined to subjective expectations to get to that age.



more than one answer to the question to what age Max can expect to live; however, more than one answer just is too much. So, the question is: What is the one and only one relevant reference class for Max?<sup>10</sup>

We find this problem in our case, too. Given one reference class, namely the events on Z-street, the chances of meeting children on the street were pretty small:

$P(\text{A hits and kills a child} / \text{A is driving at 20 mph, A is on Z-Street}) = \text{some small value.}$

Or, more generally:

$P(\text{consequence} / \text{act, circumstances}) = a$  (where  $a$  is small).<sup>11</sup>

Hence, with respect to that reference class we should say that A suffered from bad luck, perhaps even moral bad luck (case (1)). Given another reference class, namely the events in the X-neighbourhood, the chances of meeting children on the street were quite large:

$P(\text{A hits and kills a child} / \text{A is driving at 20 mph, A is in the X-neighbourhood}) = \text{some high value.}$

---

10 The reference class problem arises for all the different interpretations of "probability". See Hájek 2007.

11 The question what belongs to the act and what to the circumstances need not worry us here. No matter how we "slice the cake" and no matter whether our problem is to identify the relevant description of the circumstances or the relevant description of the act, we have to deal with the same basic problem of the relevant reference class.

Hence, with respect to that reference class we should blame A morally for not behaving appropriately. Now, which one is the correct judgment? Did he make a mistake or was he just unlucky? It seems obvious that at most one answer can be correct (and how could more than one be adequate?). The correct answer would be based on the choice of the correct reference class. But which one is the relevant reference class?<sup>12</sup>

As we will see, there is no promising (non-skeptical, “non-revisionist”, that is, non-contextualist) solution to the problem of the relevant reference class, that is, the problem of determining the one relevant reference class (see Fetzer 1977, Hájek 2007).<sup>13</sup> This leaves us with two options. Either there is no relevant reference class or there is one but so far at least (if not in principle) we don't know that there is one or which one it is. The first case suggests what one can call "nihilism" about relevant reference classes. In our driver's case we would have to say that there is no reason to choose Z-street rather than the X-Neighbourhood as the relevant reference class. If that is correct, then there is also no such thing as the one and only one probability that children will run onto the street. There is only one probability that children will run onto the street, given that we're on Z-street, and another probability that children will run onto the street, given that we're in the X-neighbourhood. If that is so, then the answer to the question is wide open and it is indeterminate whether A behaved in a responsible or in an irresponsible way. And that certainly creates a problem because the assumption that there is a determinate matter of

---

12 see also Greco 2006 for a nice series of cases (involving drunk drivers) which illustrates a closely related problem concerning moral luck.

13 A contextualist response like the one proposed below does not count as a “solution” in the strict sense which would enable us to indicate unique relevant reference classes in a straightforward way, that is, without further relativization or contextualization. It rather constitutes a “sceptical solution” which accepts the initial problem and does not dissolve it. It is in this sense only that the proposal below is meant as a “solution”.

fact which determines whether a given agent behaved responsibly or not seems to be a core part of our ordinary way of thinking and speaking about responsibility. Nihilism is hard to accept and seems very counterintuitive.

The second interpretation of our problem suggests what one can call “scepticism” about relevant reference classes (whether the scepticism is preliminary or principled). The sceptical scenario is not better. We don't know and have no reason to assume that there is such a thing as the one and only one relevant reference class. Given any particular class we don't know and have no reason to assume that it is the relevant one. If that is correct, then we don't know and don't have a reason to believe that there is just one probability that a child will run onto the street. Given one out of many values for that probability we don't know and have no reason to assume that it is the correct value. And this entails that we don't know (in at least many cases) whether an agent (like A in cases (1), (2)) has behaved in a responsible way or not. Again, this certainly creates a problem for everyone who does not want to be a sceptic about responsibility.

### *III. The Solution that Doesn't Work*

One might want to propose that Z-street is relevant here rather than the X-neighbourhood: The accident happened in Z-street and not anywhere else in the X-neighbourhood. The general idea behind this proposed way out of the problem is not new: It is the idea that the relevant reference class is the smallest or narrowest reference class (cf., e.g., Reichenbach 1949, 372-378, esp. 374). This is considered by many adherents of this idea to be

intuitively plausible. And it seems to guarantee that there is one and only one relevant reference class in all (or most) cases and thus always (or usually) a unique answer to questions about responsibility. There are two versions of this idea. According to the first version, the relevant reference class is determined by all the conditions  $D_1 \dots D_n$  met by the event  $E$ .<sup>14</sup> The probability of  $E$  we're looking for would thus be the conditional probability of  $E$ , given  $D_1 \dots D_n$ :

$P(E / D_1 \dots D_n)$ .

Given that and also given that no (or almost no) two events are exactly alike we should expect the relevant reference class to consist of just one member: the one particular situation which meets the conjunction of all the conditions specific to just one event, namely  $E$ . This raises several problems: How could we ever hope to grasp all the conditions met by the event? What if there is an infinite number of such conditions? And how are we going to deal with the fact that we don't seem to have probabilistic information about single cases?<sup>15</sup>

Let us look at Max, the carpenter, again. It would not be sufficient to only consider the reference class of carpenters; he smokes and we thus have to look at the class of smoking carpenters. Since Max also rides horses, lives in a big city and is slightly overweight, we

---

14 Reichenbach restricts the conditions to those for which we do have statistical information concerning their relation to  $E$ . However, if we want to know what really is the relevant reference no matter the limitations of our information, then we have to choose the formulation in the text above.

15 Talk about "the probability of single cases" is ambiguous here. It can either mean the probability that a certain combination of properties which is only instantiated once is  $F$ . Or it can mean the probability that a certain individual  $a$  is  $F$ . Fortunately, this difference does not matter here.

also have to look at the class of smoking, horse-riding and slightly overweight big city-carpenters. Suppose further that Max lives in Glasgow where it rains a lot. Given a special genetic precondition, he suffers more than others from the huge amount of rainfall in Glasgow. It is not difficult to see that we're ending up with a combination of properties which are so specific that it is instantiated only once (Max). Why is it a problem if Max's reference class has only one member, namely himself? First, the statistical information we have always concern a plurality of cases, not single cases. One might object that this is only an epistemic limitation and there really is – perhaps inaccessible to us – a true single case probability. Now, talk about single case probabilities is also very questionable, if not highly problematic. It is not clear at all what it is supposed to mean that there is a probability  $x$  that one individual at a time (Max, now) has a particular property (will become 71). If the most specific reference class only has one member, then one would have to defend the idea of single-case probabilities (in both of its versions mentioned in fn.15) which is highly contested, to say the least (cf., e.g., Gillies 2000). Finally, why should we take all conditions met by E into account? Some of these conditions obviously have no relevance for E. It is much more plausible to only consider those which are probabilistically relevant for E.

This leads to the second version of the idea of the smallest or narrowest reference class which is more plausible (see Reichenbach 1949, 374, Hempel 1965, 53-79, 397-403, Salmon 1966, 90-92, and also Beebe 2004, 181). According to the second version, the relevant reference class is determined by all the probabilistically relevant (for E) conditions  $C_1 \dots C_n$  met by E. A condition  $C_x$  is probabilistically relevant for E just in case

$$P(E / Cx) \neq P(E).$$

The probability of E which we're looking for would thus be the conditional probability of E, given the probabilistically relevant conditions C, ... Cn:

$$P(E / C_i \dots C_n).$$

In our case, the relevant conditions would include spatial (E happened in Z-street, in the X-Neighbourhood, etc.), temporal (in April, at 5pm, after school) and other kinds of conditions (e.g., weather conditions, health conditions and constitution of driver and child, etc.). Let us for a moment focus on the spatial parameter (just for the sake of simplicity). The narrowest or smallest reference class would be determined by the conjunction of all the probabilistically relevant spatial conditions. Since both the X-Neighbourhood and Z-street are probabilistically relevant, the relevant spatial reference class would be determined by the conjunction "It happened in the X-Neighbourhood and it happened in Z-street" which is equivalent - given the geography - to "It happened in Z-street". In other words, according to this proposal the first case description (1) is the relevant one:

$P(A \text{ hits and kills a child} / A \text{ is driving at 20 mph, } A \text{ is on Z-Street}) = \text{the probability of } E$   
we're looking for = a small value.

How convincing is this view?

First of all, I have made the assumption here that there are only two spatial reference classes. But what if there are many or even infinitely many? Also, problems concerning single case probabilities similar to those mentioned above arise, again (see the remarks on Max's case above).

However, even if single case probabilities were not to pose a problem, there would still be even more serious difficulties. First of all: Why should one assume that there is a unique reference class (of probabilistically relevant properties) such that every other such reference class would be broader? Nothing seems to motivate this assumption. There could be cases such that for every reference class (of probabilistically relevant properties) there is a narrower one.

Furthermore and more importantly, why should one accept the idea of the smallest relevant reference class in the first place? Why should we be inclined to say that in our case Z-street rather than the X-neighbourhood gives us the relevant reference class? Well, because the accident happened on Z-street and not anywhere else in the X-neighbourhood, one might respond. This is true but we need to be careful here. Z-street is the place where the accident happened but what does this imply about the relevant reference class? We are interested in what happened on Z-street at a certain time but that does not mean that the events on Z-street also constitute the relevant reference class. There is an argumentative step in need of justification here:

(1) Z-street is, most precisely, where it happened

(2) Hence, Z-street is the relevant location.

Missing the argumentative support for this step, we should remain sceptical about the conclusion. We need an independent argument that shows that Z-street is the relevant location. Just pointing out that the accident happened on Z-street is not enough. Someone with stricter ideas about cautious driving might point out that in such a neighbourhood one just has to drive very carefully whereas someone with less strict ideas about cautious driving might hold against this that Z-street doesn't require such cautious driving. Is one of the two right? And if yes, does it have anything to do with the fact that the accident happened on Z-street?

Things get even trickier. Even if we agreed on Z-street as the relevant location, the same kind of questions would arise with respect to the time of the event. Should we consider that it was a Sunday during the holidays or rather that it was 5 minutes before the ice-cream place opened? Further kinds of conditions come into the picture, too: What was the age and height of the child? What was the condition of the road or the reaction time of the driver? With respect to all parameters of such kinds the same questions about the relevant reference class will come up (I won't repeat myself here). Finally, we would have to find a non-arbitrary way of weighing all these parameters against each other.

What if Z-street is busy whereas the X-neighborhood is very quiet? Is it not obvious in such a case that A should go with the narrower reference class and drive more cautiously and slowly? Doesn't this show that the narrower reference class is the relevant one? No, at least not in general. To be sure, it is at least *prima facie* plausible that in this "reverse" case



the driver should slow down. There is thus an asymmetry between the reverse cases and our original case insofar as in the latter case it is not so plausible (if at all) that the driver should go with the narrower reference class. What explains this asymmetry? There is nothing in the probability distribution across the spatial parameter that could explain it. It has to be something else. There might be a psychological effect having to do with the salience of “nearby” risks. Perhaps more plausible is a moral explanation: We seem to adhere to a principle of specific obligation (PSO) (or some similar principle) according to which

(PSO) One ought to behave (at least) as cautiously as the most specific description (accessible to the agent) of the relevant circumstances demands.

However, in contrast we do not seem to accept a corresponding principle of specific permission (PSP) (or some similar principle) according to which

(PSP) One may behave according to what the most specific description (accessible to the agent) of the relevant circumstances allows for.

This explains the asymmetry between our judgments about the reverse and the original case. It is important to notice that this difference is based on moral principles we do, as a matter of fact, accept. This still leaves the question open whether a principle like (PSO) is true or well justified and thus can justify one choice of a reference class over all

alternatives. Even if it does, it might help us only in some cases – like the reverse case – but not in all cases. There does not seem to be a plausible moral principle which would determine a unique solution to our original case.

Finally, it will also not help to say that the reference classes the agent is aware of or can think of are the relevant ones. In our driving cases above he knows of both reference classes and even know of the corresponding probabilities. He is not missing out on any relevant fact; there is nothing he should have known but does not know.

What if the driver does not know what he should have known? Can normative considerations about what he should have known help us solve the reference class problem? No and here is why not. What should the agent know: the probabilities for the X-neighbourhood of those for Z-street? The answer to this question depends on what the relevant reference class is: the X-neighbourhood or Z-street. The driver should know the probabilities for the X-neighbourhood if that is the relevant spatial reference class and does not need to know the probabilities for Z-street if that is not the relevant spatial reference class; however, if Z-street is the relevant spatial reference class, then the driver needs to know the probabilities for Z-street but doesn't need to know the ones for the X-neighbourhood. So, normative considerations about what the agent should know depend on what the relevant reference class is and can thus not be used to solve the problem of the relevant reference class.<sup>16</sup>

However, the subject did not lack any potentially relevant knowledge in our case. Cases of non-culpable error or ignorance are different altogether; there might be no responsibility

---

16 Thanks to a referee here.

at all in such cases. Our case, again, is different. Taking all this into account, we can recapitulate that the question about the relevant reference class is still wide open.

#### *IV. Towards a Solution*

It seems that there is no solution to the reference class problem in sight which would tell us what determines relevant reference classes. What does all this imply for our puzzle? I distinguished between scepticism and nihilism above. According to the first, there are relevant reference classes but we (often) do not or even cannot know them and thus do not or cannot even have knowledge or justified beliefs concerning the responsibility of agents. Scepticism clashes with our practice of attributing moral responsibility to agents and is very hard to accept. And: Do we have any reason to assume that there are relevant reference classes when we cannot know that there are any and which ones they are? Let us rather look at nihilism. Insofar as our puzzle suggests nihilism, there is something constructive to say about it.

Here is where we have got so far. At least in some cases, if not in very many cases, the probability of an event (accident) varies with location, time and causal pre-conditions. In these cases, the probability of the event (accident) varies with the reference class we choose. The indeterminacy of the relevant reference class consists in this: There is no matter of fact which determines which reference class is the relevant one. The “facts” therefore leave it open what the probability of a certain event (accident) is. This also implies, given the principle about responsibility mentioned above, that it is

underdetermined by the facts (plus moral considerations) whether the behaviour of a person was responsible and appropriate, irresponsible and morally blameworthy, or whether the person was just suffering from bad luck. I want to argue that at least in some cases it is therefore relative to our description of the situation or on our perspective on things (“A was on Z-street” vs. “A was in the X-neighbourhood”) whether it is true to say that A has made a mistake or true to say that he suffered from bad luck. The truth conditions of our moral judgments about certain actions vary with the descriptions and the perspective of the person who makes the judgment; the latter are not completely determined by the objective situation. General moral judgments (“Given normal circumstances, it is bad/ not permitted to kill other persons if one can avoid it”) in combination with the “objective” facts of a situation do not or do not always determine completely how we have to judge a particular act in moral respects. How we judge it depends on our way of looking at things. One and the same situation might be such that different and mutually incompatible judgments are equally reasonable and appropriate.

All this looks very messy. But perhaps there is a way out of this mess? Here is straightforward advice: Embrace contextualism! Contextualism about moral terms (like “responsibility”, “blameworthiness”, etc.) is the view that the truth conditions of sentences of the form “A is *m*” (with “A” ranging over subjects and “m” over moral terms) vary with the context of the speaker of an utterance of a sentence of this form.<sup>17</sup> This view need not be restricted to linguistic cases; we can extend it to thoughts and beliefs which are not

---

17 I am using the word “truth” in a minimal sense here, that is, in a sense that does not imply any commitment to moral realism. –Contextualism about “knowledge” holds that the truth conditions of sentences of the form “S knows that *p*” vary with the context of the speaker of an utterance of a sentence of that form. For epistemic contextualism about terms like “knowledge” cf., e.g., Cohen 1987, DeRose 1999, Lewis 1996.

expressed linguistically: The contents of thoughts or beliefs of the form “A is m” (with “A” ranging over subjects and “m” over moral terms) vary with the context of the thinker. For the sake of simplicity, we can restrict ourselves to the linguistic case here.

One can motivate moral contextualism in general through the consideration of cases. Here is one. Jack has not returned a pot that he borrowed from his good friend Mary. Ann says “Jack is blameworthy for not returning the pot” while Frank says “Jack is not blameworthy for not returning the pot”. Ann and Frank agree on all the facts of the matter. They just differ in the normative standards concerning the returning of pots between close friends. A non-contextualist – an “invariantist” – would say that Ann and Frank contradict each other and that at most one of them gets the moral standards relevant in this case right. In contrast, a contextualist could say that they’re both right without contradicting each other: One sentence is not the negation of the other; “Frank is blameworthy for not returning the pot” means something different in Ann’s and in Frank’s mouth. Different kinds of moral standards in Ann’s and Frank’s case explain the differences in meaning. If the contextualist is right about our pre-theoretic judgments – our “intuitions” - about straightforward cases like the one above, then she has an advantage over invariantists who are burdened with the task to argue for the thesis that all normative standards are context-invariant.<sup>18</sup>

Moral contextualism can also be motivated in a more theoretical way. Just to pick one example for the purpose of a general illustration without having to commit myself to it here: If *ought* implies *can* and if *can* (that is, the attribution of abilities) is context-sensitive, then there is a good case for holding that *ought* (that is, attributions of

---

18 Thanks to a referee for pressing me on this. For more on error theories, see below.

obligations and duties) is context-sensitive, too. David Lewis has argued for the context-sensitivity of “can”-statements (see Lewis 1998, 166-167): I but no chimp can speak Finnish but I cannot come with you to Helsinki as a translator. The meaning of “can speak Finnish” or of “can” more generally varies with speaker’s context. If what a subject ought to do is determined by what they can do, then we should expect “ought”-statements to show a parallel context-sensitivity of meaning.

Moral contextualism should be clearly distinguished from different forms of invariantism about the content of moral expressions. Apart from the classical invariantism mentioned above I would like to briefly mention “subject-sensitive invariantism” about moral terms.<sup>19</sup> According to this view, the content of attributions of moral properties is not sensitive to the context of the speaker; however, truth values can vary because moral properties (like blameworthiness etc.) are not only determined by the usual moral parameters (moral principles etc.) but also by non-moral aspects of the situation of the subject (not the speaker). A rather uncontroversial example consists in the role mitigating circumstances play in our judgments about responsibility. Many other examples can be found in studies of concrete moral problems (e.g., in applied ethics in particular). Timmons (1993, 1996, 1999), Sinnott-Armstrong (1996) and Thomas (2006a, 2006b) defend this kind of view in interesting and controversial ways for core moral notions.<sup>20</sup>

---

19 For subject-sensitive invariantism about epistemic terms like “knowledge” see Hawthorne 2004 and Stanley 2005. According to this view, knowledge does not only depend on epistemic aspects of the subject’s situation (evidence, belief, etc.) but also on non-epistemic factors (like what is at stake for the subject). – Another form of invariantism not to be discussed here is truth-relativism: the same content can be true for one assessor and not true for another assessor (see MacFarlane 2005).

20 This view is sometimes also called “contextualism” – because of the relevance of further, traditionally neglected factors in the situation of the subject. In the early days of the discussion of epistemic contextualism, people distinguished between “subject-contextualism” and “attributor-contextualism” (see for this DeRose 1992). Only the later view is nowadays called “contextualism”.

In contrast, contextualism is a form of “variantism” about content: What speakers say when they use terms like “responsible” varies with their context of utterance. However, given the meaning of their utterances (and the content of their thoughts), there is no variation of responsibility (or other moral characteristics of the agent) across context. The latter is something subject-sensitive invariantism and relativism are committed to: They locate variation not in the semantics of the relevant terms but in the worldly situations we’re talking about with those terms. Contextualism has an advantage over such non-classical forms of invariantism insofar as it is more plausible – as I think it is - to locate variation in the semantics and not in the facts. Classical invariantism is not committed to any of this but does have problems accounting for the phenomena (see above).

Moral contextualism has, somewhat surprisingly, had much less of an impact on moral philosophy than epistemic contextualism on epistemology. I won’t speculate about the reasons why. I am also not going to defend moral contextualism in general here<sup>21</sup> but rather suggest a form of moral contextualism concerning responsibility and related concepts as a solution to our problem or puzzle. Here is what I have in mind.

The context of an attribution of responsibility involves the choice of one among many perspectives or reference classes. Very roughly: An agent’s responsibility may then vary with the perspective of an attributor. More precisely: The truth conditions of judgments of

---

We should stick with the parallel restriction in the moral case and use “contextualism” in the restricted sense explained above.

21 Dreier 1990, Norcross 2005a, Brogaard 2003, 2008, Greco 2008, Wedgwood 2006b, Montminy 2007 and Jenkins & Nolan 2010 propose contextualism for core moral terms (see also Unger 1995 and critically: Weatherson 2008, sec.4 and Schroeder 2009, 284-287). There is also growing debate on the context-sensitivity of normative ought-claims: see, e.g., Björnsson & Finlay 2010 (esp. the first two sections) and Wedgwood 2006a, 151-152. Hawthorne 2001 and Rieber 2006 have tried to apply attributor contextualism to the concept of a free action while Norcross 2005b defends a contextualist analysis of “harm”. Sinnott-Armstrong 2006 defends contrastivism with respect to justified moral belief but does not hold a contextualist version of contrastivism.

the form “S is morally responsible for x” depend on and vary with the context of the attributor, especially with chosen reference classes (for simplicity’s sake, I skip meta-linguistic talk when nothing hinges on it). This view has the great advantage of showing us a way out of our problem.<sup>22</sup>

This approach explains why we don’t end up with lots of contradictions. It is elliptical to say that both “A does deserve blame” and “A does not deserve blame” are true. If we were to make the contextual factor fully explicit, we would rather say something like the following: Both “Given description D1, A does deserve blame” and “Given description D2, A doesn’t deserve blame” are true.<sup>23</sup> There is no contradiction here. Which reference class we pick is not left to arbitrary decision but rather determined by our practical interests (and probably some other parameters, like salience, too; see below) in the situation.

Consider our initial case again. The lawyers of the parents of the child killed in the accident will probably rather tend to focus on the X-neighbourhood (“This is a very busy neighborhood, after all”) whereas the lawyers of the defendant will probably rather tend towards choosing Z-street as the relevant reference class (“This is, of course, a very deserted street”).<sup>24</sup> What factors determine the choice of a reference class? We can take a lead from the discussion about epistemic contextualism and mention two main factors

---

22 Heller 1995 proposes a contextualist response to the generality problem in epistemology – which can be seen as a special case of an epistemological version of the reference class problem (see also Heller 1999).

23 One could propose a kind of “relationalism” according to which there is a further parameter implicit in responsibility judgments, namely a parameter for reference classes. This would still qualify as a kind of contextualism.

24 For a discussion of legal implications of the reference class problem see Colyvan/ Regan/ Ferson 2001.



here: practical interests on the one hand and salience on the other hand.<sup>25</sup> The precise location of the accident might well be much more salient to the parents or to someone close to the victim than to the lawyers of the defendant (if not the defendant themselves). It is plausible to assume in our example that the parents are much more disposed to imagine the accident in its concrete detail than the lawyer.

The other relevant factor determining the choice of a reference class are practical interests.<sup>26</sup> The parents might have an interest in seeing the defendant convicted while the lawyers will have the opposite interest. Consciously or not so consciously, both parties might tend towards picking those reference classes which will make the driver look bad or not so bad. There might be more factors determining the choice of a reference class as the relevant one. Some people hold certain moral or cultural norms – like (PSO) or (PSP) (see above) – which suggest certain choices of reference classes. And there might be further psychological factors. Whatever the complete list of factors turns out to be, they will vary from speaker's context to speaker's context and thus explain the context-dependency of our judgments of responsibility.

One problem with such a contextualist view about the attribution of responsibility, blame and praise consists in the fact that agents typically have a non-contextualist view of what they are doing when they make these kinds of attributions. They certainly don't seem to think that their attributions are driven by practical interests or by psychological factors

---

25 According to some theories of epistemic contextualism, it is the conversational context and what is being mentioned in a conversation which determines what is salient to the participants (and thus also the contextually relevant epistemic standards; see Lewis 1996). This is less plausible in moral cases like our example. Salience is rather determined by other, psychological, factors.

26 see for epistemological cases Cohen 1987 and DeRose 1999 (for subject-sensitive invariantism: Hawthorne 2004 and Stanley 2005). These authors focus very much on what is at stake for the attributor: The more is at stake, the higher the standards for knowledge.

like salience. So, it seems that contextualism has to systematically attribute semantic blindness and error and ignorance to agents when they make judgments about responsibility. The lawyers in the above cases think they are disagreeing when they're not.<sup>27</sup> Isn't this a huge drawback of the view? I think this is indeed a drawback but it is not huge because every view on the subject matter will have to attribute some blindness to agents. For everyone who is right in cases like our traffic accident case (the lawyers of the defendant or the parents), there will be lots of people who are wrong about it (the parents or the lawyers of the defendant); this is not just another case of rather superficial moral disagreement but rather a case where one party is fundamentally mistaken about the correct way of making judgments about responsibility (going by the narrowest reference class or not). It is hard to see how any at least somewhat plausible view about the attribution of responsibility could avoid including some error theory (see more generally: Mackie 1977). I don't see why the burden for the contextualist should be unbearable, especially in comparison. Furthermore, the lawyers in our case above are still disagreeing about something: whether the defendant is guilty and should be sentenced. They would or could even, after some reflection, come to see that they have different standards of responsibility (and say things to each other like „You, as the defence lawyer, are, of course, more lenient!“ and „We do, of course, understand that you are entertaining very strict views about responsibility, given the terrible loss you suffered!“). They can and might come to see that they are talking past each other in a sense (even if both parties hope that the judge will share their way of looking at the case).

---

27 For this kind of problem in the epistemic case see Schiffer 1996 and, as a response, DeRose 2009, ch.5.

But how can our practice of making attributions of responsibility persist if contextualism is true? How can we explain the robustness of this practice if it is so delicately context-sensitive? Well, how can we explain the persistence of moral judgment in general, given long-standing and deep disagreements both about the nature of morality and about concrete moral problems? If we think we have a good enough explanation for the latter phenomenon, then we should also be confident that we can have one for the former phenomenon. Both are in the same boat.<sup>28</sup>

Isn't there a non-contextualist, "invariantist" way out of all this? Can't we come up with plausible, well-justified or true normative (e.g., moral) principles which tell us in each case what the relevant reference class is and thus enable us to dissolve our puzzle? Aren't there better and worse choices of reference classes? Isn't there better or worse reason to go with one rather than another reference class? It is important to keep in mind that what we would need here is a normative argument which supports one reference class against all other reference classes or, at least, against all other reference classes which suggest a different judgment about responsibility in the case at hand. Perhaps one can make such an argument in the "reverse case" above; perhaps one can argue that a principle like (PSO) is the correct one here and fortunately happens to select exactly one reference class as the relevant one. But even if such an argument should go through in cases like the reverse case (which would have to be shown in detail), there would still be many other cases, like our original

---

28 Thanks to a referee who pressed me on the last two points. - One more issue. As soon as the attributor becomes aware of the indeterminacy of the relevant reference class and of the relativity of their own view to their own interests and perspectives, something similar to Moore-paradoxality and incoherence threatens: "A is suffering from bad luck – but that's just me and my perspective on things!" sounds self-defeating. It seems that the attributor can only suspend judgment under these conditions. See, however, the above remarks on error theories.

one, where no good normative argument is in sight that could select a unique reference class as the relevant one. In short, if there is a normative solution to our problem, then it will at best be a very partial one. The indeterminacy of the reference class and thus the motivation for contextualism remains.

Similarly: Perhaps one can find a plausible account of what counts as legitimate interests which can tell us, at least in some cases, which choices of reference classes are based on legitimate and which on not so legitimate interests. And perhaps there can be an account of salience and other psychological factors which distinguishes between more informed views of the situation and less informed views and, on that basis, determines one unique relevant reference class. All this might go some way. But there does not seem to be a good reason to expect that this general kind of strategy – whether it works with normative principles, an account of legitimate interests or a theory of well-informed salience - will solve the whole problem and not only part of it. Until plausible accounts to that effect are presented, we have good reasons to remain sceptical – and to stick with a contextualist account. All this should not be taken as a denial that there could in principle be good normative arguments of such a kind but rather that there does not seem to be a developed strategy or argument yet which would resolve our contextualist concerns.

Some might suggest that we should just forget about moral contextualism at this point - all the remarks above are still not good enough and leave us with a position we cannot live with. Epistemological contextualism is seen by many as offering promising strategy against epistemological scepticism. In contrast, the moral contextualism proposed here seems to directly invite moral scepticism or nihilism about responsibility (despite all the

explanations above). Doesn't all that lead to relativism and subjectivism ("It all depends on how you look at it"), that is, into the moral philosopher's hell? I don't think so. First of all, I haven't argued and think it is false to say that our context-relativity is always around. It seems to concern only some cases – namely cases where there are two relevant reference classes such that the probability of the outcome, given action and circumstances, varies with the choice of reference classes (as in our cases (1) and (2)). In addition, the above hints as to factors determining choices of reference classes go some way into the direction of mitigating the indeterminacy of the relevant reference class. Contextualists, too, don't want to and don't have the resources to judge that Jack the Ripper is not worse off, morally, than, say, Gandhi. Neither does contextualism imply that moral principles or ideas are obsolete nor that moral judgments cannot be true or false, correct or incorrect. However, there certainly is a certain "relativity" or underdetermination of our moral judgments, according to the kind of contextualism proposed here. This might seem suspect because it seems to go against the idea of a certain "invariance" and "objectivity" of our moral view of the world: The facts are what they are and they don't change with our perspective. Contextualism goes against this intuition. This might be puzzling but it certainly does not constitute a collapse of our moral concepts. Apart from that, contextualism is able to answer our problem which is an important problem, indeed: the reference class problem, as applied to moral responsibility. To be sure: The proposed contextualism does not make the indeterminacy of the relevant reference class go away. But it explains and makes intelligible a problematic aspect of our practice of attributing

responsibility to agents. In this sense, the solution proposed here is a “sceptical solution”. Even if the puzzle has not lost all of its bite, it has lost some of it.

ACKNOWLEDGMENTS: I am grateful to conversations with and comments by a number of people (and I hope I’m not forgetting anyone): Michael Clark, Christoph Fehige, John Greco, Tony Milligan, Bob Plant, Walter Sinnott-Armstrong, Barry Smith, anonymous referees as well as audiences at universities at Aberdeen, Essen, Glasgow and Stirling.

## References

- Baumann, Peter 2008, Problems for Sinnott-Armstrong's Moral Contrastivism, in: *The Philosophical Quarterly* 58, 463-470.
- Beebe, James R. 2004, The Generality Problem, Statistical Relevance and the Tri-Level Hypothesis, in: *Nous* 38, 177-195.
- Björnsson, Gunnar & Finlay, Stephen 2010, Metaethical Contextualism Defended, Ms.
- Brogaard, Berit 2003, Epistemological Contextualism and the Problem of Moral Luck, in: *Pacific Philosophical Quarterly* 84, 351-370.
- Brogaard, Berit 2008, Moral Contextualism and Moral Relativism, in: *The Philosophical Quarterly* 58, 385-409.
- Cohen, Stewart 1987, Knowledge, Context, and Social Standards, in: *Synthese* 73, 3-26.
- Colyvan, Mark/ Regan, Helen M./ Ferson, Scott 2001, Is it a Crime to Belong to a Reference Class?, in: *The Journal of Political Philosophy* 9, 168-181.
- DeRose, Keith 1992, Contextualism and Knowledge Attributions, in: *Philosophy and Phenomenological Research* 52, 913-929.
- DeRose, Keith 1999, Contextualism: An Explanation and Defense, in: John Greco/ Ernest Sosa (eds.), *The Blackwell Guide to Epistemology*, Oxford: Blackwell, 187-205.
- DeRose, Keith 2009, *The Case for Contextualism. Knowledge, Skepticism, and Context*, vol.I, Oxford: Clarendon.
- Dreier, James 1990, Internalism and Speaker Relativism, in: *Ethics* 101, 6-26.

- Fetzer, James H. 1977, Reichenbach, Reference Classes, and Single Case Probabilities, in: *Synthese* 34, 185-217.
- Gillies, Donald 2000, *Philosophical Theories of Probability*, London: Routledge.
- Greco, John 2006, Virtue, Luck and the Pyrrhonian Problematic, in: *Philosophical Studies* 130, 9-34.
- Greco, John 2008, What's Wrong with Contextualism, in: *The Philosophical Quarterly* 58, 416-436.
- Hájek, Alan 2007, The Reference Class Problem is Your Problem too, in: *Synthese* 156, 563-585.
- Hawthorne, John 2001, Freedom in Context, in: *Philosophical Studies* 104, 63-79.
- Hawthorne, John 2004, *Knowledge and Lotteries*, Oxford: Clarendon.
- Heller, Mark 1995, The Simple Solution to the Problem of Generality, in: *Noûs* 29, 501-515.
- Heller, Mark 1999, The Proper Role for Contextualism in an Anti-Luck Epistemology, in: *Philosophical Perspectives* 13, 115-130.
- Hempel, Carl G. 1965, *Aspects of Scientific Explanation and other Essays in the Philosophy of Science*, New York/ London: Free Press/ Collier.
- Jenkins, Carrie S. & Nolan, Daniel 2010, Maximising, Satisficing and Context, in: *Noûs* 44, 451-468.
- Lewis, David 1996, Elusive Knowledge, in: *Australasian Journal of Philosophy* 74, 549-567.



- Lewis, David 1998, The Paradoxes of Time Travel, in: Peter van Inwagen & Dean W. Zimmerman (eds.), *Metaphysics: The Big Questions*, Oxford: Blackwell, 159-169.
- MacFarlane, John 2005, The Assessment Sensitivity of Knowledge Attributions, in: Tamar Szabó Gendler & John Hawthorne (eds.), *Oxford Studies in Epistemology*, vol.1, 197-233.
- Mackie, John L. 1977, *Ethics. Inventing Right and Wrong*. New York: Penguin.
- Montminy, Martin 2007, Moral Contextualism and the Norms for Moral Conduct, in: *American Philosophical Quarterly* 44, 1-13.
- Nagel, Thomas 1979, Moral Luck, in: Thomas Nagel, *Mortal Questions*, Cambridge: Cambridge University Press, 24-38.
- Norcross, Alastair 2005a, Contextualism for Consequentialists, in: *Acta Analytica* 20-2 (35), 80-90.
- Norcross, Alastair 2005b, Harming in Context, in: *Philosophical Studies* 123, 149-173.
- Reichenbach, Hans 1949, *The Theory of Probability. An Inquiry into the Logical and Mathematical Foundations of the Calculus of Probability*, Berkeley & Los Angeles.
- Rieber, Steven 2006, Free Will and Contextualism, in: *Philosophical Studies* 129, 223-252.
- Salmon, Wesley C. 1966, *The Foundations of Scientific Inference*, Pittsburgh, PA: University of Pittsburgh Press.
- Schiffer, Stephen 1996, Contextualist Solutions to Scepticism, in: *Proceedings of the Aristotelian Society* 96, 317-333.
- Schroeder, Mark 2009, Hybrid Expressivism: Virtues and Vices, in: *Ethics* 119, 257-309.

- Sinnott-Armstrong, Walter 1996, Moral Skepticism and Justification, in: Walter Sinnott-Armstrong/ Mark Timmons (eds.), *Moral Knowledge? New Readings in Moral Epistemology*, Oxford: Oxford University Press, 3-48.
- Sinnott-Armstrong, Walter 2006, *Moral Skepticisms*, Oxford: Oxford University Press.
- Stanley, Jason 2005, *Knowledge and Practical Interests*, Oxford: Clarendon.
- Thomas, Alan 2006a, *Value and Context. The Nature of Moral and Political Knowledge*, Oxford: Oxford University Press.
- Thomas, Alan 2006b, *Inferential Contextualism and Moral Cognitivism*, Ms.
- Timmons, Mark 1993, *Moral Justification in Context*, in: *Monist* 76, 360-378.
- Timmons, Mark 1996, *Outline of a Contextualist Moral Epistemology*, in: Walter Sinnott-Armstrong/ Mark Timmons (eds.), *Moral Knowledge? New Readings in Moral Epistemology*, Oxford: Oxford University Press, 293-325.
- Timmons, Mark 1999, *Morality without Foundations: a Defense of Ethical Contextualism*, Oxford: Oxford University Press.
- Unger, Peter 1995, *Contextual Analysis in Ethics*, in: *Philosophy and Phenomenological Research* 55, 1-26.
- Weatherson, Brian 2008, *Attitudes and Relativism*, in: *Philosophical Perspectives* 22, 527-544.
- Wedgwood, Ralph 2006a, *The Meaning of 'Ought'*, in: Russ Shafer-Landau (ed.), *Oxford Studies in Metaethics*, Oxford: Oxford University Press, vol.1, 127-160.
- Wedgwood, Ralph 2006b, *The Context-Sensitivity of 'Ought', and the Normativity of Rationality*, Ms.

Williams, Bernard 1981, Moral Luck, in: Bernard Williams, Moral Luck. Philosophical Papers 1973-1980, Cambridge: Cambridge University Press, 20-39.