An Examination of an Argument Against the Autonomy of Epistemic Evaluation

Miguel Ángel Fernández
Instituto de Investigaciones Filosóficas
Universidad Nacional Autónoma de México
mafv@filosoficas.unam.mx

0. Much epistemological theorizing in the 20th century assumed that there is something distinctive to epistemic evaluation of beliefs and agents that distinguishes it from other forms in which beliefs and agents may be evaluated. It is usual to express the distinctiveness of epistemic evaluation in terms of the ultimate cognitive goal of getting the truth, here are some such expressions:

Epistemic evaluation is undertaken from what we might call the “epistemic point of view”. That point of view is defined by the aim of maximizing truth and minimizing falsity in a large body of beliefs.¹

…the distinguishing characteristic of this particular species of justification [i.e. epistemic justification] is, I submit, its essential or internal relationship to the cognitive goal of truth. Cognitive doings are epistemically justified, on this conception, only if and to the extent that they are aimed at this goal.... ²

In order for evaluations to be epistemic, it must be that the person’s beliefs are measured against the standard of how good a job they do of realizing his goal of now believing truths and now not believing falsehoods.³

These are mere slogans in need of clarification and refinement, but they express the conviction of many epistemologists that in order to be epistemic an evaluation must focus on some aspects of beliefs and believers and leave other aspects of them out of consideration. For example, typically, what evidence for the truth of p a subject has is treated as a consideration relevant for making an epistemic evaluation of the subject and his belief, but what moral emotions and aspirations the subject has are typically not regarded, per se, as relevant for making such evaluation. What features exactly of believers and beliefs are relevant for an epistemic evaluation is a question that a proper

¹ Alston 1985: 83.
² Bonjour 1978: 5.
refinement of the above slogans has to answer, but I think that many philosophers that subscribe to some version of them would feel confident in practice to discriminate what is relevant from what is not to epistemic evaluation.

In recent years some views have emerged that have challenged the thesis that some features of believers and beliefs, traditionally regarded as irrelevant for epistemic assessments, are really that irrelevant. It has been argued, for example, that the *practical interests* of a believer play a crucial role in determining whether his evidence is good enough in a certain situation to make him a knower.\(^4\) It has also been argued that the *moral motives and values* of a believer can contribute to make his belief good enough as to make it count as a case of knowledge.\(^5\) And practical interests and moral motives and values are precisely the sort of things that traditional approaches exclude from what is relevant for distinctively epistemic evaluation. We might say that those novel views challenge the *autonomy* of epistemic evaluation by arguing that it is not independent of features of beliefs and believers that have been traditionally regarded as relevant to make non-epistemic (e.g. pragmatic and moral) evaluations of them but irrelevant to evaluate them epistemically.

In this paper I will examine an argument against the autonomy of epistemic evaluation; it is due to Linda Zagzebski. She announces her argument with the claim that:

…..the common view that epistemic good is independent of moral good is largely an illusion.\(^6\)

And concludes it with the claim that

It is very unlikely that epistemic value in any interesting sense is autonomous [from moral value and the wider values of a good life]\(^7\)

My subject matter in this paper is what she means by these claims and what reasons she has for endorsing them.

Before getting to the matter I should add a caveat: I will present those of Zagzebski’s views that I want to criticize as constituting an *argument* against the autonomy of epistemic evaluation, but she does not present the argument *as such*. I

\(^{4}\) See for example Stanley 2005


\(^{6}\) Zagzebski 2003b: 12.

assemble the argument from premises she clearly holds; although the premises appear at different papers, they do hold up together as the argument I present.

1. As I understand it, Zagzebski’s strategy to show that epistemic evaluation is not autonomous from the moral values and motives a subject may have, begins by arguing that some crucial structural elements underlying moral evaluation are also present in the basic underlying structure of epistemic evaluation. She conceives the basic structure of moral evaluation as an assemblage of three nodes: an agent (with his motives, desires, inclinations, virtues), the acts of the agent and the intended outcomes of the agent’s acts:

\[
\begin{array}{ccc}
\text{Agent (motives)} & - & \text{Act} & - & \text{Intended outcome}
\end{array}
\]

Let’s call this the ‘Basic Structure of Moral Evaluation’ (BSME). In discussing BSME I am going to leave aside many details of Zagzebski’s theory of moral and epistemic evaluation. For example, in her theory the good motives on which an agent acts are the motivational component of virtuous traits of the agent; the full explanatory work of the goodness of an act is done by the virtues of the agent, not only by the motivational component of his virtues. Zagzebski provides an extremely detailed account of the nature of the virtues.\(^8\) I don’t think that leaving aside those details and simplifying her theory in the way just mentioned invalidates the points I will raise, for they concern matters of structure, not of content.

The above structure simply tells us that an agent acts out of certain motives and his acts, when successful, produce a certain intended outcome. As an illustration of the normative consequences that this structure has in Zagzebski’s theory consider an agent that acts out of the good motive of compassion (understood as the aversion to the suffering of others) and his act has the intended outcome of relieving someone’s suffering. Zagzebski thinks that such an act acquires good normative properties in virtue of being an act that arises out of a good motive; if it had a different origin it would not be as good as it is. Variations in the agent’s motives affect the goodness of the act. As evidence for this claim she describes three cases where the agent’s motives vary and so do, she claims, the goodness of his acts. The three cases are as follows: (a) acts that

\(^{8}\) Zagzebski 1996: 77-258
have the consequences of compassionate acts, i.e. the elimination of the suffering of others, without having it as an aim nor as a motive; (b) acts that have the elimination of the suffering of others as an aim, without being motivated by compassion; and (c) acts that have compassion as a motive. An act of type (a) could be talking next to someone who, unbeknownst to us, gets her suffering alleviated just by hearing a human voice; an act of type (b) could be talking to that same person with the intentional aim of alleviating her suffering, but in doing so one is motivated not by the desire to alleviate her suffering for the sake of it but, say, by the egoistic desire to be relieved from the nausea produced in us by the suffering person. And an act of type (c) could be talking to that person while being motivated to do so by the desire to alleviate her suffering for the sake of it. Zagzebski’s central claim is that, other things being equal, acts of type (c) are morally more valuable than acts of types (a) and (b), and that the explanation of this inequality is the motive present in (c) but absent in (a) and (b).

Although the goodness of an act is affected by the agent’s motives, these do not affect the goodness of the act’s outcome; if the outcome of the act is good, it is good independently of the agent’s motives. This is also illustrated in the above situations, where the agent’s motives vary and the goodness of the act’s outcome (i.e. the relief of suffering) remains constant.

Now, Zagzebski thinks that we can identify a similar basic structure in the epistemic evaluation of beliefs. It is natural to present such structure as exactly analogous to the moral one:

<table>
<thead>
<tr>
<th>Agent (motives)</th>
<th>Cognitive acts (acts of belief)</th>
<th>Intended outcome (State of true belief)</th>
</tr>
</thead>
</table>

Let’s call this the ‘Basic Structure of Epistemic Evaluation’ (BSEE). In this epistemic analogue of the moral structure, the goodness of cognitive acts is also supposed to be affected by the goodness of the agent’s motives. Zagzebski illustrates the point with a tripartite division of cognitive acts analogous to the division (a), (b) and (c) above. There are cognitive acts that (a*) have true belief as a consequence, without having it as an aim nor as a motive, e.g. as when we get true beliefs by accident; (b*) cognitive acts

---

9 Note that this type of act doesn’t fully fit into the BSME model, for the actual outcome of the act doesn’t happen to coincide with the intended outcome of the act. However, for Zagzebski’s comparative purposes what matters is the relation between the agent’s motive and his acts, rather than the relation of his acts and their consequences.

that have true belief as an aim, without having the motive of the love of truth, e.g. as when we what to know the truth not out of love of truth but only because what we would then believe would please us; and (c*) acts of belief that are motivated by the love of truth. Here Zagzebski’s claim is also that, other things being equal, cognitive acts of type (c*) are epistemically more valuable than cognitive acts of types (a*) and (b*), and that the explanation of this inequality is the motive present in (c*) but absent in (a*) and (b*). As she puts it: “I propose that love of truth is a motive that confers value on acts of belief in addition to any other value such acts might have”. This shows that the motive of love of truth “is capable of conferring additional value on the acts it motivates”.

Up to here the analogy of the epistemic structure with the moral structure seems to be perfect, but the analogy breaks down when we take into account the intended outcomes of the cognitive acts. Whereas the relief of someone’s suffering is external to the compassionate act that causes it, Zagzebski thinks that a true belief is not external to the cognitive act that causes it. She thinks that if states of true belief were external to cognitive acts then, exactly as it happens in the moral case, the goodness of the intended outcome, i.e. the state of true belief, would be insensitive to the goodness of the agent’s motives to act. But that would make the agent’s motives in general irrelevant to the evaluation of his beliefs, contrary to what she wants to maintain. Given that, according to her, the goodness of cognitive acts is affected by the goodness of the agent’s motives, she seems to think that in order to maintain that the goodness of the motives do also affect the goodness of states of belief, she needs to argue that such states are somehow more intimately connected with the cognitive acts that produce them, than the intended outcomes of moral acts are connected with moral acts. The idea seems to be that such a more intimate connection between states of belief and cognitive acts would enable the goodness of the agent’s motives to affect the goodness of the states of belief through affecting the cognitive acts that produce them.

Initially, Zagzebski argues that the more intimate connection between outcome states of true belief and cognitive acts is a special case of being a property of:

It’s true that a good outcome such as relief of suffering cannot be made better by the act that produced it, much less by the motive of the act that produced it, but

---

11 Ibidem.
12 Zagzebski 2003a: 149, my emphasis.
13 Ibid.: 147.
that’s because the outcome is a state of affairs separate from the act. In the case of acts of belief the intended outcome is a property of the act itself.\textsuperscript{14}

I will come back in a moment to the interpretation of this claim, now it is important only to note that the BSEE needs a crucial modification given Zagzebski’s conception of the relation of cognitive acts and the beliefs they produce:

\[
\begin{array}{ccc}
\text{Agent} & \text{Cognitive acts} & \text{Intended outcome} \\
\text{(motives)} & \text{(acts of belief)} & \text{(State of true belief)} \\
\end{array}
\]

Let’s call this modified structure ‘BSEE*’. The relation between states of true belief and cognitive acts is not a mere causal relation where the effect is external to the cause, it is a rather that the true belief is a property of the act itself. In Zagzebski’s view it is this relation between states of true belief and cognitive acts that makes the agent’s motives relevant for the goodness of his belief states.

\section*{2.}

So far we have reviewed only the abstract structure that according to Zagzebski makes it possible that an agent’s motives affect the goodness of his cognitive acts, and thereby the goodness of his states of belief. But this is only the first half of her argument against the autonomy of epistemic evaluation, the second half of her argument consist in trying to show that specifically moral motives figure prominently among the agent’s motives that can affect the goodness of his states of belief. I will expound briefly this complementary part of her argument.

In the example we described above to illustrate how the goodness of an agent’s cognitive acts is sensitive to the agent’s motives, Zagzebski chooses the motive of love of truth to construct the illustration; but according to her there are various other intellectual motives, each of which is the motivational component of a specific intellectual virtue, such as open-mindedness and intellectual fairness, and any of those specific motives could serve the same illustrative purposes. The reason why Zagzebski focuses on the motive of love of truth is the central place it occupies among all intellectual motives:

\textsuperscript{14} Zagzebski 2003a: 151. My emphasis.
....love of truth is plausibly the primary motive underlying a wide range of intellectual virtues. If love of truth is a good motive, it would add value to the intellectual acts it motivates.\textsuperscript{15}

Zagzebski wants to assert the antecedent of that conditional statement: that the motive of love of truth is a good motive, therefore it can add value to the cognitive acts that it motivates and thereby to the states of belief that result from them. But where does the goodness of the motive of love of truth come from? She doesn’t want to deny that love of truth has intrinsic value, but in her argument against the autonomy of epistemic evaluation the claim that bears more weight is that the motive of love of truth derives some of its value from the connections it has with the constellation of motives that are constitutive of a good life, or in her Aristotelian terminology, a life of eudaimonia,\textsuperscript{16} and central to that constellation there are various moral motives. Let’s look at a couple of ways in which according to her the value of the motive of love of truth derives from its connection with moral motives:

(1) When something of moral importance is at stake when we act, and that act depends on the truth of a particular belief, then it is morally important that that belief be true. Then, “the motive for true belief in such cases is motivated by the higher-order motive to be moral or live a good life”.\textsuperscript{17}

(2) True belief is typically a prerequisite for successful practical action; so, the motive to value truth is motivated by the motive to value those practical ends we pursue in acting, and this motive in turn is motivated by the desire to have a good life.

If this is correct, one way in which moral motives could affect the goodness of a state of belief would fit into the following pattern: an agent has many motives to act, some of them are moral, when the agent acts on a good moral motive and the success of his morally good action depends on finding out the truth concerning whether \( p \), the motivation of the agent to be morally good motivates him to value/love truth. In that case, the presence of the moral motive makes the motive of love of truth more valuable for the agent than it would have been if he hadn’t had the moral motive. Then, when the agent acts cognitively on his motive of love of truth and his cognitive act leads him to believe \( p \), his cognitive act and his belief that \( p \) are better than they would have been if

\textsuperscript{15}Zagzebski 2003b: 18
\textsuperscript{16}See for example her 2003b: 23-26.
\textsuperscript{17}Ibid.: 24.
he had not acted on that motive or if that motive had not been itself made more valuable by its connection with the agent’s moral motives. Hence, the moral motives of the agent can affect the goodness of his states of belief.

3. I want to make some critical comments on each of the two stages of Zagzebski’s argument against the autonomy of epistemic evaluation, beginning with the second stage.

Assuming that an agent’s motives, his cognitive acts and his states of belief are related BSEE* depicts them, does the existence of that structure of evaluation permit the goodness of the agent’s moral motives affect the goodness of his states of belief? Perhaps we could agree with Zagzebski that the structure in question allows the moral motives of the agent to affect some goodness of his states of belief. For example, maybe the true beliefs of an agent that are the result of his acting cognitively on the motive of love of truth, which in turn is motivated by the motive to be morally good, are more admirable than the true beliefs of an agent that lacks such motivational profile. If we assume that the admirability of an object is grounded on a good possessed by the object, then the fact that the true beliefs of the former agent are more admirable than the true beliefs of the latter would indicate that the former’s possess something good that the latter’s lack. But this would be threatening to the autonomy of epistemic evaluation only if the difference in goodness between the true beliefs of those agents can be seen as an epistemic difference, or a difference with epistemic import. The defender of autonomy does not deny that an agent’s moral motives can affect the moral evaluation of the agent’s beliefs as morally admirable and worth of being imitated. Zagzebski needs to give a reason to consider that the difference in value in an agent’s beliefs introduced by his acting on the right moral motives, and reflected in their admirability, is a difference in epistemic value and not merely in moral value.

In formulating this demand to Zagzebski we must avoid a risk of begging the question. Given that the aim of her argument is to show that in the mouth of the advocate of the autonomy of epistemic evaluation the terms ‘epistemically good’ and ‘morally good’ serve to draw an illusory dichotomy, in criticizing the second stage of her argument we must not assume that precisely that dichotomy exists. But aren’t we assuming precisely this in suggesting that although the difference in value in the agents’ beliefs is moral there’s no reason to think that it is epistemic? I think that in requesting for a reason to classify the described difference in value as epistemic we are not
assuming any controversial philosophical dichotomy, we are merely giving voice to some intuitions well entrenched in ordinary evaluative practice. If there is here a party making contentious philosophical assumptions, that party is Zagzebski’s, for cases can be described that show that there is a clash between the role that her theory assigns to moral motives in affecting the epistemic goodness of states of belief and standard evaluative practice. Ordinary practice classifies the motives that Zagzebski appeals to as deprived of epistemic import. Here is a basic example.

Imagine two scenarios. In the first, S desires to lead a good life and this desire motivates S to pursue certain courses of action and value certain traits of character, for example to be generous and donate money to charities. This motive, in turn, motivates S to learn the truth about certain subject matters, for example, whether some apparent charity representatives are honest and trustworthy. S undertakes a scrupulous investigation to verify the identity of the representatives and as a result he comes to believe that they are honest and trustworthy.

In the second scenario the motivational chain that leads to S’s scrupulous investigation is entirely different. S is a selfish and corrupted man who lacks any motivation to value generosity and the performance of generous acts. Nevertheless, he has a selfish interest in appearing to be generous, because that would give him a good name, so he decides to donate money to charities. This motivates him to learn the truth about certain subject matters, for example, whether some apparent charity representatives are honest and trustworthy. S undertakes a scrupulous investigation to verify the identity of the representatives and as a result he comes to believe that they are honest and trustworthy.

Zagzebski’s theory seems to predict that S’s scrupulously formed true belief in the first scenario has more epistemic value¹⁸ than S’s scrupulously formed true belief in the second scenario, just because of the different motivational origins of those beliefs. Nevertheless, standard practice classifies both beliefs as epistemically on a par, despite the different motivational chains leading to belief in each case. Although in ordinary practice we might not use the concept ‘epistemic’ to express this parity, we would surely treat both beliefs on a par from an epistemic point of view, for example we would bet as much on one belief as we would on the other. Therefore, standard practice treats the different motives exhibited in the motivational chains as epistemically irrelevant.

¹⁸ And even that it’s an instance of knowledge.
Again, the presence of the motives in the first scenario might add some value to the acts that they motivate, and even to the beliefs that result from those acts, for example, those acts and those beliefs might be more admirable in the first scenario than in the second. But we don’t have a reason to classify such value as epistemic; on the contrary, ordinary practice gives us a reason not to classify it in that way.

Zagzebski could reply that, in spite of our intentions, our description of ordinary evaluative practice is already contaminated by the false dichotomy she is fighting against. We should stop conceiving of moral motives idiosyncratically as irrelevant for the epistemic assessment of belief. But the intuitions that guide our judgment in the above scenarios seem well entrenched and not idiosyncratic at all. The intuitions in question are so entrenched that they seem to operate even in Zagzebski’s own description of her theory, for the value of the moral motives her theory regards as the source of some of the epistemic value of a cognitive act is conceptualized as non-epistemic by Zagzebski herself:

I propose that the higher-order motive to have a good life includes the motive to have certain other motives, including the motive to value truth in certain domains......... If knowledge is true belief credited to the agent because of its place in her motivational structure, it gets value not only from the truth motive but also from the higher-order motive that motivates the agent to value truth in some domain or on some occasion. And that motive has nothing to do with epistemic value in particular; it is a component of the motive to live a good life.\(^\text{19}\)

But the value in cognitive acts that she aims to account for is what makes for the epistemic difference between the value of acts that result in knowing and the value of those that result in merely believing with truth. So, her recognition that the value of the motives at the beginning of the motivational chains is a non-epistemic value makes it hard to understand how her explanation can succeed. For, why should we suppose then that the value her explanation accounts for is the value that she aims to explain? We have reason to suppose otherwise, since the kind of value in her explanans admittedly mismatches the kind of value in her explanandum.

4. I move now to the first stage of Zagzebski’s argument against the autonomy of epistemic evaluation. As we saw above, according to BSEE\(^*\) the state of belief is not

\(^{19}\) Zagzebski 2003b: 24, my emphasis.
merely an external product of a cognitive act, but a property of it. This thesis is essential for her further claim that the agent’s motives can affect the goodness of his states of belief. If an agent’s state of belief is conceived merely as a external product of his act then the goodness of the state of belief could not be affected by the agent’s motives, just as the goodness of the state of relieved suffering is not affected by the agent’s motives because it is an external product of the agent’s act. Why should we accept that an agent’s state of belief is a property of one of his cognitive acts, rather than an external consequence of it? Moreover, what does this claim exactly mean?

An initial difficulty in interpreting Zagzebski’s claims is that she does not explain what a cognitive act is, or as she sometimes says, an ‘act of belief’. As a first approximation one might think that a cognitive act, or a series of cognitive acts, begins when one sets for oneself the question whether \( p \) and thereby begins an enquiry to determine whether \( p \). And the cognitive act, or the series of cognitive acts, ends either when one settles for oneself the question whether \( p \) or when one abandons the enquiry. Zagzebski is interested in cognitive acts that terminate in the former way, in forming a belief as to whether \( p \). So conceived cognitive acts form a very heterogeneous class that comprises all sorts of acts, from acts that involve the movement of one’s body, such as preparing a mix in the lab or looking through the window, to what we might call purely mental acts, such as solving an equation or remembering something. Let us interpret ‘cognitive act’ in this broad manner.

What could it mean that a state of belief is a property of a cognitive act and not merely an external or separate effect of it? Whatever cognitive act we might want to pick up, it seems that it is essential to cognitive acts that they are processes, things that take time and consist in the succession of changes. Now, it is clear that a state of belief can be the result of a cognitive process, just as many states are the result of processes. But this structure is also present in the case of a moral act; for example, the state of relieved suffering is indeed the result of a process, namely, the action of the compassionate agent. What is the difference between this process and its result and a cognitive process and its resultant state of belief? The terms ‘external’ and ‘separate’ that Zagzebski uses suggests that the difference in which she might be thinking is that the state that results from the compassionate act lies beyond the physical location of the process that produces it, whereas the state of belief that results from the cognitive act lies within the physical location of the process that produces it. In effect, whereas the state of relieved suffering occurs beyond the boundaries of the agent’s body, the state of
belief occurs within the boundaries of the agent’s body. This would give a clear sense to the idea that the state of belief is a property co-located with the process that produced it, whereas the state of relieved suffering is merely an external or separate result of the process that produced it. But this interpretation cannot serve Zagzebski’s purposes, for the state of relieved suffering that results from a compassionate act can have its physical location within the physical location of the process that produces it. This happens when practicing self-care I execute a compassionate act towards myself and thereby relieve some of my own suffering. In such a case the state of relieved suffering occurs within the boundaries of the agent’s body (i.e. mine), and if the above interpretation of Zagzebski’s claims were correct this would mean that the state of relieved suffering would be, in that case, a property of the act that produced it and not merely an external result of it. But then the goodness of the state of relieved suffering would have to be sensitive to the agent’s motives, since that’s what Zagzebski says that happens when the state is not merely an external result of the process but a property of it. But surely this is not correct, if it is true that the state of relieved suffering in someone else is a good state regardless of my motives for bringing it about, the state of relieved suffering in myself is a good state regardless of my motives for bringing it about. The physical location of the state that results from an act doesn’t seem to make the normative structure to evaluate the state shift from one where the motives of the agent are not relevant to one where they are. We need a different interpretation of the claim that a state of belief is a property of the act that produces it and not merely a separate consequence of it.

On a more liberal interpretation of being a property of, it is correct to say that any cause has the property of having its specific effect. On this interpretation a state of relieved suffering is the property of the act that produced it, as much as a state of belief is a property of the cognitive act that produced it. The act of taking an aspirin has as one of its properties the relief of a headache, as much as the act of speaking to someone in need of hearing a human voice has as one of its properties the relief of the person’s suffering. On this interpretation the relief of suffering in both, others and myself, is a property of my compassionate act, which subverts even more than the previous interpretation the contrast that Zagzebski needs between the way cognitive acts are related with states of belief and the way moral acts are related with their outcomes. All states produced by acts are properties of the acts that produced them, just because any cause has the property of having its specific effect. This would mean that the motives of the agent always affect the goodness, not only of his acts, but also of his acts’ outcomes,
contrary to the structural difference that according to Zagzebski distinguishes belief states from states that are the outcomes of moral acts.

The previous paragraphs provide two interpretations of the mereological claim that a state that results from a process is a property of it: a restrictive interpretation that imposes some restrictions on the spatial location of a property relative to its possessor, and a liberal one that imposes no such restrictions. Neither of those interpretations validates Zagzebski’s metaphysical claim that a state of belief is a property of the cognitive act that produced it, whereas the outcome of a moral act is merely an external or separate consequence of the act. Given that this claim is essential to her further normative claim that the (moral) motives of the agent affect the goodness of his states of belief, we should not accept this alleged implication of the metaphysical claim until we are given a proper way to understand it that has exactly the normative consequences that Zagzebski thinks it has.
References


