Epistemic Authority, Preemptive Reasons, and Understanding
Christoph Jäger
Dept. of Philosophy, University of Innsbruck
Email: christoph.jaeger@uibk.ac.at

To appear in *Episteme*
Revised final version, March 17, 2015

Abstract
One of the key tenets of Linda Zagzebski’s book *Epistemic Authority: A Theory of Trust, Authority, and Autonomy in Belief* (2012) is the “Preemption Thesis for epistemic authority”. It says that, when an agent realizes that an epistemic authority believes that p, the epistemically rational response for her is to adopt the authority’s belief and to replace all of her previous reasons relevant to whether p by the reason that the authority believes that p. I argue that such a “Hobbesian approach” to epistemic authority yields problematic results. This becomes especially virulent when we apply Preemption to cases in which the agent and the authority share their belief (same-belief cases), or in which both have either a positive or a negative graded doxastic attitude toward a given proposition. As an alternative I propose what I call a “Socratic approach”, according to which epistemic authorities will not only motivate us to adopt their beliefs, but also provide us with higher-order reasons for re-assigning our own considerations their proper place in the web of reasons for and against the view in question, thereby fostering our overall understanding of the topic.

1. Introduction

Much recent discussion in social epistemology has centered on rational disagreement among epistemic peers. Far less attention has been paid to rational engagement with epistemic authorities. An agent’s epistemic peers are her epistemic equals, i.e., people whose intelligence and information, evidence and epistemic skills with respect to some question match the agent’s own intellectual resources in the relevant respects. However, sometimes we encounter people we conscientiously judge to be our epistemic superiors, that is, people who tend to perform epistemically better than we do in a given domain. How should we react? What kinds of doxastic-attitude adjustment are rationally called for when we learn what an epistemic authority believes?
In her pioneering book *Epistemic Authority: A Theory of Trust, Authority, and Autonomy in Belief* (2012)¹ Linda Zegzebski explores these questions and presents a detailed account of the nature of, and rationality conditions for, adopting or adjusting beliefs on authority. By tackling this topic, she fills an important lacuna in contemporary social epistemology, and I learned much from studying her book. However, while Zagzebski offers many interesting and insightful analyses, one of her key claims yields problematic results. Or so I shall argue. Zagzebski advances the following “preemption thesis” (henceforth for short: *Preemption*) about epistemic reasons:

**Preemption**

“The fact that the authority has a belief *p* is a reason for me to believe *p* that replaces my other reasons relevant to believing *p* and is not simply added to them” (p. 107).

Obtaining a reason for believing something on authority, as Preemption says, need not amount to obtaining a reason strong enough to license believing what the authority believes. (I might think that the authority is more likely to obtain the truth about whether *p* than I am, but also think that she is not very likely to obtain the truth either.) However, in typical situations, Zagzebski thinks, such a reason actually suffices rationally to adopt the authority’s belief and to abandon all of one’s other reasons concerning the question. I do not wish to deny that Preemption fits certain kinds of epistemic context. Your car has broken down and you come up with a rough-and-ready theory about what’s wrong with it. Consulting your mechanic you find that her diagnosis differs substantially from your dilettante guess. You conscientiously judge the mechanic to be an authority in the field, give up your previous verdict, and discard your previous reasons as deficient. No complaint. Except that this is just one of many different kinds of situation in which we engage with epistemic authorities. There are many interesting and important kinds of epistemic context which differ substantially from this kind of case and, I argue, for which Zagzebski’s preemption view goes astray. Applied to scientific research, for example, or indeed philosophy itself, the view runs into trouble.

If Preemption holds for disagreement between the agent and the authority, it should also apply to agreement between them. But when they agree, it seems inappropriate for the agent to abandon all of his previous reasons for believing that *p* in order from now on to continue holding this belief solely for the reason that the authority shares it. Instead, the fact

¹ If not otherwise indicated, page references will be to this book. Another helpful recent investigation of the topic is McMyler (2011), who explores the relationship between epistemic authority and epistemic autonomy within the framework of the philosophy of testimony.
that the authority shares the agent’s belief is, at least *prima facie*, evidence for the agent that his own reasons are good ones. So why replace them altogether with another reason? Why abandon apparently good reasons? The rational move here for the agent seems rather to keep his own reasons and to add to them, in a cumulative case, the fact that the authority shares the belief.

As sketched so far, this argument has been silent about the authority’s reasons and their relation to the agent’s reasons. Yet I shall apply the argument both to cases in which (i) the agent’s reasons differ from the authority’s and in which they coincide, as well as to cases in which (ii) the agent is unaware of the authority’s reasons and to cases in which he is aware of them. (I shall assume that he is in any case aware of his own reasons.) Same-reasons cases in which the agent is aware that the authority also shares his reasons seem to present a particularly tough problem, indeed a paradox, for Preemption: The agent learns that an authority shares his belief, and that she holds it for the same reasons he does; yet, Preemption requires him to replace his reasons with the sole reason that the authority has that belief.

In section 2 I reconstruct Zagzebski’s view of epistemic authority and suggest some applications or embellishments, including that the approach be extended to graded belief. Section 3 discusses the abovementioned problems for Preemption in more detail, applying them to what I call (i) the switching problem, (ii) the problem of competing authorities, and (iii) the problem of unhinging proper bases. In section 4 I outline my alternative account of Socratic epistemic authority.

2. Zagzebski on epistemic authority

Zagzebski’s project has two main components. One – the focus of this paper – is an account of the epistemically appropriate response to discovering that an epistemic authority holds certain beliefs. The other – to be kept in mind as a background to the present discussion – is an account of the conditions under which it is epistemically appropriate to take someone to be an epistemic authority to begin with. Here Zagzebski brings in the notion of epistemically conscientious self-reflection. The core idea is that, in spite of certain skeptical problems pertaining to epistemic circularity, in normal circumstances it is rational for you to trust that your own cognitive faculties are reliable. Moreover, typically you can conscientiously believe that other normal human adults have the same natural desire for truth and the same intellectual capacities than you have (p. 55). Combined with the principle that one should treat like cases alike, Linda argues, this gives us first the claim that in ordinary circumstances one has a
prima facie reason to trust others as well. The relevant “Principle of Epistemic Trust in Others” says, roughly, that when “I am led to believe that others have the same [cognitive, C. J.] property I trust in myself (to the same degree as I have myself), I have a prima facie reason to trust them as much as myself” (p. 68).²

Now, we often have reason to believe that another person is more likely to form true beliefs about some question in a given domain than we would be if we attempted to figure things out ourselves (p. 105). In such cases, Zagzebski argues, “the conscientious thing to do is to let the other person stand in for me” and to adopt her beliefs, irrespective of what my previous reasons and beliefs were. This is what it means to accept another person as an epistemic authority. In general, according to Zagzebski,

“an epistemic authority is someone who does what I would do if I were more conscientious or better than I am at satisfying the aim of conscientiousness – getting the truth” (p. 109).

A conscientious agent is one who does her best to achieve her epistemic goals, which above all include forming and sustaining true beliefs and avoiding false ones. Lest this conception of conscientiousness as “doing one’s best” appear overly subjective, we may remind ourselves of Zagzebski’s prominent arguments elsewhere that there are epistemic virtues, which are abilities that reliably achieve epistemic ends.³ Since other aspects of virtue epistemology appear as a backdrop to Epistemic Authority, we may suppose that conscientiousness also involves the exercise of epistemic virtues when one does one’s best.

According to Zagzebski’s general characterizations, epistemic authority might structurally be characterized as a three-place relation between a given subject (or set of subjects) S, another subject (or set of subjects) S*, and some doxastic domain D. So if someone is an epistemic authority for you in some domain D, she need not be an epistemic authority for me in that domain; and if someone is an authority for you in D, she need not be an authority for you in some other domain D*. We may add that authority is also time-

² This argument raises a number of interesting questions about trust in others which I shall not pursue in this paper. In particular, it may be questioned whether my trust in my own epistemic faculties really does impose a general consistency requirement on me to trust that others who have similar faculties operating in similar situations are epistemically reliable as well if, given the problems of epistemic circularity, trust in my own faculties is only pragmatically or psychologically required. For a detailed examination of such arguments and various bridge principles that seem to be required in this context, see Fricker (2014). Fricker argues that analogy arguments from epistemic self-trust to epistemic trust in others must assume successful empirical cases for epistemic similarities between myself and others which in typical situations are either not to be had or will not establish what is needed.

relative. Since our epistemic skills and virtues and the amount of available evidence and information about a given topic vary over time, S* may be an authority for S at a given time $t_0$ but no longer play that role at some later time $t_1$, or *vice versa*.

All this is fine as far as it goes. However, as it stands, the approach is silent on how a subject should form or adjust graded beliefs when coming across an authority, who may herself hold graded beliefs on the topic in question. Since this is a very common situation – often we are confident, or learn that an epistemic superior is confident, to a certain degree that $p$ – I propose extending Zagzebski’s account to graded belief. Second, the fact that one *has* a sufficiently strong reason for believing that $p$, given by the fact that an authority believes (to degree D) that $p$ – henceforth for short: “has an authoritative reason for believing that $p$” – need not result in one’s actually believing $p$ *on the basis of* this reason. Reasons one has for believing something need not be reasons *for which* one actually believes it. I propose spelling out Zagzebski’s preemption view in terms of epistemic basing, such that when the subject conscientiously adopts the authority’s belief and “replaces” all of her own reasons relevant to the question, she makes (her awareness of) the fact that the authority believes that $p$ the reason on which she bases her belief, in some technical sense of epistemic basing. Both my remarks on graded belief and on epistemic basing are intended as constructive comments. However, they will also pave the way for some critical questions. In the remainder of this section I say a bit more about applying Linda’s ideas to graded belief; some questions concerning basing will be discussed in the next section.

Why should a theory of epistemic authority say something about how to adjust graded beliefs when encountering an authority (who may herself hold a graded belief about the topic at hand)? Consider a situation fitting the unqualified description “S believes that $p$”. Suppose that S comes across an epistemic authority S* who also believes that $p$, but suppose that S’s belief is fairly weak whereas the authority is highly confident that $p$. In terms of coarse-grained belief, both subjects may qualify as “believing that $p$”, in which case Preemption does not require S to change anything about his doxastic attitude itself. (Preemption only says that S must alter his reasons.) Yet S ought to adjust his attitude. A natural extension of the preemption view as stated by Zagzebski, I maintain, is to prescribe that S become more confident that $p$ while discarding his previous reasons for $p$. A tripartite distinction between belief, disbelief, and belief suspension is unable to capture the kind of belief revision that such cases call for.

Zagzebski could account for graded beliefs by saying that, if you realize that an epistemic authority has a doxastic attitude of degree D toward $p$, this fact gives you a reason
preemptively to adopt or sustain a doxastic attitude of that same degree D toward p. For example, she might follow the familiar practice of modelling degrees of belief in terms of credences or subjective probabilities. Maximal confidence (regarding it as certain) that p corresponds to assigning a subjective probability of 1 to p; minimal confidence or believing it to be impossible that p corresponds to assigning a probability of 0 to p. It is contentious how we should map coarse-grained belief onto subjective probabilities. But one way to do so – a way which admittedly works with high idealizations – is to count doxastic attitudes with a probability assignment of $0.5 < \Pr(p) \leq 1$ to p as believing, and any attitude between a probability assignment of $0 \leq \Pr(p) < 0.5$ as an instance of disbelieving p (which can be taken to be equivalent to believing $\neg p$ with a strength equivalent to $1 - \Pr(p)$). In this picture belief suspension may be identified with a probability assignment of 0.5. It is commonly granted that actual epistemic subjects will not normally assign exact probabilities to propositions or at least won’t be able to perfectly identify any probabilities they do assign. Hence working with confidence intervals may be more realistic. This is not the place to delve deeper into an analysis of graded belief. It is up to the preemptionist to come up with an appropriate model here. Suffice it to say that something along the above lines could provide us with a more fine-grained picture of what it means to adjust beliefs in light of the views of epistemic authorities. It also gives us a slightly more detailed picture of how initial disagreement and initial agreement between agent and authority may be described. Let us say that S, with his doxastic attitude A, disagrees with an epistemic authority S* regarding her doxastic attitude A*, if either (i) the former is a positive attitude falling within the realm of “believing” (i.e., it assigns a probability of $0.5 < \Pr(p) \leq 1$ to p), while the authority’s attitude is one of belief suspension or of disbelief (i.e., it assigns a probability of $0 \leq \Pr(p) \leq 0.5$ to p); or if (ii) the authority’s attitude is one of belief while the agent’s attitude is one of disbelief or of belief suspension. Moreover, let us say that S weakly agrees with an authority S* if both of their attitudes are either positive (and thus constitute belief) or negative (disbelief), while their degrees of confidence or the respective confidence intervals differ. They fully agree if the degrees of confidence or confidence intervals coincide. (Here I will also speak loosely of their attitudes agreeing or disagreeing.) Let this latter case also cover belief suspension. In the next section I argue, among other things, that even when there is only weak agreement, other things being equal the rational and conscientious thing for the agent to do upon learning what an authority believes will often not be to abandon his reasons for his own attitude. Although he may have to re-evaluate their epistemic force, because they are reasons for the same kind

---

4 Perhaps one could term this just as well “weak disagreement”, but let’s stick to the above terminology.
of attitude the authority has toward p, they will not generally qualify as epistemically obsolete.

3. Problems with preemption

I now present three specific problems for Preemption. (i) First, there is the problem of competing authorities. Suppose you have what you consider to be good reasons for believing that p. (Let p be some controversial philosophical or scientific claim, for example.) You also have reasons against p, but since your positive reasons outweigh your negative ones, it has thus far been rational (conscientious, epistemically apt) for you to adopt a positive doxastic attitude toward p. In a fresh attempt to put your views on a secure basis, you look once more at what others think about the topic and find that S, whom you conscientiously judge to be an epistemic authority in the field, at least weakly agrees with you that p. However, to your dismay you also find that S*, whom you conscientiously judge to be an epistemic authority in the field as well, disagrees with you. How should you react?

In such cases the definite description “the authority” in Preemption does not apply. Yet, situations of this stripe appear to be widespread. For example, think of trying to collect advice from several medical doctors, builders, mechanics, etc. Zagzebski herself mentions such scenarios. “There are also cases of competing authorities”, she writes,

“cases in which the authority is more likely to get the truth than I am, but so are several other persons. There is no reason for me to trust this particular authority more than others. A partial response to this situation is that an authority would be aware of competing authorities, in which case the authority’s own response to the presence of competitors is relevant to what I should do if I am conscientious” (p. 111).

I have two worries. First, as a general stipulation, the claim that the authority would be aware of competing authorities is not conceptually required by Zagzebski’s notion of epistemic authority. An authority need not even be aware of his or her own authoritative role for the subject. Thus, in order for the stipulation that the candidates for authority are aware of each other’s views not to be ad hoc, we need more of an argument.

Second, even if we assume for the sake of argument that there would be such awareness, Zagzebski’s “partial response” is, as far as I can see, left somewhat in limbo in the
book. In any case, I doubt that the idea points in the right direction. The problem is that often authorities, even if they are aware of each other, remain in stable disagreement. This might not hold for trivial kinds of empirical expertise (if two mechanics disagree about what’s wrong with your car, they will at some point discover who is right); but it does hold for fields such as philosophy, politics, religion, and so forth, which are very important to us. On Zagzebski’s model of preemptive authority, both in the case of mutual awareness as well as when your authorities are unaware of each other’s opinions, you seem to have three options when they remain in stable disagreement. Either you preempt your own reasons concerning whether p and adopt S’s view. Or you preempt your own reasons and adopt S*’s view. A third option would be to adopt neither view, but instead to take the authorities’ contradicting beliefs as neutralizing each other. Here too the spirit of preemptionism would seem to require suspending all of your own reasons (whatever this might amount to). However, all three alternatives seem unacceptable. If you preemptively adopt S’s position, you cannot at the same time rationally adopt S*’s position, and vice versa. In either case you would thus hold a belief that contradicts a belief held by someone you conscientiously judge to be an epistemic authority. On the other hand, if you preemptively suspend belief then you adopt an attitude that conflicts with the attitudes of both of your authorities.

The preemptionist might deny that the story is coherent and argue that what I called “authorities” cannot actually play that role in the example. However, recall that Zagzebski herself explicitly acknowledges that we will often encounter “competing authorities”. But suppose for the sake of argument that preemptionism were to rule out the possibility of competing authorities. In that case, the above reflections show either that at least for certain areas of inquiry, including philosophy itself, there is an important sense or kind of “epistemic authority” which the preemption approach does not capture. According to this other notion, cases like the one just sketched can occur, and they should motivate the subject neither to switch to one of the testifiers’ positions nor to suspend belief while preempting her own reasons, but to continue thinking about the issue and to (re)assess the total reasons relevant to the question. Alternatively, the argument may be taken to show that in domains such as philosophy there is no such thing as epistemic authority. However, I think that we do have a notion of authority which applies to such fields, and that ruling out epistemic authority for

---

5 But see Zagzebski’s discussion of disagreement in chapter 10 of Epistemic Authority, which does engage with related issues.
6 In fact this problem is not confined to two competing authorities. There may be more of them, and if we think in terms of graded belief their competing views may be spread over a wide variety of different doxastic attitudes toward a given question.
them would be a high price to pay. In section 4 I will sketch the direction in which I think such a different notion of epistemic authority may be explored.

(ii) A second problem is what we may call the *switching problem*. Suppose that you used to believe that p but one day come across a learned and intelligent friend or colleague who presents you with some novel arguments for not-p. Impressed with the quality of these arguments and with your colleague’s intelligence, methodological skills, breadth and depth of information, etc., you conscientiously come to regard him as an authority in the field. He disbelieves that p. Complying with Preemption, you rationally and conscientiously adopt his belief that not-p. However, suppose that a few weeks later you learn that your authority has changed his mind. To your surprise, he has switched to your previous position and come to believe that p. (Indeed, perhaps this has happened because he has read some of your work on the topic. An extreme case would be that it has left such a mark on him that he has come to regard you as an authority in the field. However, for present purposes we don’t need this twist.) Occasionally changing one’s views, you rightly judge, does not by itself undermine a person’s intellectual authority, so you rationally continue to regard your colleague as an authority. How should you react?

According to Preemption you ought to follow your authority and switch back to your prior position, yet this time your belief must not even in part be based on your earlier reasons. Since you ought to believe on authority, all of your previous reasons must remain inert and be substituted by the sole reason that the authority believes that p. But this is a counterintuitive result. Switching back and forth between two attitudes, for the sole reason that an authority alters his position, while discarding other reasons relevant to the question, does not appear to be the right kind of doxastic behavior. In fact, the case is still underdescribed. For example, it may be that the authority switches back to your belief that p, and for the same reasons you used to have for it. Let us label cases where reasons are shared *same-reasons* cases, and contrast them with *different-reasons* scenarios. (The reasons the agent and the authority have may or may not be good ones.) Two further conditions that combine with these cases are the agent’s being aware of the authority’s reasons, and his not being aware of them. Obeying Preemption would yield paradoxical results when applied to either situation. Even in same-reasons cases the agent must not reactivate his previous reasons, *even though the authority has adopted them*, and even if the agent learns about this. Preemption prescribes that the agent switch back for the sole reason that the authority has come to believe that p. This does not seem right. But nothing in Zagzebski’s preemption view, it seems, implies that such cases could not occur or that in such scenarios the (former) authority loses her status as an authority.
(iii) I used phrases such as believing something “for a reason”, “with a reason”, or “on the basis of some reason”. A third problem for Preemption arises if we spell out such phrases in more detail and consider cases in which the agent’s own original reasons for a belief are good ones. Thus, when the agent adopts the authority’s belief for the sole reason that the authority holds it, he may disconnect his belief from other good reasons. This problem is more general than the switching problem: Preemption may require you to cancel adequate epistemic basing relations between reasons and beliefs. Call this the problem of unhinging proper bases.

The epistemic basing relation is a matter of some controversy. For present purposes it will suffice to remind ourselves of some central topics of the debate and to work with some general ideas on which many parties agree. To begin with, if you have an epistemically adequate reason or ground for a belief, in the sense that it is an element of your noetic profile, this need not result in your actually forming the corresponding belief. Second, you may have an adequate ground and also hold the corresponding belief, but still not do so as a result of having the ground. A classic example is Keith Lehrer’s superstitious lawyer who has, and is aware of, epistemically good reasons for believing that his client is innocent, but who also consults the cards in order to discover the truth about the case. What actually makes him believe that his client is innocent is his faith in the cards. Lehrer argues that the lawyer’s belief is justified.

Following familiar practice let us say that a belief B is based on the ground (or set of grounds) G if G is the ground (set of grounds) for which B is held. Pace Lehrer, many authors have argued that this relation should be interpreted in causal terms. However, a belief’s being properly based, as we may say, requires more than the agent’s having reasons that cause him to hold it. Arguably, a doxastic condition must be fulfilled as well, to the effect that the subject also holds beliefs about the justifying force of these reasons. Note that all this is still compatible with the possibility that the reasons for which you believe p fail to be objectively adequate. You may base your belief on a reason which you mistakenly think is a

---

7 Of course one’s beliefs do not always rest on adequate bases. But sometimes they do, and the argument to follow is that in such cases unhinging beliefs from their reasons amounts to epistemic regression.

8 Sometimes the term “reason” is reserved for propositional justifiers, while “ground” serves as a more general term that also covers non-propositional justifiers such as experiences, perceptual inputs, perhaps emotions, etc. For present purposes not much hangs on this distinction, and I shall use the terms more or less interchangeably.


10 For a helpful discussion of the debate until the mid 1990s see Korcz (1997); a recent overview is Neta (2011), who discusses causal accounts sympathetically, but ultimately leaves it open whether he thinks they are correct.
good one. In light of these reflections I shall for present purposes adopt the following principle:

**Proper Basing**

S’s (graded) belief B is properly based on a given ground (or set of grounds) G iff (i) G is the ground for which S holds B and (ii) S has a true and rational belief to the effect that G sufficiently supports B.  

Equipped with these clarifications, consider Zagzebski’s analysis of epistemic authority. According to her official characterization, we recall, the fact that the authority believes that p is, under certain conditions, “a reason for me to believe p that replaces my other reasons relevant to believing p” (p. 107). How exactly are we to understand this? What Preemption demands, I suggest, is, not only that the agent incorporate the authority’s belief as a reason to believe that p into his doxastic system. In addition, what Zagzebski seems to have in mind is that the authority’s belief should typically constitute a reason for which one adopts, and on which one should base, that belief, in the technical sense of “epistemic basing” just outlined.

But this is still incomplete. Note, second, that Zagzebski emphasizes that the replacement requirement does not prescribe ignoring one’s previous reasons (see p. 113). What is required instead, she thinks, is that these reasons somehow “stop operating” (Zagzebski, personal communication). Here is how she explains her view in the book. “Let us look now”, she writes, “at what it means to take someone else’s belief as a preemptive reason to believe the same thing.”

“It might appear that this requires doing something psychologically impossible: ignoring our own reasons for and against the belief. But letting a reason preempt my other reasons does not require ignoring my other reasons. In fact, it is because I am not ignoring them that I see that the belief of the authority has a certain status vis-à-vis my other reasons. If I stop at the red light because that is what the law requires, I let the fact that the law requires it be my reason for stopping. I do not ignore the fact that I would prefer not to stop because I am in a hurry, or that I believe it is generally safer to stop, or that I do not want to take the chance of getting a high-priced traffic citation. But if I stop because the law says to do so, that reason has the status of being my

---

11 Note that this principle does not require B to be true.
reason for stopping. It can be the reason even though I am quite capable of reciting many reasons for and against stopping” (p. 113).

How should we interpret the idea that what is required is not ignoring my own reasons but still being able to recite them, yet letting the authoritative reason take over the status of being “my reason”? It is because of the authoritative reason, Zagzebski writes, that she stops at the red light. So this is the reason for which I act. It is natural to construe this practical case after the manner of the epistemic one: Preemption prescribes to cancel all basing relations between the subject’s activity and her own, non-authoritative, reasons that would otherwise operate. According to Zagzebski the same would apply to epistemic cases. Here too the agent ought to make (his awareness of) the fact that the authority believes that p the sole epistemic basis for his own belief that p, thereby unhinging this belief from all other reasons, including good ones.

However, this is an epistemologically strange result. To begin with, it is unclear whether it is even psychologically possible to sustain reasons or grounds on which one previously based one’s belief, keeping them in one’s mind and still acknowledging their epistemic significance, but unhinging them from their role as grounds for which one holds the belief. If I were asked to do this, I would not know how to proceed. And even if such cases were psychologically possible, are they realistic or phenomenologically plausible? For example, consider my perceptual belief that the bird in the tree is a red cardinal. Let this belief be partly based on a certain visual experience, i.e., on my seeming to see a red crested bird of a certain size, and so on. Suppose that since I am not entirely confident whether it is a red cardinal I ask an ornithologist who happens to be around (and whom I accept as an epistemic authority) for his opinion. He confirms that the bird is a red cardinal. According to Preemption, I would have to sustain my belief that the bird is a red cardinal, but I must stop basing it even in part on the visual experience which previously factored among my justifying grounds. I do not think that this is how things actually work in such cases. And I would not know how, while still staring at the bird, to make my perceptual ground an epistemically inert and impotent element of my noetic system.

Moreover, switching from psychological to normative considerations, wouldn’t the agent be epistemically better off if he based his belief both on the fact that the authority shares his belief and on his own reasons (or set of reasons)? Basing a belief on two good reasons is better than basing it on one. Consider Zagzebski’s practical case. My belief that the law dictates stopping at the red light may be a good reason for doing so. But an even better reason
is the conjunction of my beliefs that the law prescribes stopping and that stopping is safer for
the driver and for other travelers and that by stopping one avoids risking a ticket, and so on. In
epistemic cases the situation is analogous. The fact that an authority believes that p is a good
reason to hold that belief. But if you also have good independent reasons for believing that p,
then basing your belief on these other reasons as well will put you in an even better epistemic
position. An epistemic authority, though by definition generally more successful in attaining
the truth, is not infallible. Hence there is no a priori certainty that her reasons are conclusive;
the case for p can usually be strengthened by supplementing the authoritative reason with
other good reasons.

In general, unhinging beliefs from good reasons will constitute not epistemic progress
but epistemic regress. Consider the following variant of our superstitious-lawyer example.
Suppose that on the second day of the trial the lawyer is cured of his faith in the cards and
comes to base his belief in his client’s innocence solely on the epistemically proper reasons he
has. He has improved his epistemic situation. However, suppose that on the third day of the
trial he talks to an authority who turns out to share his belief in the client’s innocence. (For
element, the lawyer may talk to an absolutely reliable witness who knows that the client is
innocent.) Again, Preemption demands that the lawyer sustain his belief but change its basis.
Granted, he need not, in Zagzebski’s words, ignore his previous grounds; yet he should
disempower them and dismiss them from their role as reasons for which he believes that his
client is innocent. Once more, I submit, this is a counterintuitive and epistemologically
unattractive result. The lawyer would renounce the fruit of an epistemic achievement. 12

Such examples suggest a more general moral. If belief on authority is appropriate here,
then Preemption does not seem to capture what relating to epistemic authorities often looks

---

12 Here is a somewhat more detailed picture of epistemic progress in terms of basing. The two worst
positions are case 1 and case 2: 1. S believes that p, but bases this belief on no grounds at all.
Moreover, S does not even have any good grounds for the belief, in the sense that such grounds form
part of S’s noetic system. 2. S believes that p on the basis of epistemically bad grounds. Moreover, S
does not even have any good grounds for the belief. (Example: S believes that p solely because of
wishful thinking.) Arguably, the order of these first two positions in terms epistemic failure is not
clear. Which one is worse? However we may want to answer, I suggest that with each of the following
steps the subject makes epistemic progress: 3. The grounds for which S believes that p are
epistemically deficient, but at least S has adequate grounds in the sense that they stay S’s mind.
However, S is not aware of their epistemic significance. 4. The grounds for which S believes that p are
epistemically deficient. But S does have adequate grounds and is aware of their epistemic significance.
(This is what happens with Lehrer’s superstitious lawyer.) 5. The set of grounds for which S believes
that p contains both epistemically deficient and epistemically adequate grounds. S is aware of the
latter’s epistemic significance. 6. The grounds for which S believes that p are constituted exclusively
by epistemically good grounds whose epistemic significance S is aware of. This list describes some
salient cases, but I am not claiming that it is exhaustive. (For example, yet another case may be that
what causes S to believe that p is constituted both by epistemically deficient and epistemically
adequate grounds, while S is not aware of the latter’s justifying force.)
like. Consider once more the cardinal scenario. For example, the authority will have similar perceptual grounds for judging that the bird is a red cardinal. (He, too, is appeared to cardinalishly, as some would say.) More precisely, here the agent can rationally assume, or even knows, that the expert-authority, although being an epistemic superior in the domain, shares not only the agent’s belief but also some of his grounds or reasons. But this means that the appropriateness of these reasons is confirmed and that the agent, rather than acquiring a reason to preempt them, acquires a reason to keep them epistemically active. An analogous point arguably holds for scenarios with exclusively propositional reasons. In general, if an authority shares some or even all of your reasons, the rational thing to do is to stick to them. Same-belief-same-reasons cases appear to be especially clear examples in which Preemption fails.\(^\text{13}\)

So far, these reflections were cast in terms of what I called “full agreement”. But a similar argument applies to weak agreement in terms of graded belief. Suppose that S thinks that p is more likely than not-p but is nonetheless not very certain that p. S bases his graded weak belief on reason r. Suppose that S encounters someone, S*, whom he conscientiously judges to be an epistemic authority in the domain to which p belongs. Moreover, S learns that S* is absolutely certain that p, and that S* bases his firm belief on the very same reason r which S has for his weak belief. Even then Preemption rationally requires S to cease basing his belief on r and from now on to believe that p, presumably as firmly as S* does, solely for the reason that S* firmly believes that p. But this is an epistemologically strange result. In the present case, rather than acquiring a reason for dropping r, S will typically acquire a reason to attribute even more epistemic weight to r than before.

I conclude this section with one final observation in support of my argument as laid out so far. A key argument which Zagzebski advances for Preemption is that, if you accept someone’s belief as authoritative but still weigh your own reasons against the fact that the authority has that belief, this will worsen your “track record in getting the truth” (p. 114). Applied to graded belief and weak agreement this argument amounts to the following: If you employ your own reasons for believing that p to degree D, while being aware that an authority believes that p to degree D*, and D and D* differ, then if you don’t adopt the authority’s doxastic attitude you worsen your track record of forming and sustaining appropriate doxastic attitudes. Zagzebski takes her cue from Joseph Raz, claiming that his arguments concerning practical authority apply, mutatis mutandis, to epistemic authority. Raz argues that “if …

\(^\text{13}\) This argument may gain even more bite if we consider the fact that typically the agent may assume that the authority, precisely in virtue of playing that role, does share his, the agent’s, reasons. An epistemic authority will be likely to be aware of good reasons and to put them to work.
there is no other relevant information available then we can expect that in the cases in which I endorse the authority’s judgment my rate of mistakes declines and equals that of the authority.” However:

“In the cases in which even now I contradict the authority’s judgment the rate of my mistakes remains unchanged, i. e., greater than that of the authority. This shows that only by allowing the authority’s judgment to pre-empt mine altogether will I succeed in improving my performance and bringing it to the level of the authority” (Raz 1988, p. 68).

In order for this thesis to be applicable to graded belief and weak agreement let us, instead of talking about “contradicting the authority”, say more generally that if a subject’s degree of belief differs from that of the authority, these attitudes conflict with each other. However, is the Raz-Zagzebski argument sound?

It is true that adopting the authority’s (degree of) belief will probably improve my track record of forming and sustaining true beliefs and avoiding false ones. But how does this support the view that I ought to preempt my previous judgment or attitude, thereby taking my original reasons out of operation? In a helpful discussion of Zagzebski’s preemption view Katherine Dormandy argues for what she calls a total-reasons view (Dormandy 2015). She claims that the Raz-Zagzebski argument is fine so far as it goes, i. e., that it does show that “believing in accord with Preemption raises the agent’s probability of believing truly” (p. 6). Yet, she argues, basing one’s beliefs on all of the reasons one has supporting them is an epistemic good that will be missed out if one preempts. It is not clear to me that basing one’s beliefs on all of the reasons one has supporting them is an independent epistemic good, i. e., a good that can conceptually be separated from epistemic goods deriving from the truth goal. In any case, suppose that an agent knows that the proportion of true beliefs delivered by a certain belief-forming method she typically employs in a given domain is 60%. She also knows that a certain authority lands true conclusions 90% of the time. Does this example suggest Preemption? I do not think so.

My awareness that the authority believes that p has feedback effects on how I rationally evaluate my own reasons. If I believe that the authority believes that p and that she is right 90% of the time, then this gives me a reason to think that, although the method which originally led me to believe that p only delivers true beliefs 60% of the time, this time I was successful. The fact that the authority is highly confident that p tells me something about this
particular application of my epistemic method: It provides a reason to think that, although generally there is a 40% chance that using my method leads me to a false conclusion, this instance falls within the 60% range in which things go well. If that is so, however, I can and should rationally include my own original reason I had for my doxastic attitude among the reasons on the basis of which I am now 90% certain that p.

4. Socratic Authority and understanding

When the subject and the authority initially disagree, matters are slightly more complicated. Suppose that you initially believe that not-p but, following an authority, come to believe that p. In such cases too, I submit, there is a sense in which your original reason for believing that not-p has a role to play in explaining and justifying your new belief that p. Mulling over such cases, however, points to a notion of authority that differs from Zagzebski’s target concept – or at least to certain aspects of epistemic authority that Preemption fails to capture. According to the notion I have in mind, an epistemic authority for a given subject is someone who not only succeeds more often in attaining the truth, but who also is able to foster the subject’s overall insight into the problem under consideration. This kind of authority, I maintain, plays a significant role in many important areas of intellectual activity, including philosophy itself. I call this kind of epistemic or intellectual authority Socratic authority. If someone is a Socratic authority for a subject in a given domain, he not only has a higher ratio of true to false beliefs in the domain than the subject does. He also displays superior methodological skills and insights which enable him properly to assess evidence, reasons, methods of thinking and investigation, and so on, and to communicate such insights to others. Engaging with a Socratic authority will thus typically not only (rationally) motivate an epistemic agent to adjust his beliefs if they did not match the authority’s beliefs; it will also enable him to see what was right and what was wrong with his own grounds and reasoning methods, and what is right about the grounds and methods favored by the authority. A Socratic authority, in other words, serves not only as a source of maximizing true beliefs at the object level, but also as a source of understanding.

The past decade or so has seen growing epistemological interest in understanding. Zagzebski herself has helped pioneer this territory, and one of her key epistemological concerns has always been to broaden traditional perspective on epistemic goals and virtues.14

My proposal to explore a concept of epistemic authority which integrates the aim of understanding should thus chime well with Linda’s general epistemological interests. However, I think that – especially from the perspective of her own concerns – the preemption view, put forth as a general account of epistemic authority, is too narrow. As the problems discussed in the preceding section show, the account overgeneralizes and neglects an important dimension of what epistemic authority typically amounts to, particularly when it is doing the heaviest epistemological lifting.

A philosophical attempt to understand understanding will have at least three main aims: It will (i) distinguish important types of understanding; (ii) provide insight into their nature; and (iii) explain why we value understanding or why, indeed, we often tend to value understanding a topic more than having true beliefs or even knowledge about it.\(^{15}\) I will not say much about (iii) but will aim to demarcate the kind of understanding at issue in epistemic authority and outline its most important features. The objects and kinds of understanding are manifold. The kind of understanding I want to focus on resides in the vicinity of what Elgin calls “the understanding of a topic” or “subject matter”, or what Zagzbeski picks out as understanding a “theoretical framework”.\(^{16}\) Many issues about understanding are controversial. For example, it is contentious how to conceive of the relation between understanding and knowledge, whether understanding is factive, whether it is essentially propositional, or whether it is transparent in the sense that it entails understanding that one understands. For present purposes I shall not take sides in these controversies. However, I will assume that understanding has two general features about which there is widespread consensus. First, the kind of understanding at issue is not an all-or-nothing affair, but admits of degree. Second, it involves something like grasping systematic connections among elements of a complex whole, or gaining insight into certain relations between items within a larger body of information.\(^{17}\)

The dependence relations I am concerned with are support relations between reasons and doxastic attitudes. Understanding a theoretical subject matter such as an abstract scientific

---

\(^{15}\) Cf. Kvanvig (forthcoming).


\(^{17}\) For example, according to Zagzebski (2009, p. 144; cf. 2001, p. 241) “understanding …involves grasping relations of parts to other parts and perhaps the relation of parts to a whole”. Kvanvig (2003, pp. 192–93) writes that “understanding requires … an internal grasping or appreciation of how the various elements in a body of information are related to each other in terms of explanatory, logical, probabilistic and other kinds of relations”. Elgin (2007, p. 35) sees it as “primarily a cognitive relation to a fairly comprehensive, coherent body of information”. Grimm (2011, p. 86) argues that “the object of understanding … can profitably be viewed as a kind of system or structure – something … that has parts or elements that depend upon one another in various ways”. Greco (2012, p. 123) characterizes understanding as “knowledge of a system of dependence relations”.

or philosophical question involves appreciating the relative weight of reasons relevant to the possible positions on this question. Somewhat more precisely, the degree to which a subject understands a topic increases with the degree of her awareness of the epistemic weight of the reasons pertaining to the question. Accordingly, I suggest the following principle:

**Weight-of-Reasons Principle of Understanding**

The degree to which S understands a subject matter is proportional to S’s awareness of the relative epistemic weight of the total available reasons relevant to propositions belonging to that subject matter.

Consider some examples. Suppose I believe, with a certain degree of confidence, that the last dinosaurs became extinct some 66 million years ago and that this happened primarily because of massive volcanic eruptions and their ensuing effects. Then I consult an expert geologist and paleontologist whom I conscientiously believe to be an authority on the question and learn that he believes that the main factor accounting for the extinction is that a large asteroid hit the earth. Suppose he is right. Then, if I switch to the authority’s belief, I improve my track record of believing truths. But merely switching beliefs will not give me much additional understanding. Arguably, even in that situation I am in a position to improve my understanding somewhat, for the fact that the authority rejects the volcano hypothesis allows me to infer that my reasons in favor of it don’t pass muster, or that they are outweighed or undermined by competing reasons. However, I clearly gain more understanding if I also learn what the authority’s positive reasons for the asteroid hypothesis are, and I would improve my understanding even further if I also learned why my previous reasons for favoring a rival hypothesis are deficient, why other rival hypotheses are explanatorily inferior as well, and so on. In general, in such cases there is typically a *web of reasons* which is constitutive of the subject matter, and gaining insight into the epistemic dependence relations between these reasons and the degree to which they support certain propositions gives us understanding. This kind of understanding is typically one of our main rational aims when we encounter, and exchange views with, what I called a Socratic authority.

Or consider an adaptation of Kvanvig’s much discussed Comanche example.\(^\text{18}\) Suppose that, due to my scant knowledge of Native American history – maybe my main source are James Fenimore Cooper’s *Leatherstocking Tales* – I believe that the Apaches dominated the southern plains (say, for some 100 years starting in the 1770s). I then learn

---

\(^{18}\) Kvanvig (2003), pp. 197-198.
from an authoritative source that it was the Comanches. The rational and conscientious thing for me to do is to switch to this latter belief. However, if I want to deepen my understanding of Native American history, then I also want to learn (for example) why it was the Comanches rather than the Apaches, what explains the Comanches’ dominance and, *vice versa*, what other phenomena their dominance explains, etc. Here too my old reasons factor into the epistemic base of my new attitude (i.e., my belief that the Comanches dominated the southern plains) in the sense that I assign them their proper place in the web of reasons relevant to the issue.

Finally, consider an example from philosophy. Suppose an authority has convinced me that, contrary to my earlier beliefs, free will is compatible with determinism. According to the Weight-of-Reasons Principle, I understand the free-will problem to the extent that I am aware of the connections between, and quality and relative weight of, the available reasons for and against compatibilist accounts of free will and determinism. (For example, these reasons might include, on the pro-compatibilist side, Frankfurt-style arguments against the principle of alternate possibilities, Strawson-type arguments that reactive attitudes are independent of whether determinism is true, Lewis-style arguments for local-miracle compatibilism; and on the incompatibilist side the consequence argument, arguments for source incompatibilism, and so on). These examples indicate that, when we engage with epistemic authorities, at least with what I have labeled “Socratic authorities”, we often – and rationally – do much more than just adopt the authority’s belief. We assess the total reasons that are on the table, which includes re-evaluating our previous reasons for or against the belief under consideration.

Note that the notion of understanding at issue here coheres well with some general and widely endorsed features of the concept. In addition to accounting for the facts that understanding admits of degree and that it concerns dependence relations among items within a complex subject matter or body of information, the view of “understanding in the realm of reasons” I am outlining can also neatly account for other general features of understanding. Above all, these include (i) a “modal” property and (ii) the way in which understanding is connected to a subject’s ability to answer why-questions. Concerning (i), the proposal is in line with what Grimm calls the “modal sense” of that particular kind of “grasping a structure” that understanding involves. “Mentally to grasp how the different aspects of a system depend upon one another is to be able to anticipate how changes in one part of the system will lead (or fail to lead) to changes in another part”, he writes. Understanding is thus a capacity “to anticipate how certain elements of the system would behave, were other elements different” (Grimm 2011, p. 89). I think that Grimm is right here. Moreover, this kind of modal feature is
also present when we weigh reasons within a system of reasons and propositions supported by them: We understand the system to the extent that we grasp how changes in our evaluation of the epistemic force and conclusiveness of the reasons would recommend different doxastic attitudes.

(ii) Another important notion for investigation is understanding why. Typically this notion is applied at the object level. (Why did the house burn down? Why did the Comanches dominate the southern plains?) However, we can also ask at a meta-level why it is appropriate to hold a certain doxastic attitude toward a given proposition, or why it is more appropriate to adopt this particular attitude rather than that. One understands the problem or question whether p to the degree to which one’s doxastic attitude toward p approximates the doxastic attitude that is, in light of the total available reasons, epistemically appropriate.

One question we started with in this section was: In different-belief cases, in what way might the reasons for which I used to believe that not-p figure in the epistemic basis on which, after my exchange with the authority, I now believe that p? Suppose the epistemic basis of my belief that not-p consists of the reasons r1, r2, and r3. I learn that someone whom I conscientiously judge to be an authority in the field holds the contrary belief that p. As it happens, the authority convinces me that my reasons do not in fact support not-p (at least not to the degree I thought), or that they do support not-p but that there are stronger reasons r4, r5, and r6 – reasons I may or may not have considered yet – which support p. The rational thing to do for me is to adopt the belief that p. However, in such a situation the reasons for which I switch to believing that p, and which help me understand why it is rational to believe p rather than not-p, include not only r4, r5, and r6, but also my view – gained from my exchange with the authority – that r1, r2, and r3 do not sufficiently support not-p or are undermined by r4, r5, and r6. I now believe the contrary (with a certain degree of confidence) because I have learned that the reasons favoring not-p are weaker than those favoring p. Perhaps in such cases of initial disagreement the way in which my own prior reasons figure in the epistemic basis of my new belief is somewhat more complicated than in cases of agreement. However, even then my previous reasons factor into the basis of my new belief by being embedded in higher-order epistemic assessments of the form “r1 is not a good (enough) reason for

---

19 Cf. for example Pritchard (2010); Greco (2012). Pritchard (2010), p. 74, goes so far as to call this the paradigm usage of “understands” and contrasts it, following Brogaard (2005), with a “more holistic usage which applies to subject matters”. Grimm (2014) suggests construing understanding primarily as understanding of causes, but eventually concedes that a more general and perhaps less misleading claim might be that understanding consists of something like “knowledge of dependency relations”.
believing not-p”, “From r2 it follows that …, hence r2 does not in fact sufficiently support the belief that not-p”, and so on.

I conclude these observations by considering our question from the authority’s point of view. Zagzebski discusses her claims primarily from the (non-authoritative) agent’s perspective. “Since I am approaching the issue of authority from the point of view of the subject”, she writes, “I will assume that authority is a power that the subject can recognize and to which she can rationally respond. It is not necessary for my account that the authority sees herself as an authority for a particular subject” (p. 103). However, even though this assumption may not be necessary for Zagzebski’s dialectical goals, it will still be illuminating to look at contexts in which an authority is aware of her role. Arguably, the authority herself would typically not want her epistemic inferior simply to believe that p because she, the authority, believes that p. Instead, she would typically want the epistemic agent to re-examine his own reasons in light of what he learns from the authority, to re-assess their epistemic weight, and to give them their proper place in the space of reasons for and against a given proposition. At least, this is what a Socratic epistemic authority, who aims to promote the agent’s overall understanding, would want the agent to do.

5. Conclusion

In the constructive parts of this paper I have suggested that Zagzebski’s theory of epistemic authority be extended to graded belief, and that we can best interpret her preemption thesis in terms of epistemic basing. Adopting an authority’s doxastic attitude toward p and preempting all of one’s own prior reasons relevant to whether p – which, as Zagzebski insists, does not involve ignoring these other reasons – amounts to epistemically disempowering those reasons and unhinging them from the epistemic basis of the revised attitude. In the critical parts of this paper I have argued that, in various important cases that qualify as exchanges with an epistemic authority, this is not in fact the most conscientious and rational thing to do. Paradigm examples include those involving scientific hypotheses, morality, politics, religion, and last but not least philosophy itself. I ran through various kinds of case involving both coarse- and fine-grained belief: initial full agreement, initial weak agreement, and initial disagreement between an epistemic agent and the authority, and I outlined how the subject’s total reasons should figure in the basis of what she ought to believe when she learns what an authority believes.
Even if Zagzebski were to accept the main lines of my argument, one dialectical move open to her might be to reply that all of this just shows that there is no such thing as epistemic authority in fields such as the sciences, philosophy, and other domains of intellectual activity with similar features. However, in my view such a response would be phenomenologically implausible, and it would not sit well with Linda’s general views about where to apply the concept of epistemic authority. After all, she does apply it to religion and ethics, of all things. Moreover, even scientists and philosophers occasionally accept others as epistemic authorities, at least temporarily. If so, Preemption, if construed unrestrictedly as a thesis about epistemic authority in general, is to be rejected.

I conclude that all of this suggests that many important areas of intellectual activity are instead captured by a different concept of epistemic authority, namely by what I call “Socratic epistemic authority”. In contrast to a Hobbesian epistemic authority, a Socratic epistemic authority will not motivate us simply to preempt our previous reasons for or against a certain belief, but will help us understand the problem or subject matter connected with the belief and, if necessary, on the basis of a deepened understanding to correct our attitudes. The kind of understanding under consideration lies in the vicinity of “understanding a theoretical framework”. The aim is to understand why a certain doxastic attitude toward p, or a certain degree of confidence in p, is appropriate, given the total reasons relevant to believing that p.

The gist of all this, I take it, is not alien or hostile to Zagzebski’s general epistemological interests. By contrast, in recent decades she has herself been a leading figure among those epistemologists who have broadened our perspectives on epistemic goals and values, and she was among the first epistemologists who (re)discovered understanding as an epistemic goal. My suggestions may therefore be seen as an attempt to apply some of Zagzebski’s earlier insights to her present inspiring foray into important new territory in social epistemology.20

References


---

20 Portions of this paper have been presented at a book symposium on Zagzebski’s Epistemic Authority, organized by Nancy Snow, at the 88th annual meeting of the American Philosophical Association, Pacific Division, April 16–20, 2014, in San Diego, and at the conference “Neue Perspektiven der Erkenntnistheorie” (“New Perspectives in Epistemology”), organized by Pedro Schmechtig, May 23–25, 2014, at Dresden University. For helpful discussions I am indebted to the audiences, and especially to Katherine Dormandy, Thomas Grundmann, Benjamin McMyler, Linda Zagzebski, and an anonymous referee for Episteme.


