


ARTICLE

# Preemptive Authority: The Challenge From Outrageous Expert Judgments

Thomas Grundmann 

Department of Philosophy, University of Cologne, Albertus Magnus Platz, 50923 Köln, Germany  
Email: [thomas.grundmann@uni-koeln.de](mailto:thomas.grundmann@uni-koeln.de)

(Received 4 July 2021; accepted 5 July 2021; first published online 23 August 2021)

## Abstract

Typically, expert judgments are regarded by laypeople as highly trustworthy. However, expert assertions that strike the layperson as obviously false or outrageous, seem to give one a perfect reason to dispute that this judgment manifests expertise. In this paper, I will defend four claims. First, I will deliver an argument in support of the preemption view on expert judgments according to which we should not rationally use our own domain-specific reasons in the face of expert testimony. Second, I will argue that the preemption view does not leave room for rejecting an expert judgment simply because it is outrageous. Third and finally, I will argue that outrageous expert judgments are ambiguous. Whereas some of them should be rationally rejected by laypeople, others are true and rationally acceptable. So, being outrageous is not, in and of itself, a reason to reject the judgment. Finally, I will argue that there are resources available to the preemption view that enable the layperson to reject some but not all outrageous expert judgments. This is sufficient to overcome the challenge from outrageous expert judgments to the preemption view.

**Keywords:** Expert; epistemic authority; preemption; undercutting defeat; social evidence

Experts are highly trustworthy testifiers in their respective domains of expertise. Typically, they are better informed and have superior reasoning skills in these domains than we (laypeople) do. Under normal conditions, we thus rationally defer to their testimony. However, the situation is completely different when experts deliver judgments<sup>1</sup> that seem outrageous in their domains of expertise. With “outrageous judgments” I refer to judgments that strike us as obviously false. When your physician advises you to take 4,000 pills per hour for the rest of your life, when you see a vision scientist who is telling you that she is invisible, or when your trusted TV weather man tells you that it will rain millstones in your area tomorrow, then it seems rational not to follow those experts’ lead. On the contrary, the very fact that your trusted expert tells you something that is so obviously false seems to be an excellent reason to dispute her authority on this matter. This is a core intuition that many people, including most philosophers,<sup>2</sup> share.

<sup>1</sup>In this paper, I use the term “judgment” to remain neutral regarding the controversial question whether the testifier’s belief or her assertion constitutes the testimonial reason. No reference to the agential connotation of judgment is intended. For the latter aspect, see, e.g., Sosa (2017).

<sup>2</sup>See, e.g., Elga (2007: 483), Zagzebski (2012: 116), Lackey (2018).

Interestingly, the intuition is highly relevant for the assessment of two conflicting views about the evidential status of expert judgments for laypeople. According to one of these views, experts are simply *heavyweight epistemic sources* (Fricker 2006; Lynch 2016; Lackey 2018) whose judgments have a high evidential weight relative to a layperson's own reasons. However, this does not mean that the layperson should not rationally use her own reasons. Her final verdict always results from adding the testimonial reasons provided by an expert judgment to her further reasons and weighing all available reasons according to general principles of aggregation. And it might well be that, under specific conditions, the highly confident judgment of the layperson will ultimately outweigh a putative expert judgment. An outrageous expert judgment may just be a case in point. If the layperson is confronted with an obviously false judgment of the expert, the thought is that, in this specific case, it is more rational for the layperson to dispute the expert's judgment and doubt her authority, than it is to defer to it (Lackey 2018: 234, 238). On the alternative view, experts are (*preemptive*) *authorities* (Zagzebski 2012; Constantin and Grundmann 2020). Their judgment that *p* does not only deliver a strong reason for the layperson to follow suit, it also normatively screens off all the other reasons which the layperson may possess with respect to *p*. So, on this view, the expert judgment is not only a strong, but also the only relevant reason for assessing *p*. If this is correct, there seem to be no rational resources left to rationally dispute even the most outrageous expert judgments. On the authority view of experts, we thus seem to be unable to rationally reject outrageous expert beliefs such as the ones mentioned above and that would be a highly counterintuitive consequence of this view. In Lackey's words:

[The authority view] fails to provide the resources for rationally rejecting an authority's testimony when what is offered is obviously false or otherwise outrageous. (Lackey 2018: 234)

One might even argue that this amounts to a *reductio ad absurdum* for preemptive authorities.

In this paper, I will investigate how serious this challenge to the authority view on expert judgments really is. In section 1, I will define the key concepts of the debate. In section 2, I will present what I take to be the best motivation for the authority view. In section 3, I will argue that Zagzebski's own attempt to reconcile the dismissal of outrageous expert judgments with preemptive authorities does not work. In section 4, I will argue that, properly understood, the authority view does not preempt all reasons that are relevant to the assessment of some proposition *p*, but only domain-specific reasons. I will then show that in some, but not all cases of outrageous judgments, the layperson still has sufficient resources to dispute outrageous expert judgments. In section 5, I will address the question whether it is generally problematic for a given view on expert judgments that it cannot rationally dismiss outrageous judgments. I will argue that for a judgment to be outrageous does not generally exclude its truth and that perfectly rational judgments can seem outrageous to laypeople, simply because a layperson's lack of knowledge or training lets it appear outrageous and not because it really is a silly judgment. In section 6, I will discuss some major objections to the view defended here. I will conclude by claiming that the authority view has the resources to dismiss expert judgments that are outrageous for domain-independent reasons and that accepting the rest is not a big bullet to bite. Outrageousness in expert judgments in itself is not a good reason to reject these judgments.

## 1. Key concepts

### 1.1. Experts

Experts often have a reputation for being highly trustworthy and intellectually skillful people in their domain of expertise. According to the reputational conception of expertise (presented by Stichter 2015: 126), being an expert is nothing over and above being a person with this kind of good reputation. More objective conceptions of experts claim that experts must *be* highly trustworthy and intellectually skillful, such that they deserve their good reputation. Among these objective conceptions, there are those which see the main virtue of experts in their ability to facilitate understanding in laypeople (Jäger 2016) or colleagues (Croce 2019), and those which take the reliability of the expert's judgment as the relevant feature. Goldman (2018) claims that the expert is someone whose judgment in a specific domain is much more reliable than that of the average person.<sup>3</sup> This assessment already shows some signs of relativity. What counts as average depends on the social environment and the time. Fricker (2006) defends an explicitly *relative notion of expertise*, according to which E is an expert relative to some subject S if E has a "superior ability to determine the truth" (p. 235) regarding that subject. In this paper I will use this relative notion of objective expertise, such that someone is an expert compared with someone else only if she is epistemically superior to that someone.<sup>4</sup> Moreover, expertise is not only relative to people but also relative to domains. Weather forecaster Myra may be much more reliable in her weather predictions than Tom, while Tom may be doing much better with respect to abstract mathematics. This seems to suggest that expertise is relative to domains of objects. But that is not quite right. If we understand expertise as superior reliability, then it results from being better informed and being more skillful in one's use of methods (Goldman 2001; Constantin and Grundmann 2020). A good example may be astrophysics: In contrast to a layperson, an expert can predict certain astrophysical events – such as the return of a comet – much more accurately. But when the comet appears in the night sky, she is not better than the layperson in determining by eyesight that the comet has reappeared, although this is also an assertion about the domain of astrophysical objects (for a similar view, see Goldman 2001). So, I will work with the following definition of experts:

*E is an expert relative to S regarding a domain D iff E is more reliable than S in judging about D on the basis of the domain-specific method M and relevant input.*

<sup>3</sup>This view obviously needs some further tweaking. First, we wouldn't classify someone as an expert whose highly reliable judgment is fully explained by her trust in a genuine expert or books on the topic. Second, ancient astronomers such as Ptolemy will be classified as experts although their astronomical judgment will be less reliable than the relevant judgment of today's laypeople. A promising line to respond to these problem cases is (i) to require a reliable ability to judge on the basis of domain-specific evidence and (ii) to require an objectively (not time-indexed) highly (conditionally) reliable reasoning capacity (in the domain) and the possession of more relevant evidence than average at the time. Being an expert would thus become a time-relative property. Accordingly, Ptolemy was an expert astronomer in ancient times since he possessed an objectively strong reasoning competence in his domain and more astronomical evidence than average (in his days). The actual Ptolemy would not be an expert astronomer today. However, if Ptolemy lived today he would be an expert astronomer because much more astronomical evidence would be available to him today than it was in ancient times. Third, to be an expert in some domain requires more than mere epistemic superiority with respect to laypeople. For example, a philosophy student with a low C-average is doing better than a student with a low D-average. Nevertheless, we do not consider the C-average student an expert, even relative to the D-student. The epistemic superior must also pass a certain objective threshold in order to qualify as an expert (see Goldman 2018: 5).

<sup>4</sup>My use of the notion "expert" here differs significantly from the non-relativistic use of "expert" in Constantin and Grundmann 2020. However, our view has not substantially changed. It is only a terminological decision.

It is worth noting here that the reliability of the expert also depends on her extensive pool of evidence. Typically, an expert is not only more skillful in her reasoning but also has considered all the relevant evidence, including all of the relevant evidence that is available to the layperson.

### 1.2. Defeat and preemptive reasons

Defeaters are reasons that remove justification. They can do this in two different ways. A rebutting defeater for a belief that *p* is a reason to believe the contrary of *p* such that it outweighs the prior reasons for believing that *p*. In contrast, an undercutting defeater is a reason that indicates that certain pieces of evidence do not properly support a target belief that *p* and thereby renders these pieces of evidence rationally unusable. For instance, if Fred has visual impressions of red widgets in front of him he is *prima facie* justified in believing that these widgets are red. But when a trustworthy colleague tells him that the widgets are illuminated by red light, his belief ceases to be justified. In this case, the testimonial evidence renders Fred's color impressions rationally unusable for the assessment of the widget's color. Reasons that screen off other reasons from being rationally usable in the assessment of some target proposition will be called preemptive reasons. It turns out that, in contrast to rebutting defeaters, undercutting defeaters constitute preemptive reasons. Preemptive reasons are thus not particularly mysterious.

### 1.3. Preemptive authorities

Zagzebski has introduced the concept of an epistemic authority who is – by definition – a source of preemptive reasons.<sup>5</sup> The basic idea here is that by making a judgment on some proposition in her role as an authority, she generates a reason for the judgment's recipient that is evidence for what she says and at the same time screens off all the other domain-specific reasons relevant to the assessment of the relevant proposition. The preemptive reason here doubles as a reason: It is itself a reason for belief in the relevant proposition and a reason that preempts all other reasons regarding that proposition. So far, it is all a matter of definition. Now, the non-trivial and more controversial claim will be that whoever is identified as an expert by a subject *S* will become a (preemptive) authority for *S*.<sup>6</sup> When the identified expert issues a judgment within her domain of expertise, this judgment will constitute a reason for *S* to defer to the expert and to ignore her further domain-specific reasons regarding *p*. We will see in the next section how this identification of recognized experts with authorities can be motivated.

### 1.4. Outrageous judgments

Outrageous judgments are judgments that cause a certain negative reaction in us. We take them to be obviously false, absurd, or we feel the strong inclination to reject these judgments. Here are some typical cases:

1. Your physician tells you that you should take 4,000 pills an hour for the rest of your life (Zagzebski 2012: 116).

<sup>5</sup>Compare Zagzebski (2012: 102): "What is essential to authority is that it is a normative power that generates reasons for others ... to believe something preemptively." For a different notion of authority that comes close to identified relative expertise, compare what Constantin and Grundmann (2020: 4115) call "grounded authority".

<sup>6</sup>Here, identification is understood as a fallible procedure. In this sense, it is sufficient for expert identification if the epistemic agent possesses (fallible) reasons for considering someone an expert.

2. Your calculator displays: “ $2 + 2 = 5$ ”.
3. On TV, the CNN weather man predicts: “It will rain potatoes in the Rhineland tomorrow afternoon” (inspired by Elga 2007: 483).
4. Your trusted pastor tells you: “Women are morally inferior to men” (Lackey 2018).
5. Artificial intelligence people sometimes claim: “Knowledge has turned out false.”
6. Your favorite philosopher asserts: “There are no moral truths/no numbers/no experts.”
7. After a dinner of five friends, two of them calculate everyone’s share in their heads, given that the total amount of the bill is 242 €. Paul claims: “It’s 450 € for each of us.” (For Christensen’s extreme restaurant case, see Christensen 2007: 199).
8. A biologist believes in the truth of a certain theory that contains an obvious contradiction.

The first thing to be said about outrageous judgments is that outrageousness is not an absolute property of a judgment. A judgment’s outrageousness is always relative to persons and times. The judgment that MMR-vaccination causes autism may be outrageous for medical scientists but not for anti-vaccinationists. And that it is not true that two events are either simultaneous or they are not might have sounded outrageous to Einstein before but not after he discovered relativity theory in physics. *So, when a judgment is outrageous it is outrageous for a specific subject S at a time t.*<sup>7</sup>

Moreover, if the judgment that  $p$  is outrageous for  $S$ ,  $S$  seems to disagree about  $p$ . But this is at best a necessary condition. (Shortly, I will argue that, strictly speaking, even this is false.) For illustration, suppose you and your friend do some mental calculation and disagree about the result. Suppose further that you disagree only slightly in that the resulting numbers are quite close to each other. In such a case, you would presumably not assess the result of your friend as being outrageous, although you disagree with him.

This naturally leads to a stronger condition: The judgment that you find outrageous must strongly deviate in its content from your own judgment. One issue with this is that it is difficult to see how the degree of content deviation can be measured when the content is not numerical. Apart from this problem of scope there is the further problem that only in some cases of outrageous numerical judgments, it is true that the relevant number significantly deviates from the one that is taken to be the correct one, e.g. in the Extreme Restaurant Case. However, there are other cases such as “ $2 + 2 = 5$ ” that do not satisfy this feature. 5 does not largely deviate from the correct number 4.

One might think that what really determines whether a judgment is outrageous is whether one disagrees about a fundamental truth. Some examples of outrageous judgments satisfy this condition. Whether there are numbers, moral truths or experts matters for a great number of other propositions and is, in this sense, fundamental. However, examples such as the Extreme Restaurant Case or “ $2 + 2 = 5$ ” do not satisfy this condition.

Alternatively, one might suggest that outrageous judgments are such that one does not understand why they are true. But this condition is not sufficient either. Whenever we disagree about a judgment concerning some proposition we do not understand why the (from our perspective false) proposition is true. However, I have already argued that disagreement is not sufficient for a proposition to be outrageous for us.

A seemingly superior proposal is this one: The judgment that  $p$  is outrageous for  $S$  (at time  $t$ ) iff  $S$  is highly confident (at  $t$ ) that  $p$  is false.<sup>8</sup> This characterization seems to

<sup>7</sup>Outrageousness is also a matter of degree. A judgment can be outrageous to a smaller or greater degree. To keep things simple, I do not address this aspect in the further paper.

<sup>8</sup>Should the class of outrageous judgments be more inclusive? One might want to include cases of Moore-paradoxical assertions (“It is raining but I don’t believe that it is raining”), performative

cover the complete list of examples above. In all cases, the subject seems to be highly confident that the judgment she is confronted with is false. Notice that the subject S need not be highly confident about the truth of her actual belief in order to find a conflicting judgment outrageous. This is so because the conflicting judgment is in most cases not the negation of the actual judgment but some contrary judgment. For instance, S might be moderately confident in the truth of the result of her mental calculation in the Restaurant Case, namely that everyone's share is 48€. At the same time, S can be extremely confident that their share is not 450€.

Despite the initial plausibility of this view, the outrageousness of a judgment for some subject S cannot involve that S confidently believe that p is false. First, it seems possible and is sometimes even rationally required to accept outrageous judgments. Consider a case in which a student in an introductory logic course, using his newly acquired logic skills, finds some argument clearly invalid. Surprisingly, his logic instructor tells him that his impression is misleading and that the argument is in fact valid. Since the instructor is in a hurry, she cannot explain to the student why the argument is valid. From the student's point of view, the instructor's judgment is outrageous. What he takes to be a clear case of an invalid argument is claimed by her to be valid. It nevertheless seems rationally required of the student to accept the instructor's verdict about the argument. Now, you can only accept a proposition if you are confident that it is true. Since it is possible that one can accept an outrageous judgment without its outrageousness being lost, outrageousness cannot be defined in terms of confidently believing that the target proposition is false.

Secondly, it seems possible that S finds p outrageous although she herself believes (or even knows) p at the same time. Accordingly, S might classify even one of *her own* judgments as outrageous. Suppose your visual impressions strongly support a certain judgment. Then you learn that you are misled by a cognitive illusion or hallucination. You now know that things are not as they visually appear to you. Although you know better, the misleading visual impressions are still in place, however, and cause a strong inclination to believe what the impressions suggest. Your tutored belief is then in massive cognitive dissonance with what your visual reasons (which you know to be misleading) suggest. This tutored belief is thus hard for you to hold, it is an outrageous judgment. Similar cases are cases in which your everyday experience is in tension with your physical theory (e.g., relativity theory) or in which your tutored belief is in tension with your pre-theoretic intuitions.<sup>9</sup> The lesson to be learnt from cases like these is that outrageousness cannot be defined in terms of confidence (which is a degree of belief), but must be defined more liberally. I therefore suggest the following definition:

*A judgment that p is outrageous for subject S (at t) if and only if S has the strong impression (at t) that p is false.*

That some proposition strongly seems false to S must be understood non-doxastically. It is a strong inclination to believe or something like a strong epistemic feeling of falsity.

contradictions ("I don't speak now"), contradictory propositions ("I know that I don't know anything"), or meaningless utterances ("Green numbers fly low"). However, I take it that Moore-paradoxical assertions, performative contradictions, and contradictory propositions all belong into the class of obviously false judgments. Meaningless utterances are not considered because they do not constitute propositions.

<sup>9</sup>Imagine that you put a rope tightly around the earth's equator. Question: if this rope were 12 meters longer, what would be the average distance of the rope from the surface? Intuitively, people tend to reply that the distance would be very close to zero. In fact, it would be 1.9 meters. That is a deeply surprising provable result that seems outrageous to the layperson.

The appearance of outrageousness can but need not be based on reasons. Consider again the case of a student in the introductory logic course. Even if some argument appears to be clearly invalid to him and anyone's conflicting judgment will seem outrageous, his assessment will not rely on reasons.<sup>10</sup> If it is based on reasons, those reasons can be misleading, but they can also be known to be misleading and thus be defeated. Even defeated reasons may cause the appearance of outrageousness, however. Hence, outrageousness is a psychological rather than an epistemological property of judgments.

## 2. The motivation for experts as preemptive authorities

What is the rationale behind the claim that whenever someone has identified an expert, this expert has the status of an authority for them? Suppose Sara is a pupil who does not know anything about relativity theory. One day, she learns from her teacher that the expert physicist Einstein discovered that it is not true that two events are either simultaneous or they are not. This is clearly a reason for Sara to defer to Einstein. But what happens to Sara's background reasons regarding this simultaneity claim? What happens to her everyday experience of simultaneity? If Sara is clever, she will understand the situation in the following way: As an expert physicist, Einstein clearly has superior reasoning skills in the field of physics, but he has – most probably – also considered all the relevant evidence that is available to laypeople such as me. Moreover, it is highly likely that Einstein responded rationally to the total available domain-specific evidence because he has excellent domain-specific reasoning skills. However, it is also true that when Sara considers domain-specific evidence on her own and aggregates it with the evidence of Einstein's judgment, she will deviate (at least slightly) from his judgment. This becomes more explicit when we talk about degrees of belief or credences. When adding the evidence of the expert judgment to the domain-specific evidence (either publicly available data or Sara's private experiences), it is highly likely that Sara will end up with a credence that differs (at least slightly) from the expert's credence. After all, the pool of total evidence Sara relies on differs from Einstein's, due to the evidence she adds to his judgment. All these basic facts are fully disclosed to Sara. But then she can see that by weighing further domain-specific evidence with the evidence of the expert judgment she will deviate from what *she herself* takes to be the clearly superior judgment of Einstein who has already considered all of Sara's own evidence. Hence, she recognizes that considering these further pieces of evidence would lead to a judgment that is not fully rational. She might factor in evidence that has already been taken into account by the expert (this would be a case of double counting), or she might use evidence that is clearly irrelevant or false from the expert's perspective, or she might make a mistake in weighing the evidence on her own. In any case, she has evidence that the expert's judgment is the result of the rational weighing of all the relevant evidence. At this point, the pupil should realize that her domain-specific reasons are not rationally usable because they would lead to a judgment that is irrational to the extent that it deviates from the expert judgment. So, taking the expert's judgment seriously does not only generate a reason to follow his lead but also generates a reason that preempts the pupil's further domain-specific reasons.<sup>11</sup>

<sup>10</sup>For this reason, I do not accept the suggestion (made by Jason Kawall in his comments) that a non-psychological, epistemic definition of outrageous judgments is possible. Accordingly, outrageous judgments would be judgments against which the layperson has strong domain-specific reasons. Even if such a definition were plausible it would not change my overall assessment that outrageous judgments need not be rationally rejected since strong domain-specific reasons may be undermined by testimonial reasons.

<sup>11</sup>One might worry that the details of the rational assessment of the proposition by the expert should not play any rational role for the layperson's evaluation unless they are fully disclosed to the layperson. However, even if the layperson does not know all the details of the expert's cognitive perspective, she



Before I present this line of thought as a semi-formal argument, I must highlight some relevant background assumptions. The identified expert judgment generates a preemptive reason for the pupil’s reasons with respect to p under the following three necessary conditions: (a) ONLY IF there is no reason to believe that the pupil has domain-specific reasons on her own *that are not shared with the expert*, (b) ONLY IF there is no reason to believe that the expert judgment is based on domain-independent reasons, and (c) ONLY IF there is no reason to believe that different equally rational answers to the total evidence are possible in this particular case. It seems obvious that the expert judgment will not determine the rational response to the relevant evidence, if Sara has gathered new, so far unknown evidence (e.g., she might have made a relevantly new observation that is not shared with the expert), or if the expert judges about propositions outside his area of expertise, or if different responses to the same evidence can be equally rational (as permissivism<sup>12</sup> claims). However, these would be rather atypical cases. Typically, when a layperson confronts an expert, these conditions will be satisfied, and the layperson will know this.

Here is, then, the argument for the preemptive character of a recognized expert judgment:

- (P1) When the layperson, in her own judgment, considers pieces of the domain-specific evidence (considered by the expert) in addition to the expert judgment, the resulting judgment will deviate from the expert judgment.
- (P2) Most likely, when the layperson’s judgment deviates from the expert judgment under these conditions, it will be irrational.

Therefore,

- (C) When the layperson, in her own judgment, considers pieces of domain-specific evidence (considered by the expert) in addition to the expert judgment, the resulting judgment will, most likely, be irrational.

This argument is clearly valid. It has the form:

- (1)  $p \rightarrow q$
- (2)  $q \rightarrow r$
- 
- (C)  $p \rightarrow r$

We can use a conditional proof to show its validity:

- 1 p assumption
- 2  $p \rightarrow q$  premise (1)
- 3 q  $\rightarrow$ Elim : 1, 2
- 4  $q \rightarrow r$  premise (2)
- 5 r  $\rightarrow$ Elim: 3, 4
- 6  $p \rightarrow r$   $\rightarrow$ Intro: 1-5

---

has some evidence of how experts generally assess the evidence (including the layperson’s evidence). As it seems to me, this unspecific evidence of evidence is sufficient to support my argument.

<sup>12</sup>Kelly (2013).



Now, let me clarify the premises. (P1) appears particularly plausible if one thinks about judgments in terms of graded beliefs or credences. If the layperson considers further evidence apart from the expert judgment, then this will have at least a minimal weight that, in the aggregation, will make at least a minimal difference. (P2) claims that the expert judgment (most likely) sets the standard for a rational response. This is only true, if the three conditions explicated above are satisfied. It is true (a) only if there is no reason to believe that the expert judgment is based on domain-independent reasons, (b) only if there is no reason to believe that the layperson has additional domain-specific evidence that is not shared with the expert, and (c) only if there is no reason to believe that permissiveness holds for a rational response in this case. This qualification has two significant consequences: First, what is preempted are the domain-specific reasons of the layperson only. Second, these domain-specific reasons are preempted, only in so far as the layperson has no reasons to believe that they are not considered by the expert. In this case, the evidence of the layperson is a subset of the expert's evidence. All pieces of the layperson's evidence that are not taken to be considered by the expert do not get preempted. This means that preemption through the expert judgment is always restricted in the indicated ways.

The argument sketched above is just one among other arguments<sup>13</sup> that have been delivered to support the preemption view.<sup>14</sup> Without going into further details here, it seems to be the most promising argument. Its main advantage is that it explains preemptive reasons as undercutting defeaters – a phenomenon that everybody should be happy to accept.

### 3. Why Zagzebski's attempt to cope with the dismissal of outrageous expert judgments fails

On Zagzebski's view, any reasonable view on the nature of expert judgments must cope with the challenge from outrageous expert judgments. For her, it must be possible to reject outrageous judgments. She also claims that the authority view has the resources to cope with this challenge.

There is another source of resistance to preemption ... Suppose the authority's belief is something outrageous. Can't that count as a defeater of your authority belief? Yes it can, but that does not count against preemption. (Zagzebski 2012: 116)

What does Zagzebski's response to the challenge look like? She believes that trust in experts is based on reasons and that we can use these resources to reject the thesis that experts are authorities, if they hold outrageous beliefs.

[I]t remains the case that a general trust in ourselves leads us to trust in authority, and the judgment that someone is an authority can be withdrawn. (Zagzebski 2012: 116)

So, Zagzebski claims that preemptive authority is justified and can also be defeated. But how is this going to work in detail? One natural worry here is that you cannot use preempted reasons as a defeater for the outrageous expert judgment. This is so because

<sup>13</sup>See, for example, Zagzebski's (2012: 114–15) and Raz's (1988: 68–9) track record argument or Keren's (2007: 374) argument from double counting.

<sup>14</sup>For a comprehensive defense of this view see Constantin and Grundmann (2020).

preempted reasons are not rationally usable. However, Zagzebski seems to claim exactly this, namely that preempted reasons can be used against their defeater (as defeater-defeaters), if the authority judgment is too absurd in light of the preempted reasons. There is another option here for the proponent of the authority view, namely to claim that preemption is limited to domain-specific reasons and that we can use the domain-independent reasons that are untouched by preemption to reject outrageous judgments and authority. But Zagzebski does not take this route. This becomes clear when she addresses the 4,000-pills case. There, she argues that the reason why we (as laypeople) can reject the judgment is not that it is physically impossible to follow the expert's recommendation (as in a case with 40,000 pills) but that the judgment is outrageous (Zagzebski 2012: 116, fn. 12). She thinks that reference to the 40,000 pills case blurs this point. On my interpretation, when it becomes physically impossible to take all these pills, we have domain-independent reasons (which are not preempted) to reject the judgment. In contrast, Zagzebski seems to argue that you can even use your domain-specific reasons to defeat preemptive authorities. It is this claim that I find hard to swallow.

Why is it impossible to defeat the authority judgment on the basis of the defeated domain-specific reasons? There are two main objections to this. First, that a reason is preempted means that it is rationally unusable and thus counts for nothing. It seems obvious to me that you cannot use a reason that has literally zero evidential weight to overrule a judgment. In order to do that the reason would still need some minimal weight, which preempted reasons do not have. The evidential role of such reasons can be illustrated by the epistemological mechanism of undercutting defeaters that are paradigm cases of preempting reasons. Consider again the case of red impressions under red-light illumination. When we are informed that the widgets that appear red to us are illuminated by red light, the evidential weight of the red impressions is undermined. In such a situation, red impressions are preempted reasons. Can we still use these impressions to defeat the defeater? No. It will not help, if we ourselves and other people experience the redly illuminated widgets as red again and again. Even an enormous amount of extremely vivid red impressions would not overrule the undercutting information about the red lights. The explanation for this is that the red impressions have zero weight and thus cannot make a difference to what is rational or justified to believe on the matter.<sup>15,16</sup> A defeater of the undercutting defeater would have to be source-independent. Such a defeater-defeater might, for example, be that another person informs us that our original testifier was lying to us about the red lights.

Second, when we use preempted reasons to reject authority judgments, we confuse them with reasons that were overruled but not preempted. When a reason is overruled, it still counts. It is simply weaker than the contrary reason. But it might be aggregated with further reasons such that, as a result, it overrules its defeater again. In the case of

<sup>15</sup>It is therefore simply irrelevant whether or not the first-order evidence is stronger than the higher-order undercutting defeater. A sufficiently strong undercutting defeater will screen off the first-order evidence no matter how strong it is.

<sup>16</sup>Here I deal with somewhat idealized cases in which the epistemic agent has sufficiently strong reasons for identifying someone as an expert. In some messy real-world cases, only weak (for example, second- or third-hand) evidence of expertise may be available. In these cases, evidence of expert judgments generates partial defeaters at best, which leaves open the option to reject outrageous judgments of *putative* experts simply because they are outrageous. I do not want to exclude these kinds of non-ideal cases. It may still be true that in the majority of real-world cases, we have sufficiently strong reasons for expert judgment to generate a full undercutting defeater. Thanks go to Jason Kawall, Maria Lasonen-Arnio, Giacomo Melis and Jim Pryor for pointing out to me that undercutting defeat is not always all-or-nothing but rather a matter of degree.

rebutting defeaters there is a continuing aggregation process. But treating reasons like this differs fundamentally from treating them as preempted (for this objection, see also Lackey 2018: 237–8).

The lesson to be learned here is that even the authority view can explain how outrageous judgments of experts can sometimes be defeated. But this explanation must always refer to domain-independent reasons that are not preempted. This also means that the expert's outrageous judgments cannot be dismissed simply in virtue of being outrageous. Everything depends on the right reasons for their outrageousness.<sup>17</sup> (More on this in the next section.)

There is one further concern about the authority view that needs to be addressed here. Jennifer Lackey objects to the authority view that it does not have the resources to *identify* outrageous beliefs as such. As she puts it:

The testimony of an authority replaces all of one's other reasons on the topic. Given this, in virtue of what can the testimony of an authority strike one as crazy? (Lackey 2018: 235)

I think that there is a clear answer to the question she poses. On the authority view, the “replacement” of reasons is an exclusively normative issue. The “replaced” reasons do not vanish metaphysically but are rendered rationally unusable. Now, if these reasons are still there, they also can explain why the expert judgment appears outrageous to us. Judgments may feel decidedly false to us on the basis of reasons that are already defeated. Something like this happens for example in the case of sensory or cognitive illusions where the truth is very hard for us to believe. In the definitional part of this paper, I defended a non-doxastic reading of outrageousness. On that reading, reasons that are preempted can still create the impression of obvious falsehood.

#### 4. How outrageous judgments by preemptive authorities can be rationally rejected

One thing that can be learned from the discussion of Zagzebski's view is that if we accept the authority view, we cannot use domain-specific reasons to reject an experts' outrageous judgment. In this section, I will explore to what extent there are further resources available to a layperson that may justify the rejection of these outrageous judgments. I will proceed by first investigating whether there are some general structural limitations to preemption. Then, I will discuss the cases listed above one by one to find out whether there are specific resources available in each of these cases to reject the expert's outrageous judgment.

<sup>17</sup>It is sometimes argued that even judgments of experts may be false due to performance errors, biasing factors, partly domain-independent reasons or the lack of serious belief on behalf of the expert. These error possibilities cannot be ruled out by the fact that the judgment is made by a genuine expert about the relevant domain. The outrageousness of judgments may be thought to be a reliable indicator of these kinds of errors. However, this does not follow. Compare the following argument: When we base our judgment that the widgets are red on visual impressions of their redness, our justification is undercut as soon as we are informed by a trustworthy informant that the widgets are illuminated by red light. Of course, we do know that our informants are not infallible. They might confuse the actual situation with another situation in which the red illumination is in fact present; they might be misinformed by a further informant; they might be bribed or just telling a joke. All of this is clearly possible. However, it would be completely inappropriate if we took the intensity of the red-impressions as an indicator of these kinds of errors. We can generally never use the undercut evidence as a reason to defeat the undercutting defeater. Further indicators of misinformation, bribery or joking that are independent of the already undercut evidence are needed. The same applies to the case of expert judgments. Thanks to Paul Boghossian and Jason Kawall for pressing me on this point.

One principled limitation of preemption is *First Person Authority*. When a physician, after a brain screen and further testing, tells you that you don't have pain although you feel terrible pain yourself, the physician's judgment is outrageous for you and you seem to have every right to reject this judgment. Interestingly, this verdict can be accepted by the proponent of the authority view. This is because the physician's judgments about your pain are not based on the excellent method on which your judgments about your pain are based upon. Whereas the physician only has third-personal access to your pain, your own judgment relies on first-personal introspective access. This methodological difference explains why your reasons in this case are domain-independent reasons, since domains are, as I have argued above, methodologically rather than metaphysically individuated. So, if your physician tells you "You don't have pain" when you feel terrible pain, you can rationally reject this outrageous judgment, and you can do this even if you are a proponent of the authority view.

A second limitation of preemption concerns expert judgments that are undermining trust in experts. Suppose someone whom you regard as an expert in the field of psychology tells you that she is no expert, that there are no experts or even that persons do not exist. Or suppose an expert philosopher tells you that there are no preemptive reasons or that there are no meaningful judgments. In these cases, the expert tells you something about her field of expertise that, if true, would undermine your trust in her preemptive authority. You cannot, at the same time, accept her as an authority and believe her preemptively. The class of judgments characterized by this feature thus cannot, for structural reasons, be accepted on authority.

It is tempting to claim that there is a further class of reasons that cannot be preempted through authority – a priori reasons. There is hardly any current rationalist who believes that a priori justification is generally infallible or incorrigible.<sup>18</sup> For instance, if we do mathematics in a non-empirical way, there is always the possibility of making a performance error or of further progress through the revision of some axioms. However, some philosophers hold the view that a priori justification cannot be defeated by empirical reasons.<sup>19</sup> If this were true, a priori reasons could never be preempted by the testimony of experts. This would have the welcome consequence that one could use one's own a priori insights to reject outrageous mathematical expert judgments such as " $2 + 2 = 5$ " or Paul's judgment in the Extreme Restaurant Case. But there is good reason to doubt that a priori justification is empirically indefeasible and thus cannot be preempted by expert judgment. There are three main reasons that support the empirical defeasibility of a priori justification: First, our beliefs can be justified by more than one source. That, for instance, my pencil lies on my desk is something that I can justify by looking at it, by remembering that I put it there a short time ago, or by relying on a trustworthy testifier. But if beliefs can be justified by different sources why shouldn't they be defeasible by more than one source? This general feature of epistemic overdetermination<sup>20</sup> is a prima facie reason for believing that a priori justified beliefs can be defeated by a priori as well as empirical reasons. Second, suppose that you try to make a complicated calculation. You are a bit unsure about its result and, hence, you ask an expert mathematician for advice. She tells you that you are almost right but have to revise the numbers a little bit. To me, this testimony looks like a perfect rebutting empirical defeater of my a priori reasoning. Third, there are examples of empirical undercutting defeaters for a priori justification. Suppose that neuroscientists discover that a certain pattern of neurological activity NA is

<sup>18</sup>A representative of current fallibilist rationalists is, for example, BonJour (1998).

<sup>19</sup>See Kitcher (2000: 67).

<sup>20</sup>For more on this see Casullo (2003: 120).

correlated with erroneous mathematical proofs. In a specific situation, you write down a certain mathematical proof and the neuroscientists tell you that NA has recurred. Intuitively, your belief based on the proof has been undercut.<sup>21</sup> In light of these reasons, one should accept that a priori justification is empirically defeasible. It is, therefore, not excluded from preemption. A more specific case of rationally indefeasible beliefs would be beliefs that are meaning-constitutive. According to this view, it is impossible to believe that Archibald is a bachelor without believing (or being disposed to believe) that Archibald is unmarried. If one were to give up the latter belief one would cease to believe that Archibald is a bachelor. If there are beliefs that are constitutively related to concepts, these beliefs could not be defeated by expert judgments. In contrast, we could always use these beliefs to defeat the conflicting beliefs of experts. However, whether there are such beliefs involved in conceptual competence is highly controversial.<sup>22 23</sup>

What are, apart from these structural limitations of preemption, the domain-independent resources for rejecting outrageous expert judgments? As far as I can see, there are three major sources of evidence typically available: social evidence, analytic evidence, and evidence from domain-independent disciplines. Let me address these sources in due order. First, consider *social evidence*. We often do not only know what one expert judges in her field of expertise, but also know how many other experts in the field agree or disagree with this judgment. This “social” information can be a domain-independent basis for a rational rejection of an individual expert judgment. If, for instance, the majority of relevant experts rejects the view of one individual expert, then this is a good reason for the layperson to reject this judgment as well. Sometimes, even social information about the agreement among laypeople can justify the rejection of expert judgment. Suppose there are experts and laypeople on mental math. Let us assume that the expert has a truth ratio of 90% and that the laypeople are still minimally competent such that they have a truth ratio slightly above 50%. If 1,000 laypeople all do a given calculation independently and if they fully agree about the result, then information about this fact will outweigh a diverging expert judgment.<sup>24</sup> Second, consider *analytic evidence*. Suppose an expert physicist writes a book about physics which contains a number of fallacies, contradictions or conceptual inconsistencies. According to the authority view, we as laypeople cannot reject her views on the basis of our physical reasoning. However, as long as she is not an expert in logic or linguistics, we can use our domain-independent logical or conceptual reasons to challenge her view. This is clearly true, if the lay physicist is an expert logician. Third, consider *evidence from domain-independent disciplines*. Some expert judgments have implications for other disciplines. For example, if a weather forecaster predicts that it will rain potatoes tomorrow, her claim has also, for example, physical implications regarding the mechanics of potatoes and weather. According to the authority view, we are not permitted to use our domain-specific reasons (about meteorology) to dismiss this claim. But we can nevertheless draw on our knowledge of physics to do so.

<sup>21</sup>For this case see Casullo (2003: 117).

<sup>22</sup>A proponent of this view is, for example, Boghossian (2003); Williamson (2007) is strongly opposed to this view.

<sup>23</sup>There may be even further propositions that are structurally indefeasible such as the constitutive rules of rational discourse. Compare the German transcendental pragmatists such as Karl-Otto Apel and Wolfgang Kuhlmann.

<sup>24</sup>For the mathematical calculation of probabilities in Condorcet-like situations see Landmore (2013: 166–9).

Let me apply these abstract categories to the eight examples on the list of outrageous expert judgments and see to what extent we can reject these judgments on the basis of the remaining domain-independent reasons. The first example is the case of your trusted physician who tells you that you should take 4,000 pills an hour for the rest of your life. This judgment seems outrageous. What are the resources available to the proponent of the authority view that may allow her to claim that a layperson can reject it? Whereas taking 40,000 pills an hour might be physically impossible, taking 4,000 pills may not be. But there is social evidence available to most laypeople that permits rejecting this judgment. After all, you would have heard about prescriptions of such a large number of pills, if any had been given out.<sup>25</sup> This indicates that your physician represents a minority view among the relevant experts. And this is, in turn, a domain-independent reason to reject her prescription. However, we can revise the case such that you live in an informationally isolated community in which there is no access to further medical research. Then, one can no longer argue that there are indicators that your physician represents a minority view. Under these conditions, you would not have any reason to reject his judgment. It thus would be rational for you to accept that should take 4,000 pills an hour for the rest of your life. Of course, one might argue that this is a case in which the practical stakes are high. Your very survival might depend on the right decision here. One might then argue that in a high-stakes case the threshold for justification is so high that a single expert opinion cannot meet it. You would then need to ask additional physicians for advice to pass this threshold. Alternatively, one might argue that, even if the threshold for justification is met, the threshold for action (namely, following the physician's prescription in practice) may not be. But this applies equally to all medical judgments and not specifically to the outrageous ones. Even in the case of a medical treatment that is perfectly in line with our expectations we then would have to ask additional physicians for advice. In conclusion, one has to admit that there may be variants of the 4,000 pills case in which it would be rational to follow the physician's outrageous prescription.

How should we think about the second example in which your expert calculator displays: " $2 + 2 = 5$ "? A straightforward response would be that the belief that  $2 + 2 = 4$  is constitutive to the meaning of basic numeral concepts and hence, in principle, not empirically defeasible and we thus can use our own belief to dismiss the calculator's calculation. But as we have seen above, the view that some kind of justification is empirically indefeasible is controversial. However, there are further domain-independent reasons to reject the calculator's result. We have social knowledge of the majority view among mathematicians on whether  $2 + 2 = 5$ . We have learnt in school and from the textbooks that  $2 + 2 = 4$ , which further indicates this fact. And we may rely on this domain-independent social knowledge when we reject the calculator's result as too exotic. One may also argue that the expert calculator is superior to us with respect to complicated mathematical problems rather than simple ones, such as " $2 + 2 = 4$ ". If this is correct, the calculator cannot be taken to be an expert with respect to these kinds of simple problems, even relative to mathematical laypeople. Then, however, laypeople are not rationally prohibited to use their own mathematical reasoning when solutions to simple mathematical problems become controversial.<sup>26</sup>

What about the forecaster's judgment that it will rain potatoes in the Rhineland tomorrow? First of all, one might find this judgment outrageous because, strictly speaking, it involves a conceptual inconsistency. It is analytically true that it rains raindrops, i.e., aggregations of water or at least liquids, when it rains. So, we might just use our – undefeated –

<sup>25</sup>See Goldberg (2010: Ch. 6).

<sup>26</sup>Thanks to Aleks Knocks for drawing my attention to this dialectical move.



conceptual competence to reject the forecaster's claim. But one can also understand her use of "rain" metaphorically. It rains potatoes in this sense if and only if potatoes will fall from the clouds, and that is conceptually possible. It also seems meteorologically possible that a strong tornado lifts up all kinds of objects from the surface that will fall down shortly after that.<sup>27</sup> An expert forecaster might be able to exactly predict such an event. Nevertheless, it is *physically* highly improbable that the wind will pick up nothing but potatoes and it is also *physically* highly improbable that the wind picks up any potatoes since potatoes are objects that grow underneath the surface and are thus not a primary target of strong updrafts. So, this is a clear case in which a layperson can use her domain-independent physical and general knowledge to reject the forecaster's judgment.<sup>28</sup>

The next example is the case of a locally highly trusted moral expert such as a pastor. Suppose such a "moral expert" tells you: "Women are morally inferior to men." Should you, as a "layperson", adopt this attitude? The case is supposed to be set up in such a way that we take the pastor's judgment to be false. Would it still be rational for a woman living in the relevant community to defer to this moral authority without thinking about the moral issue herself? If the intuitive verdict is "no", then there are still resources available to the authority view that can be used to dismiss this judgment. The most natural suggestion in this vein would be that we have social knowledge of a majority of other moral experts that disagree with the pastor. However, we can revise the case in such a way that the community is, again, informationally isolated. In this version, no member of the community has access to disagreeing moral experts. The proponent of the authority view would thus be forced to admit that laypeople then should follow the false judgment of the pastor. One simply has to admit that rational belief is fallible and sometimes misleading.

When an AI-scientist tells you that knowledge has turned out false, you may rely on your – unpreempted – analytic (conceptual) competence to reject this judgment. It is simply an analytic truth that knowledge requires truth.

What about the case of an expert philosopher who advocates an error theory about morality, numbers or even experts? If the expert disputes expertise, we have the case of a judgment that undermines trust in authority and thus *in that very judgment*. You therefore cannot believe what the expert judges preemptively. Error theories about numbers or morality must be treated differently. Here, again, we have social knowledge that tells us that we should not follow this exotic minority view in philosophy.

Next, consider the Extreme Restaurant Case. In this case, Paul diverges with his judgment to an extreme extent from our intuitive judgment and from all other judgments that are within the range of minimally plausible results. His answer is thus clearly outrageous. According to the authority view, we cannot rely on our mental calculations and the credences related to it. But there are clearly further resources available that allow the proponent of the authority view to reject this judgment. A proper subset cannot be larger than the set. This is a *generally* accepted set-theoretical truth. But Paul's judgment conflicts with this truth (a truth which he most probably accepts himself). So, one might argue that we can detect a logical inconsistency among Paul's mathematical judgments. And this is a domain-independent reason to reject his judgment. Of course, Paul might restore consistency by giving up the set-theoretical judgment. But we could object to this move by referring to the majority view among the mathematical experts.

<sup>27</sup>In the initial case of millstones this would be *physically* almost impossible.

<sup>28</sup>Steffen Koch raised the objection that it is highly likely that a first-rate meteorologist will also have some competences in physics such that she is superior to laypeople even in this domain. However, the layperson always has access to social evidence of physicists who take events like the raining potatoes to be highly unlikely.



In the final example, a biologist believes a theory that involves an obvious contradiction. As laypeople, we can easily reject this theory by relying on our domain-independent logical skills.

To sum up, one can say that six out of eight cases of outrageous expert judgments are such that we can reject them without relying on the preempted, domain-specific reasons. In the other two cases, we were unable to reject the expert judgment on the basis of domain-independent reasons, although in both cases the expert judgment is probably false. But this need not be an overall bad result since we have to keep in mind that rationality is not truth-guaranteeing.

## 5. Outrageousness and rationality

Initially, there seems to be a strong conflict between the core intuition that outrageous expert judgments should be rejected, simply because they are outrageous on the one hand and the authority view of experts according to which our own reasons are preempted by the expert judgment and hence cannot be used to reject the outrageous expert judgment on the other hand. By now, it has turned out that the alleged conflict is much weaker than expected. First, identified expert judgments<sup>29</sup> do not preempt all our reasons but only our domain-specific reasons. Second, the reasons that are not preempted are sufficient to justify the rejection of outrageous expert judgments in many, though not in all cases. In this section, I want to focus on the remaining conflict and will come to a rational assessment.

It is true that the authority view recommends accepting outrageous expert judgments even in some cases in which this judgment is false. So, you might think that it would be advisable to reject all outrageous expert beliefs and, as a consequence, also the authority view. However, there are also outrageous expert beliefs that strike the laypeople as obviously false but that are nevertheless true (and well justified). That it is not true that two events are either simultaneous or that they are not is something that follows from one of the best confirmed theories in physics, relativity theory, but it is very hard for the layperson to accept this implication. Something similar is true for ontological indeterminism and non-locality within quantum physics.<sup>30</sup> Here are two further, more mundane examples. The first is about the collapse of the World Trade Center at 9/11. According to the data on this collapse, it took the Twin towers 14 to 16 seconds to collapse. In comparison, it would take something to free-fall from the top of this building 9 seconds. The difference is so small that laypeople do not understand how the buildings could collapse so quickly without a controlled demolition, i.e., without controlled explosions on many of its floors at the same time. Keep in mind that the World Trade Center was a very solid construction. The outrageousness of this fact is clearly expressed by Rosie O'Donnell, one of the "Truthers", who characteristically express incredulity:

Do you know how fast it took those towers to fall? Nine seconds ... You know how fast it would have taken something to free-fall from the top of that building? Nine seconds. It's physically impossible.<sup>31</sup>

However, according to the community of experts, simulations have proven that a time like the actual time of the collapse was to be expected under the given circumstances.

<sup>29</sup>Keep in mind that I do not use this term in a factive manner. It can always turn out that what I identify as expert judgment does in fact not result from expertise.

<sup>30</sup>For many more examples of scientific truths that seem counter-intuitive, unnatural or outrageous to the laypeople, see McCauley (2011: 105–17).

<sup>31</sup>Cited from Dunbar and Reagan (2006: 42–3). Obviously, O'Donnell did not report the exact numbers but speeded the collapse up a bit to make her point more vivid.

Here we have a clear case of an expert judgment that appears outrageous to some laypeople although it is correct.

My second example is the Monty Hall Puzzle. Suppose you are in a game show. Behind one of three doors is a brand-new car that the player will win if she picks the correct door. The player picks, say, door number one. Then the host opens one of the other doors, say door number three, showing that this is not the correct door. The host now offers the player the option to switch from door 1 to 2. Should she accept this offer to raise her chances? A number of scientific studies has shown that an average of 85% of people do not switch (Bruns and Wieth 2004). For them the chances seem equal and hence there is no reason to switch. However, as expert mathematicians have proven and also simulated, the player's chances are raised from  $1/3$  to  $2/3$  if she switches.<sup>32</sup> This is an outrageous but true expert judgment.

Often but not always, outrageous expert judgments turn out true. So, outrageousness is not a guarantee of falsity. On the other hand, the psychological studies by Kruger and Dunning (1999) and Dunning (2005) suggest that the attribution of competence to oneself and the attribution of incompetence to others is strongly biased and not competently produced when the relevant person is incompetent herself. As Dunning shows, incompetent people are blissfully unaware of their own incompetence because the meta-cognitive cues of competence (such as fluency or the availability of reasons) are not reliably exploitable as evidence by incompetent people. Hence, incompetent laypeople tend to assess themselves as competent. When they are highly confident in this self-assessment they will classify most opposing views as outrageous.

The impression of outrageousness is clearly not a mark of falsity (too often outrageous expert beliefs turn out true). Moreover, outrageousness itself cannot be treated as a reason to reject the outrageous judgment. However, epistemic feelings probably have an adaptive epistemological function (Hookway 2003). They may serve as a heuristics for cognitive regulation. For example, we select good inferences because they feel compelling, we choose among alternative explanations the one that feels most appropriate, or we assess relevance on the basis of immediate salience. Under appropriate conditions the spontaneously occurring feeling of outrageousness may indicate that the considered judgment is massively irrational. What triggers this emotion may be a strong cognitive dissonance with the agent's background beliefs or natural appearances (McCauley 2011) or massive disfluency in cognitive processing (Alter *et al.* 2007; Thompson *et al.* 2011; Levy 2017). Outrageousness may thus provide us with a reliable heuristics of selecting strongly irrational judgments when conditions are normal. However, this heuristics may easily turn into bias if one confronts expert judgments. This *anti-expert bias* results from the fact that expert judgments are most likely in conflict with laypeople's background beliefs (since they are innovative or even revolutionary) and processed less fluently (because they are unintuitive). As I have already argued, being in conflict with laypeople's (prima facie rational) background beliefs does not render expert judgments irrational, if one keeps in mind that they preempt the laypeople's domain-specific reasons. Given the anti-expert bias in the mechanism behind the outrageousness impression, we should not expect rational guidance from them when we confront expert judgments. If we really want guidance in these cases, we have to look at the reasons behind the impression of outrageousness.

## 6. Objections and replies

In this section, I will discuss three objections to the view defended here. According to the first objection, preemption generally requires deference to the judgment of

<sup>32</sup>For a comprehensive discussion, see Rosenhouse (2009).

epistemic superiors, even if their judgment is only slightly more competent. Such an epistemic behavior may easily have epistemically terrible consequences. Consider Bernard Williams and Philippa Foot, and imagine that one of them – it needn't concern us who – is slightly more reliable in her judgments about moral philosophy than the other. It would be clearly disastrous for philosophical practice if Williams or Foot simply deferred in their judgments to their counterpart since it would be the end of all fruitful debates in philosophy. This looks like a *reductio* of the preemption view.<sup>33</sup>

In reply, two things can be said. First, what is epistemically required and what is instrumentally valuable with respect to certain epistemic ends are two completely different things that may conflict under specific conditions. For example, conciliatory views on peer disagreement rationally require that the parties suspend judgment under fully disclosed peer disagreement. This rational requirement obviously conflicts with the fact that only sticking to one's guns throughout intellectual controversies promotes intellectual progress. However, this does not show that there is no rational requirement to suspend judgment in the face of peer disagreement in the first place. So, having epistemically bad consequences does not automatically speak against being rationally required. Second, expert controversies among near-but-not-perfect peers are often such that the opposing parties do not fully share their evidence. Typically, the parties rely on their own insights, original arguments or different perspectives on the topic under discussion to support their views. This keeps the debate going. If that correctly describes their epistemic situation, the conditions for preemption are not fully satisfied. It was argued above that expert judgment generates a preemptive reason ONLY IF there is no reason to believe that the pupil has domain-specific reasons on her own *that are not shared with the expert*. However, the debating experts do have such a reason since they can reasonably expect of one another to possess not yet fully shared evidence. If, on the other hand, all arguments and pieces of evidence have been fully disclosed and no further dynamics are missing, deference to the epistemic superior, if she is identifiable, seems fully appropriate.

Here is a second objection: The argument that is intended to support general preemption does not really succeed since it supports non-aggregation rather than preemption. What is really prohibited is that we use our own domain-specific reasons over and above the reasons provided by the expert testimony. Now, consider a situation in which the layperson's judgment that is based on domain-specific evidence is confirmed by the expert judgment. In that situation, the layperson may argue that she has good reason to trust her own judgment. Since aggregation is prohibited, she may not use the expert judgment as additional evidence. Hence, when the layperson agrees with the expert, non-aggregation can be implemented without preempting the layperson's use of her own domain-specific reasons.<sup>34</sup>

Although the situation looks symmetrical when the expert agrees with the layperson, it is not. For the layperson, it is much safer to rely on the expert judgment than on her own domain-specific reasoning, even in cases of agreement. This is so because even if the layperson correctly assesses the evidence she may do so accidentally or in an incompetent manner. Since the layperson takes the expert to be much more competent than herself, it is always safer for her to follow the expert's lead.

Finally, there is the worry that preemption and subsequent deference to the expert abstracts away from the temporal setting of a conversation between layperson and expert. Typically, the layperson will not only request ready-made verdicts about some issue but also will ask for further explanations in order to improve her understanding

<sup>33</sup>Thanks to Jason Kawall for coming up with this case.

<sup>34</sup>Thanks to Peter Brössel for raising this worry.

step by step.<sup>35</sup> One might even argue that the layperson should suspend judgment unless her judgment about the issue is based on firm understanding.

It is true that laypeople often do not only strive for correct answers to given questions but also demand understanding which involves, among other things, the ability to base judgments on domain-specific reasons. However, it is not true that one should not form a belief unless one has some degree of understanding. In the face of expert testimony, just the reverse is true. In a first step, the layperson just follows the expert's lead and is bracketing the rational use of her own domain-specific reasons. At stage two, the layperson uses her off-line abilities to simulate understanding of what she already believes on the basis of expert testimony. If her simulated domain-specific reasoning leads her regularly to what the expert has already told her, she has become an expert herself and can now base her judgments on her own understanding. Correctly interpreted, striving for understanding is an epistemic goal over and above forming true judgments about the world and does not exclude the temporal priority of the latter.

## 7. Conclusion

We started out with the worry that the authority view has no resources to reject outrageous expert beliefs. It has turned out that this is not generally true. In many cases, unpreempted, domain-independent reasons are available to the laypeople – reasons that justify the rejection of the target judgment. On the other hand, it must be admitted that such a rejection of outrageous expert judgments is not always a live option for the authority view. But that is not fatal to this view. The main reason for this is that appearing outrageous is not in itself a reason to reject the outrageous judgment. Rather, outrageousness is an epistemic feeling that is typically sensitive to reasons beyond itself. However, being typically sensitive to reasons does not imply that all these impressions of obvious falsehood are rational. Sometimes they result from biases such as the anti-expert bias. In those cases, rejecting outrageous judgments is not rational. Therefore, it is not a decisive objection to the authority view that it does not permit the rejection of outrageous judgments in 100% of the cases. Admittedly, when we accept outrageous expert judgments, they sometimes later turn out to be false. Nevertheless, such acceptance can be rational because rationality is not an infallible guide to truth. In a nutshell: The authority view can resist the challenge from outrageous expert judgments.<sup>36</sup>

## References

Alter A.L., Oppenheimer D.M. and Epley N. (2007). 'Overcoming Intuition: Metacognitive Difficulty Activates Analytic Reasoning.' *Journal of Experimental Psychology: General* **136**, 569–76.

<sup>35</sup>Thanks to Catarina Dutilh Novaes and Jason Kawall for addressing this question.

<sup>36</sup>Earlier versions of this paper were presented at the Cologne Brownbag Epistemology Series in May 2019, at the Episteme Conference 2019 in Skukuza, Mpumalanga, South Africa in June 2019, and at a workshop on Epistemic Conflicts at the University of Cologne in July 2019. I am particularly grateful to Jason Kawall for the written version of his carefully worked out and enormously helpful critical comments on my paper. Very helpful comments from and extensive discussions with the following philosophers enabled me to work out the final version of this paper: Sven Bernecker, Paul Boghossian, Sofia Bokros, Peter Brössel, Jan Constantin, Catarina Dutilh Novaes, Anna-Maria Eder, Alex Guerero, Paul Irikefe, Aleks Knocks, Steffen Koch, Jennifer Lackey, Maria Lasonen-Arnio, Jasper Lohmar, Kirk Ludwig, Giacomo Melis, Jean Moritz Müller, Giulia Napolitano, Laurie Paul, Jim Pryor, Luis Rosa, Paul Silva, Asbjørn Støglich-Petersen, Eyal Tal, Alex Worsnip. Many thanks to Jennifer Lackey for inviting me to the Episteme Conference. Finally, I would like to thank the German Research Foundation (DFG) for supporting my research on the topic of this paper as part of the project 'Disagreement in Philosophy. Semantic and Epistemic Foundations'.

- Boghossian P.** (2003). 'Blind Reasoning.' *The Aristotelian Society: Suppl.* 77, 225–48.
- BonJour L.** (1998). *In Defense of Pure Reason. A Rationalist Account of A Priori Justification.* Cambridge: Cambridge University Press.
- Bruns B. and Wieth M.** (2004). 'The Collider Principle in Causal Reasoning: Why the Monty Hall Dilemma is so Hard.' *Journal of Experiential Psychology: General* 133, 436–49.
- Casullo A.** (2003). *A Priori Justification.* Oxford: Oxford University Press.
- Christensen D.** (2007). 'Epistemology of Disagreement: The Good News.' *Philosophical Review* 116, 187–217.
- Constantin J. and Grundmann T.** (2020). 'Epistemic Authority: Preemption Through Source Sensitive Defeat.' *Synthese* 197, 4109–30. doi:10.1007/s11229-018-01923-x.
- Croce M.** (2019). 'On What it Takes to Be an Expert.' *Philosophical Quarterly* 69, 1–21.
- Dunbar D. and Reagan B.** (2006). *Debunking 9/11 Myths. Why Conspiracy Theories Can't Stand Up to the Facts.* New York, NY: Hearst Books.
- Dunning D.** (2005). *Self-Insight. Roadblocks and Detours on the Path to Knowing Thyself.* New York, NY: Taylor and Francis.
- Elga A.** (2007). 'Reflection and Disagreement.' *Noûs* 41, 478–502.
- Fricker E.** (2006). 'Testimony and Epistemic Authority.' In J. Lackey and E. Sosa (eds), *The Epistemology of Testimony*, pp. 225–50. Oxford: Oxford University Press.
- Goldberg S.** (2010). *Relying on Others.* Oxford: Oxford University Press.
- Goldman A.** (2001). 'Experts: Which Ones Should You Trust?' *Philosophy and Phenomenological Research* 63, 85–110.
- Goldman A.** (2018). 'Expertise.' *Topoi* 37, 3–10.
- Hookway C.** (2003). 'Affective States and Epistemic Immediacy.' *Metaphilosophy* 34, 78–96.
- Jäger C.** (2016). 'Epistemic Authority, Preemptive Reasons and Understanding.' *Episteme* 13, 167–85.
- Kelly T.** (2013). 'Evidence Can be Permissive.' In M. Steup and J. Turri (eds), *Contemporary Debates in Epistemology*, pp. 298–312. Oxford: Oxford University Press.
- Keren A.** (2007). 'Epistemic Authority, Testimony and the Transmission of Knowledge.' *Episteme* 4, 368–81.
- Kitcher P.** (2000). 'A Priori Knowledge Revisited.' In P. Boghossian and C. Peacocke (eds), *New Essays on the A Priori*, pp. 65–91. Oxford: Oxford University Press.
- Kruger J. and Dunning D.** (1999). 'Unskilled and Unaware of it.' *Journal of Personality and Social Psychology* 77, 1121–34.
- Lackey J.** (2018). 'Experts and Peer Disagreement.' In M. Benton, J. Hawthorne and D. Rabinowitz (eds), *Knowledge, Belief, and God: New Insights in Religious Epistemology*, pp. 228–45. Oxford: Oxford University Press.
- Landemore H.** (2013). *Democratic Reason. Politics, Collective Intelligence, and the Rule of the Many.* Princeton, NJ: Princeton University Press.
- Levy N.** (2017). 'Nudges in a Post-Truth World.' *Journal of Medical Ethics* 43, 495–500.
- Lynch M.** (2016). *The Internet of Us: Knowing More and Understanding Less in the Age of Big Data.* New York, NY: WW Norton.
- McCauley R.M.** (2011). *Why Religion is Natural and Science is Not.* Oxford: Oxford University Press.
- Raz J.** (1988). *The Morality of Freedom.* Oxford: Clarendon Press.
- Rosenhouse J.** (2009). *The Monty Hall Problem.* Oxford: Oxford University Press.
- Sosa E.** (2017). *Judgment and Agency.* Oxford: Oxford University Press.
- Stichter M.** (2015). 'Philosophical and Psychological Accounts of Expertise and Experts.' *Humana Mente: Journal of Philosophical Studies* 28, 105–28.
- Thompson V.A., Prowse Turner J.A. and Pennycook G.** (2011). 'Intuition, Reason, and Metacognition.' *Cognitive Psychology* 63, 107–40.
- Williamson T.** (2007). *Philosophy of Philosophy.* Oxford: Wiley-Blackwell.
- Zagzebski L.** (2012). *Epistemic Authority: A Theory of Trust, Authority, and Autonomy in Belief.* Oxford: Oxford University Press.

**Thomas Grundmann** is Professor of Philosophy at the University of Cologne, Germany. He has published numerous articles on general epistemology, applied epistemology, and philosophical methodology. Among his books are *Die philosophischen Wahrheitstheorien* (2018) and *Analytische Einführung in die Erkenntnistheorie* (2017). His current research addresses issues from social epistemology such as epistemic

authority and experts, the limits of critical thinking, epistemic autonomy, disagreement, and post-truth. He is the co-author of *The Epistemology of Fake News* (2021) and *Experimental Philosophy and Its Critics* (2012).

---

**Cite this article:** Grundmann T (2021). Preemptive Authority: The Challenge From Outrageous Expert Judgments. *Episteme* 18, 407–427. <https://doi.org/10.1017/epi.2021.30>