

ON THE VERY IDEA OF UNDERCUTTING DEFEAT

Erhan DEMIRCIUGLU

ABSTRACT: My aim in this paper is to cast doubt on the idea of undercutting defeat by showing that it is beset by some serious problems. I examine a number of attempts to specify the conditions for undercutting defeat and find them to be defective. Absent further attempts, and on the basis of the considerations offered, I conclude that an adequate notion of undercutting defeat is lacking.

KEYWORDS: evidential defeat, undercutting defeat, rebutting defeat, evidential support, misleading evidence

1. Undercutting Defeat: “The Unpolished Conception”

In order for a given case to be a case of evidential defeat, the following conditions (or their respective instantiations)¹ must simultaneously obtain:

- (i) The subject S has a piece of evidence E that supports p .
- (ii) The subject S has a piece of evidence E' such that E and E' together do not support p .²

A widely recognized distinction, famously drawn by Pollock,³ is between two kinds of evidential defeat, viz. between rebutting and undercutting defeat. The distinction depends on two different ways of satisfying (ii). Rebutting defeat occurs when a subject has some evidence that supports a proposition and also has some other evidence that supports the negation of that proposition, in which case the subject's

¹ For convenience's sake, I sometimes drop this otherwise important qualification below and pretend that such variables as E and p are constants. This gives rise to an innocuous shifting between talk of truth and talk of satisfiability (e.g., between the truth of (i) and the satisfiability of (i)).

² Conditions (i) and (ii) together define what one might call “total defeat,” where E' *totally* defeats the support E provides to p . There is also what one might call “partial defeat,” where E *partly* defeats the support E provides to p . In what follows, nothing substantive hangs on the distinction between total and partial defeat and the discussion will proceed on the definition of (total) evidential defeat provided.

³ J. L. Pollock, *Contemporary Theories of Knowledge* (Totowa, NJ: Roman and Littlefield, 1986).

total evidence does not support the proposition in question. In a case of rebutting defeat, the following condition obtains:

(iii*) The subject has a piece of evidence E' that supports not- p .

It is the joint truth of (i) and (iii*), which one might call “counterbalancing,” that makes a given case a case of *rebutting* defeat. Furthermore, counterbalancing guarantees the truth of (ii), which combined with (i), entails that a given case is a case of *evidential* defeat.

Now, here is a paradigm case presented as involving undercutting defeat in the literature. Entering my friend’s reading room, I see some books that appear red. On this basis, I form the belief that there are red books in the room. My friend then informs me that the books are intricately illuminated by red light. Despite the fact that my friend’s testimony is *not* evidence that there are no red books in the room, my total evidence does not support the belief that there are red books in the room. In this purported case of undercutting defeat, my friend’s testimony “attacks”⁴ and “severs”⁵ the evidential connection between my experience as of seeing red books and the belief that there are red books (rather than “attacking [the belief] itself”⁶). So, given my friend’s testimony, my experience is not “an indication of the truth of [the belief].”⁷ The support my experience provides to the belief is destroyed, as it were, by my friend’s testimony. In a case of undercutting defeat, then, the following condition is suggested by the foregoing remarks:

(iii**) The subject has a piece of evidence E' that destroys the support E provides to p .

It is, it is claimed, the joint truth of (i) and (iii**), which one might call “destruction,” that makes a given case a case of *undercutting* defeat. Furthermore, destruction guarantees the truth of (ii), which combined with (i), entails that a given case is a case of *evidential* defeat.

The notion of undercutting defeat is underdeveloped because there is no account available in the literature of what it is to destroy a given evidential support (or, equivalently, what it is to sever an evidential connection). It is clear that what is meant by “destruction” (or “severing”) cannot simply be the joint truth of (i) and (ii) since that does not distinguish undercutting from rebutting defeat. Then, what

⁴ Pollock, *Contemporary Theories of Knowledge*, 196.

⁵ Thomas Kelly. “Evidence,” in E. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, 2016, URL: <https://plato.stanford.edu/archives/win2016/entries/evidence/>

⁶ Pollock, *Contemporary Theories of Knowledge*, 196.

⁷ Stewart Cohen, “Justification and Truth,” *Philosophical Studies* 46 (1984): 279–295, 290.

exactly does “destruction” mean here? A satisfactory answer to this question is curiously lacking.

However, there is a glaring and more serious, albeit as we shall see related, problem I would like to point out about the characterization above of undercutting defeat. The problem is that (i) and (iii**) cannot be simultaneously true: if E' destroys the support E provides to p , then E does not support p , and *vice versa*. So, if what is required for a case to be a case of undercutting defeat is the joint truth of (i) and (iii**), then there is no such thing as undercutting defeat. Undercutting defeat appears to demand an impossible trick to be accomplished: you cannot have your cake and eat it too, and you cannot have E supporting p and E' destroying the support E provides to p too. And, note that rebutting defeat, which requires counterbalancing but not destruction, faces no analogous threat.

Let me call the conception of undercutting defeat according to which undercutting defeat requires the joint truth of (i) and (iii**) “the unpolished conception.” I take it that the argument just offered conclusively shows that the unpolished conception of undercutting defeat is inconsistent. Of course, the intention here is not to attribute something as overtly problematic as the unpolished conception to such noteworthy advocates of undercutting defeat as Pollock and others. Still, my impression is that the unpolished conception is not entirely off the mark and it is at least suggested by some unguarded definitions of undercutting defeat. And, more significantly, the failure of the unpolished conception brings explicitly into view the task of providing an adequate conception of undercutting defeat, one in which something along the lines of destruction of evidential support plays a central role. Destruction or something like it is what distinguishes undercutting from rebutting defeat. However, the problem is that evidential support destruction seems to guarantee that a condition for evidential defeat (namely, (i)) is not satisfied. So, the challenge is this: how can there be a case of evidential defeat in which evidential support destruction plays a central role?

Can the challenge be met? Clearly, either (i) or (iii**) (or both) must go and be replaced by some other condition, in a way that respects the idea of evidential support destruction. In the next section, I will consider a readily available attempt along these lines and argue that it does not work either.

2. “The Diachronic Conception” Considered

According to what I shall call the “diachronic” conception, the following conditions must obtain for a given case in order for it to be a case of undercutting defeat:

- (i) At t_1 , S has a piece of evidence E that supports p .
- (iii) At a later time t_2 , S acquires a piece of evidence E' , which destroys the support

that E (previously) provided to p .

The diachronic conception holds that the duo of (ii) and (iii) is consistent.⁸

I want to make two points about the diachronic conception, one short and the other longer. As for the former, let us start with observing that the qualification “later” in (iii) is essential: the diachronic conception is consistent only if t_2 is not the same time as t_1 . More specifically, if t_1 and t_2 were the same time, then the diachronic conception would suffer from the same inconsistency problem that we have seen afflicts the unpolished conception. Now, and this is my short point, the diachronic conception cannot account for some cases that, if there is such a thing as undercutting defeat, should qualify as cases of undercutting defeat. Consider, for instance, a slightly revised version of “the red lighting case” presented in the previous section. In this version, let us assume, I acquire both pieces of evidence *at the same time*: the time when I see some books that appear red is the same as the time when my friend informs me that they are illuminated by red light. If there is such a thing as undercutting defeat, then this case is presumably an example of undercutting defeat. However, the diachronic conception entails that it is not (because (iii) does not obtain). So, the diachronic conception is inadequate.

The longer point I wish to make will take some time to develop. First, here is a note about how (iii) is to be understood. As it stands, (iii) is ambiguous between two different readings: is what destroys the (previous) support E provides to p , the evidence E' itself or the subject's acquisition of E' ? Let us reconsider the red lighting case. Is what is supposed to destroy the support my experience provides to the belief that there are red books around, my friend's testimony itself or my 'acquisition' (or 'awareness' (or whatever is required for possessing evidence)) of that testimony? It is clear that under those circumstances in which my friend's testimony is present while I am unaware of it (if, for instance, my friend is only engaged in soliloquy and says that the books are illuminated by red light 'all too silently'), the support my experience provides to the belief in question must remain intact. So, the support my experience provides to the belief is not destroyed when my friend says what he does but when I become aware of what he says. It is my *acquisition* of my friend's testimony, and not merely the testimony itself, that is supposed to have the destructive effect on the support my experience provides to the belief.⁹ The

⁸ The diachronic conception might also replace (ii) by (iii), which reads: at a later time t_2 , S acquires a piece of evidence E' such that E and E' together do not support p .

⁹ Of course, there must be another side to the story: if my friend's testimony did not have the *potential* to destroy the support in question, then my acquisition of it would not destroy it. One way to put the point is, then, this: what explains the fact that my acquisition of the testimony has the putative destructive effect is that the destructive potential of my friend's testimony is

diachronic conception assumes, then, what I will call “the destructibility thesis,” according to which a subject’s acquisition of a piece of evidence might destroy the support another piece of evidence provides to a proposition. If the destructibility thesis is false, then the diachronic conception fails.

In the remainder of this section, I will argue that the destructibility thesis is false. Before that, however, let me note why the destructibility thesis might appear to be true. I grant that the main idea behind the diachronic interpretation is intuitive: a subject can have E that supports p *until* a certain time (viz., the time when she acquires E' which destroys the support). Before the acquisition of E' , E supports p . After its acquisition, however, it *no longer* does. Here is an analogy. Think about “the support” at a certain time a wife provides to her husband. At a later time, the wife learns that he is cheating on her. After that time, she no longer supports her. We can say that the wife’s learning about her husband’s disloyalty “destroys” the support she previously provided to her. Similar considerations might appear to apply to evidential support relations: the support E provides to p might get destroyed when the subject acquires E' – or so it is held by the diachronic interpretation. Why not say, for instance, that my experience as of seeing red books in the room supports the belief that there are red books there *until* my friend informs me about the red lighting, and after that, it does not because my friend’s testimony *destroys* the previous support?

Whatever its intuitive appeal might be, however, the destructibility thesis is false. First, note that there is a good reason to be suspicious of the analogy between evidential support relations and “wifey” support relations. When the wife learns that her husband is cheating on her, she goes through a certain change (e.g., she now believes that his husband is cheating on her, she is disappointed and angry with him, and so on). This change explains why she *no longer* supports her. However, when a subject acquires E' , E does not go through any change. True, the subject herself goes through a certain change (more specifically, her total evidence changes); but, the relevant point is that E is the same as before (and this is the relevant point because the relevant question is whether E continues to support p after the subject acquires E' (and not whether her total evidence does)). And, if E supports p before the subject acquires E' , and also if E does not go through any change after the subject acquires E' , then how can E cease to support p after the subject acquires E' ? How can the subject’s acquisition of E' possibly destroy the support E provides to p , if E remains the same as before?

actualized by my acquisition of it. Despite this, however, the main point stands that it is, on the diachronic conception, my acquisition of the testimony that has the putative destructive effect.

Having noted a good reason to think that the analogy breaks down between evidential and wifey support relations, here is now a direct consideration against the destructibility thesis. It is clear that if E supports p , then E supports p whether I, or any other subject, has E . And, it is equally clear that if E does not support p , then E does not support p whether I, or any other subject, has E . That is, evidential support relations hold independently of how pieces of evidence are distributed among subjects, which I will call “distributional independence.” It seems clear that what a given piece of evidence supports does not depend on *who* has that evidence or even on whether *anyone* has that evidence or not. The question “does E support p ?” can be answered without raising the question “who has E ?” or “does anyone have E ?” And, note that rejecting distributional independence entails that by *redistributing* pieces of evidence among the subjects, we can change evidential support relations that hold between those pieces of evidence. And, since, as I believe many would agree, that cannot be right, distributional independence must be accepted.

However, distributional independence is violated by the destructibility thesis: if a subject’s *acquisition* of a piece of evidence can have a destructive effect on an evidential support relation, as the destructibility thesis says it can, then *contra* distributional independence, that relation is not independent of how evidence is distributed among subjects. Given distributional independence, then, the destructibility thesis is false and evidential support relations cannot be destroyed by the *acquisition* of further piece of evidence.

I maintain that the argument from distributional independence just rehearsed shows that the destructibility thesis must be rejected. However, it might be argued that the diachronic conception does not need to assume a thesis as strong as the destructibility thesis, that a weaker version of the destructibility thesis would do. According to what we might call “the weak-destructibility thesis,” a subject’s acquisition of a piece of evidence might destroy *for her* (whatever this might plausibly mean, and as opposed to *in general*) the support another piece of evidence provides to a proposition. Since the weak-destructibility thesis is consistent with distributional independence, the response goes, the diachronic conception resting on the former is not threatened by an argument from the latter.

I agree that the weak-indestructibility thesis is consistent with distributional independence: in particular, the idea that a given evidential support relation might be destroyed *for a subject* by her acquiring further pieces of evidence is consistent with the fact that evidential support relations hold *in general* independently of how pieces of evidence are distributed among subjects. However, the weak-indestructibility thesis and distributional independence together leads to an absurdity. Suppose that everyone acquires E' , in which case the evidential support

relation between E and p is *destroyed for everyone*. However, distributional independence entails that the evidential support relation between E and p goes undestroyed by the fact that everyone acquires E' and hence that it continues to *hold in general*. However, this is absurd: what does it mean to say that the evidential support relation between E and P is destroyed for everyone, despite the fact that it continues to hold in general? Since I believe distributional independence is true, the weak-destructibility thesis must be rejected in order to avoid the absurdity.

On the basis of the considerations above, I maintain that both the destructibility thesis and the weak-indestructibility thesis are false. And, if they are false, the diachronic conception that rests on one or the other fails.

I would like to conclude this section with two clarificatory remarks. First, my point against the destructibility thesis does not cast any doubt on rebutting defeat (or evidential defeat in general). This is because rebutting defeat does not require that evidential support relations be destroyable by the acquisition of further pieces of evidence. Second, I don't deny that there might be a sense in which evidential support relations might be destroyed. If, for instance, that E supports p is a contingent fact, then E might support p at a certain time and not support it at a different time. My point is, to emphasize, merely that evidential support relations cannot be destroyed by the acquisition of further pieces of evidence.

3. Other Conceptions Considered

We have seen that neither of the two conceptions – namely, the unpolished conception and the diachronic conception – that can be more or less directly gleaned from the literature on undercutting defeat works. I will now proceed to assessing some other possible conceptions of undercutting defeat, which might further be offered as revised versions of the unpolished conception.

Let us first recall the two conditions for undercutting defeat on the unpolished conception:

- (i) The subject has a piece of evidence E that supports p .
- (iii**) The subject has a piece of evidence E that destroys the support E provides to p .

The problem that afflicts the unpolished conception is, let's recall, inconsistency: (i) and (iii**) cannot simultaneously hold. In this section, I will consider a number of revisions to (i) and (iii**) and argue that none of the conceptions that ensue from those revisions works.

Here is, then, a revised version of (i):

- (i₂) The subject has a piece of evidence E that defeasibly supports p .

Does the attempt to characterize undercutting *defeat* in terms of (i₂) and (iii^{**}) work? No, and the problem is obvious: such an attempt is circular, just as the attempt to characterize evidential defeat in terms of (i₂) and (ii) would be circular. Furthermore, it seems clear that there is no need to appeal to the notion of defeasibility in our characterization of evidential defeat because (i) and (ii) appear to be just fine. And, a retreat to (i) and (ii) would take us back to square one.

Here are two further ways to revise (i):

- (i₃) The subject has a piece of evidence E such that that the subject does not have E' is necessary for E to support p .
- (i₄) The subject has a piece of evidence E such that that the subject does not have E' is sufficient for E to support p .

However, neither of these works. Condition (i₄) is evidently unsatisfiable because that a subject does not have E' cannot suffice for E to support p : there is no piece of evidence such that a sufficient condition for that evidence to support a proposition is that a subject having that evidence not have another piece of evidence. Surely, *something else* is needed for an evidential support relation to hold. And, the problem with (i₃) is that it does not entail that E supports p , and, the joint truth of (i₃) and (iii^{**}) leaves it open that E does not support p , in which case there is nothing to be evidentially defeated.

Furthermore, and more significantly, both (i₃) and (i₄) is beset by the same fundamental problem. Neither is true given distributional independence, viz. that evidential support relations hold independently of what pieces of evidence subjects happen to have. Different subjects might have different pieces of evidence at different times, and whether a piece of evidence supports a proposition is not a function of the contingencies that affect the distribution of evidence among subjects. If so, contra (i₃) and (i₄), that the subject does not have E' can neither be a necessary nor be a sufficient condition for E to support p . It follows that neither (i₃) nor (i₄) is true.

Now, let us turn to (iii^{**}) and replace it by the following:

- (iii₂) The subject has a piece of evidence E' that E does not support p .

It is clear that (i) and (iii₂) are jointly satisfiable. However, the problem is that if (i) and (iii₂) are true, then E' is misleading evidence regarding E : what E' says, as it were, regarding the evidential connection between E and p is false. If so, the subject has E that supports p and also E' that mistakenly says that E does not support p . Since E supports p and E' mistakenly says that E does not support p , it turns out that the subject's total evidence (viz. E and E') supports p (cf. If I win a race, and you

mistakenly say that I lose it, then I still win).¹⁰ However, if so, (ii) is not true. So, the joint truth of (i) and (iii₂) guarantees that what we thereby have cannot be an example of *evidential* defeat (and *a fortiori* cannot be an example of undercutting defeat).

Here is another way to modify (iii**), inspired by Feldman:¹¹

(iii₃) The subject has a piece of evidence E' that E does not support p in this case.

As it stands, however, (iii₃) is inconsistent with (i), on the plausible assumption that (i) purports to express a general truth. So, we also need to modify (i), and a straightforward way to do this is as follows:

(i₅) The subject has a piece of evidence E that supports p in this case.

It is clear that (i₅) and (iii₃) are jointly satisfiable. However, this attempt falls prey to the very same objection that afflicts the previous attempt. It is that if (i₅) and (iii₃) are true, then E' is misleading evidence regarding E in this case: what E' says, as it were, regarding the evidential connection in this case between E and p is false. If so, the subject has E that supports p in this case and also E' that mistakenly says that E does not support p in this case. Since E supports p in this case and E' mistakenly says that E does not support p in this case, it turns out that the subject's total evidence (viz. E and E') supports p in this case. However, if so, the accordingly revised version of (ii) (which reads: The subject S has a piece of evidence E' such that E and E' together do not support p in this case) is not true. So, the joint truth of (i₅) and (iii₃) guarantees that what we thereby have cannot be an example of *evidential* defeat (and *a fortiori* cannot be an example of undercutting defeat).

4. Conclusion

I have proposed and examined a number of attempts to specify the conditions for undercutting defeat and have shown that all these attempts fail. The unpolished conception suffers from internal inconsistency; the diachronic conception fails to account for some cases that must evidently count as cases of undercutting defeat, *if* there are any, and is inconsistent with distributional independence, i.e. the fact that evidential support relations are independent of what pieces of evidence subjects

¹⁰ It is true that the subject might be *rationaly* misled by E' to believe that E does not support p and thereby abandon the belief that p , but this is compatible with the point that her total evidence supports p . How can what rationality demands from a subject might come apart from what her evidence supports? For an answer, see, for instance, Christensen's (see "Higher-Order Evidence," *Philosophy and Phenomenological Research* 81 (2010): 185-215) "bracketing" account of higher-order defeat.

¹¹ Richard Feldman, "Respecting the Evidence," *Philosophical Perspectives* 19 (2005): 95-119, 113.

Erhan Demircioglu

happen to have. Furthermore, the other conceptions I have examined turn out to be either circular or inconsistent with distributional independence, or the conditions they propose fail to capture what is required for evidential defeat. Absent further conceptualizations, and on the basis of these considerations, I conclude that an adequate notion of undercutting defeat is lacking.