Article:
Elstein, D (2007) A new revisability paradox. Pacific Philosophical Quarterly, 88 (3). 308 - 318. ISSN 0279-0750

https://doi.org/10.1111/j.1468-0114.2007.00294.x

A NEW REVISABILITY PARADOX

Daniel Y. Elstein

This is the peer reviewed version of the following article: Elstein, D. Y. (2007), A New Revisability Paradox. Pacific Philosophical Quarterly, 88: 308–318, which has been published in final form at http://dx.doi.org/10.1111/j.1468-0114.2007.00294.x. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for self-archiving.

Abstract: In a recent article, Mark Colyvan has criticized Jerrold Katz's attempt to show that Quinean holism is self-refuting. Katz argued that a Quinean epistemology incorporating a principle of the universal revisability of beliefs would have to hold that that and other principles of the system were both revisable and unrevisable. Colyvan rejects Katz's argument for failing to take into account the logic of belief revision. But granting the terms of debate laid down by Colyvan, the universal revisability principle still commits Quineans to holding that one belief is both revisable and unrevisable: the belief that some beliefs are revisable.

1. The old 'paradox'

Quine (1980; 1960) popularized a radical holism, which held that all beliefs are in principle revisable in the face of experience. Quinean holism contains a principle of the universal revisability of beliefs (UR). It it tempting to employ the same diagonalisation strategy that is held to have refuted logical positivism in an attempt to refute holism: for UR is itself revisable, according to UR. There is something initially implausible about that. The argument which Quine gives for holism does not itself appeal to experience. So it would appear that if we know that Quinean holism is true, we must know it a priori. Quineans could respond, however, that since they hold that logic itself is revisable, no philosophical argument establishes its conclusion indefeasibly. So the argument for holism does depend on experience, and so Quineans are not committed to their knowing a priori that holism is true.

Katz (1998, pp. 72-74) has argued that UR does create a paradox. His idea is that any epistemological system which has UR as one of its principles must take those principles as both revisable and unrevisable. Katz sees Quine's epistemological system as having three principles, of which UR is one and the other two are a principle of non-contradiction (N) and a principle of simplicity (S). The revisability of the principles is an obvious consequence of UR itself; why must they also be unrevisable? Katz's thought is that any belief revision is a consequence of accepting an argument which has the epistemic principles as premises. So any argument for the revision of a principle has that principle as a premiss. Thus the argument is unsound: its conclusion contradicts one of its premises, and so it cannot have premises and a conclusion which are all true. But if there can be no sound argument for the revision of one of the principles of a system, then those principles are not revisable (or so Katz maintains).

Whilst Katz's argument is fascinating, I agree with Colyvan (2006) that he has not pinpointed a paradox, and so not refuted holism. But he does point in the direction of a genuine paradox for holism. In the remainder of this first section I will explain what is

wrong with Katz's argument. The second section will introduce my new paradox, and show how it escapes Colyvan's objections to the old paradox. The third section considers the impact of the new paradox on holism. I will suggest that the new paradox forces holism to be restricted in a way that may undermine much of its motivation.

We could challenge Katz's argument in various places. First consider his reasoning for the unsoundness of any argument for revising one of the principles. We might admit that the most basic form of reasoning for the revision of one of the principles involves taking that principle as a premise, but hold that this reasoning can be augmented as follows so as to eliminate the occurrence of the principle as an undischarged assumption. We start with an argument of the form P, … ⊨ ¬P, where P is the relevant principle. But we can convert this into a proof of ¬P from the other premises (not including P) by introducing P as a supposition, deducing ¬P within the scope of that supposition using the earlier argument, and then discharging the supposition and obtaining ¬P by ¬-introduction. A similar point is discussed by Colyvan (2006, p. 3). Now this maneuver won't always work, because it relies on non-contradiction and a bit more of classical logic besides. So it doesn't block Katz's argument against there being a sound argument for revising N, though it does seem to show that his reasoning does not apply to revising UR. This doesn't in itself resolve the paradox, since it's enough for Katz that a system containing both N and UR as principles is paradoxical.

Suppose we grant Katz that there cannot be a sound argument for revising N for the foregoing reasons. Does that mean that N cannot be revised? One way of avoiding Katz's conclusion, mentioned by Resnik and Orlandi (2003, p. 305), would be to say that what we are looking for is an argument against believing N, not an argument for N's falsity. Suppose that N is true; then there isn't anything to stop an argument with N as one of its premises and the advisability of not believing N as its conclusion from being sound. This is unconvincing, however, since there can't be good epistemic reasons for disbelieving something which aren't reasons for thinking it false (or less likely to be true).

A better objection to Katz is to note that in a framework where logical principles are up for grabs we have to be careful about what we use as our background logic. In particular, there is reason to doubt Katz's assumption that the way to discover whether a principle is revisable is to find out whether there is a sound argument for its revision. When Colyvan (2006, p. 5) considers this question, he points out that 'If there is no formal notion of logical consequence at all then Katz's argument simply fails, because his argument clearly requires some notion of logical consequence.' What Colyvan suggests is that we use some non-monotonic, belief-revision logic. But he goes on to explain why this does not help Katz:

> [S]uppose we're operating in some belief-revision logic. Then the truth or falsity of N is beside the point. What's important is whether N is rationally defensible. Now let's suppose that at time t_1, prior to making inference (1) [the inference from the prior belief set and principles to ¬N], N was rationally defensible, but at some later time t_2, after making the inference (1), N was not rationally defensible. Katz's challenge, I take it, is to provide a rational defense of ¬N at time t_2. (2006, pp. 5-6)

And as Colyvan says, this task isn't too hard, because after revising we have a belief set and epistemological principles which include ¬N (that was the revision!), so trivially there is an argument from our belief set and principles to ¬N.

Colyvan's argument seems correct as far as it goes, but there is a further point we need to consider. We have an answer to the question of how we can take ourselves to be justified in believing ¬N after the revision. But how do we take the revision itself to have been justified? The idea must be that the original argument for ¬N showed that the pre-revision principles were problematic – they included a rule against contradictions but had a consequence relation which allowed for the derivation of a contradiction – whereas the post-revision principles have no such problem. So we are now in a better epistemic situation than we were in before. It would be foolish to expect that in such a case there would be any explicit argument for the revision which we could endorse both before and after, because the revision involved a change in the consequence relation (insofar as we are considering only monotonic logics). As Resnik and Orlandi suggest (2003, p. 305), what would be a problem is if the revised principles prohibited the revision we had just made. But they do not: there is no reason to think that ¬N prohibits revision of N, or that ¬UR prohibits revision of UR.

Katz's argument fails because there is no plausible test of the revision of one of the principles that finds anything wrong with that revision. The revision is putatively legitimate because the ex ante position was unstable, the ex post position is stable (in that there is a trivial argument for the revised belief), and the revision is not prohibited either before or after it has taken place. Perhaps some reply is available to Katz's defenders, but rather than consider that issue in more detail, I will show how the arguments used against Katz point the way to a more plausible paradox.

2. The new paradox

We now have something like a recipe for paradox. Find a belief with the following property: revising that belief leads to a new belief set which prohibits that revision. The most obvious candidate is an existential revisability claim (ER) that some beliefs are in principle revisable in the face of experience. Consider what happens when ER is revised: the resulting belief set contains ¬ER, and thus prohibits all revisions. But then from the post-revision point of view one cannot take the revision to have been justified. And of course it is known in advance that any revision to ER will be without an ex post justification.

According to the Quinean, who holds UR, and thus holds ER to be revisable, experience could give us reason to believe that experience never gives us any reason to believe anything. Or alternatively: some experience could make it a reasonable option for me to start believing that no experience could ever make it a reasonable option for me to change my beliefs. But this is obviously impossible. Only by manifest self-deception could I come to think that some experience gave me reason to believe that no experience could ever give me reason to believe anything which I did not already believe. Such a thought is self-undermining. By reductio then, ER is unrevisable. But UR entails that ER is revisable, so the combination of UR and ER is paradoxical, and clearly UR is the one that has to go.

Note that the new paradox is simpler than Katz's in several respects. It relies on no distinction between principles and other beliefs: it is enough that a belief set contain both UR and ER for it to be paradoxical, regardless of whether either is a principle. The argument does not require any fixed consequence relation; rather it can be seen as relying on a form of reflective equilibrium which holds that changes in belief must be

endorsable after they have occurred.

The dialectical situation is that Colyvan (2006, p. 5) has presented Katz with with a dilemma: the argument against Quine must be conducted using either a non-monotonic belief-revision logic, or an informal reflective equilibrium framework. The latter is no help to Katz, so Colyvan concessively concentrates on the former. And he goes on to frame Katz's challenge as the demand for a rational defense of the negation of the allegedly unrevisable principle post-revision. Colyvan claims to have met this challenge by pointing out that a trivial argument can be given from any belief to itself. But with the new paradox in view, it is Colyvan who has to face the dilemma. It would not be a concession to my argument to assume that there is a fixed consequence relation and that the reflective equilibrium approach can be ignored. I am not in reflective equilibrium if I believe that my last change of belief was mistaken, because in such a situation I will backtrack to my previous doxastic state.[1] Unless Colyvan has a different understanding of what reflective equilibrium involves, it is now a concession to him to assume that a belief-revision logic will provide a way of settling the dispute. And there is some reason to think that the belief-revision logic approach can't be best. When we are considering whether to believe ¬ER, it seems inappropriate to take any belief-revision logic as fixed, since any belief set including ¬ER is committed to a particular (highly non-standard) belief-revision logic, according to which all belief-revisions are incorrect.

Even if we do concede to Colyvan that a belief-revision logic is appropriate, that still leaves us with the question of whether belief in ¬ER can be justified post-revision. Now there is presumably some connection between revisability and justification; that is what is involved in UR being a principle of Quine's epistemology. Presumably this connection amounts to the possibility of defining revisability in terms of justification: a belief is revisable in the face of experience iff some experience could in principle justify revising that belief. But then ¬ER says that no experience could in principle justify the revision of any belief. So in particular according to ¬ER a revision which brought about belief in ¬ER cannot have been justified. Recall that Colyvan's suggestion was that since we can always argue from a belief back to itself, a trivial justification of any belief will be available from a belief set containing it. Can there really be such circular justification? Colyvan suggests in a footnote why he thinks there can:

> [T]he circularity is simply the circularity of proving an axiom from itself or justifying an axiom by its consequences. This may be the best we can do when asked to provide justification for fundamental principles or axioms. (2006: 9 n14)

This is unobjectionable as far as it goes: we have to start our epistemic enterprise from somewhere, and it is reasonable to hold that we are justified in continuing to believe what we started out believing, so the original beliefs must have a circular default justification. But ¬ER isn't an axiom or an original or fundamental belief, at least in the case I describe. It is something that one has come to believe, and it is easy to see post-revision that coming to believe it involved a mistake. A good epistemology may well allow that original beliefs have a default justification stemming simply from the fact that we start off believing them, but it will not allow clear mistakes to become entrenched; this forces us to acknowledge a distinction between beliefs for which a circular justification is acceptable and those for which it is not. In a belief revision logic, a circular justification should not be acceptable for beliefs arrived at via revision: it is hard to defend as rational a belief that was formed by a process now taken to be irrational when the only thing counting in favour of maintenance of that belief is the

belief itself. It appears then that a Colyvan-style defense of the revisability of ER would fail even on its own terms.[2]

Colyvan (2006, p. 8) compares revising UR to democratically voting for a dictatorship: unwise but not paradoxical. Whilst it is worrying if holists take traditional empiricism to be like a dictatorship, the analogy is apt for the new paradox in that an epistemology including ¬ER really would be autocratic. But revising ER is still not quite like voting for a dictatorship; it is like taking such a democratic vote to (alone) legitimate dictatorship. That is paradoxical, because if democracy is what confers legitimacy on government, a dictatorship (which does not allow for democratic government) cannot be legitimate, and if democracy does not confer legitimacy then a fortiori it cannot confer it on a dictatorship.

3. Holism and normativity

Taking myself to have improved on Katz's objection to holism, I now want to consider the wider picture of how all this fits in with other objections to holism, and what resources the holist might have in response. Once it is acknowledged that a truly universal revisability principle is untenable, the natural weakening of holism is to restrict the scope of UR.[3] There is an independent line of argument against including logic within the scope of UR: Wright (1986) claims that classical logic (or at least the principle of non-contradiction) appears to be required in order to make sense of recalcitrance. Recall that Quine's position is that when faced with a recalcitrant experience we may adjust our beliefs in various ways, which is meant to show that no experiment can test a particular belief rather than the whole web of beliefs. This position takes it for granted that there is such a thing as a determinately recalcitrant experience. But if we are allowed to revise our logic, it seems that we can make any total theory compatible with any possible experiences by holding that there is no problem if the total theory predicts one experience and we have another. In other words, including logic in holism makes it possible in any case to avoid revision to the theory simply by denying that one's experience is recalcitrant. And this collapses empiricism, since there is no longer a genuine distinction between experiences which are incompatible with one's total theory and those which are not.[4] Whether or not this argument succeeds in showing that UR should not range over logical principles, the new paradox creates trouble even for a suitably restricted version of holism, since ER is not plausibly a logical truth.

Any restriction on UR sufficient to exclude ER will presumably have to exclude itself too, since UR and ER are beliefs of the same kind (the only difference between them is a change of quantifier). So a form of holism immune to the new paradox will also be untouched by the old paradox. Since some responses to Katz's paradox are more concessive than Colyvan's, a restriction to UR has already been suggested, and it will work as a way of avoiding the new paradox too. The idea of the restriction is that UR and ER are both normative claims.[5] This suggests a simple principle which excludes them:

(UR*) All non-normative beliefs are revisable in the face of experience.

Whilst UR* does the job of providing a non-paradoxical form of holism, I think that Quineans would be unwise to rest easy with such a solution. If they can utilise the normative/non-normative distinction to avoid trouble, so can their opponents.

For Quine the rival of his holism, and target of his attacks, was a more moderate empiricism which allowed for a priori, unrevisable knowledge by way of analytic truth. A key plank of the argument for holism is thus Quine's crucial objection to analyticity: that 'analytic' cannot be satisfactorily defined (1980, §4). But what if we say that 'analytic' is a normative term? (It is plausible that claims of analyticity are normative, because they conflict with UR*, which is itself normative, and presumably anything contradicting a normative claim is normative.) Suppose that the primary meaning of '"All vixens are female" is analytic' is to express a commitment to never revising the belief that 'All vixens are female' is true. If this were right then there would be no problem with understanding how the meaning of 'analytic-for-L' is derived from the meaning of 'analytic' (rather than vice-versa, as Quine suggests).

So can we hold that there are analytic truths just in case there are (non-normative) sentences like 'All vixens are female' which we are committed to never revising?[6] Those who challenge the link between analyticity and revisability (and thus the link between semantic and epistemological holism) would deny it. We could admit that there are analytic truths whilst still leaving open the possibility that we are wrong about what those truths are. And my suggested defense of analyticity seems to get this wrong: if I rule out the possibility of revising, then I ignore the possibility of error, and I also ignore the possibility that we could have a simpler (and so better) total theory by revising. To see why this criticism is mistaken we have to distinguish between revising one's belief, and revising one's commitment to not revising that belief. On my view, saying that 'All vixens are female' is analytic rules out revising one's belief that all vixens are female, but it does not rule out revising one's commitment to not revising, which is to say that it does not rule out revising the belief that 'All vixens are female' is analytic. I am saying that claims of analyticity commit one to not revising the belief that is claimed to be analytic, not to not revising the claim of analyticity itself.

An opponent might say that this distinction collapses, because to be committed to the revisability of the claim that a belief is unrevisable just is to be committed to the revisability of that belief. To see that this is an error, consider the commitment to the universal moral wrongness of lying, and suppose that someone has that commitment, but takes it to be revisable (perhaps because she recognizes her own fallibility in moral argument). If the collapse objection were correct then she would be inconsistent, because taking her commitment to be revisable would involve not being committed to not lying in all circumstances. Since she is clearly not inconsistent, there is no collapse. Perhaps it seems that the analogy cannot be a good one, precisely because it appears that there is collapse in the analyticity case and not in the lying case.[7] My thought is that we can use the analogy to explain what is involved in a claim of analyticity. The claim that lying is necessarily wrong commits one to holding that an act is wrong simply in virtue of being an instance of lying, and thus to disregarding any other factors which might count against such acts being wrong. But this does not preclude reconsideration of the principle that lying is always wrong in light of surprising cases of lying which may arise. Following the norm means holding certain judgements (whether particular acts of lying are wrong) to be independent of a certain kind of evidence (the other properties of those acts); but holding the norm does not require one to hold that the other properties of acts of lying are irrelevant to the correctness of the norm itself. When I hold that it is analytic that vixens are female, I commit myself to holding true (inter alia) the sentence 'All vixens are female', regardless of the empirical data; but that does not require me to think that the empirical data is irrelevant to whether it is correct to

accept such a norm.

Another way of seeing this point is to note that in coming to think that 'The original metre bar in Paris is one metre long' is not analytic (i.e. revising the commitment) I do not automatically come to think that 'The original metre bar in Paris is one metre long' is not true (revising the belief). There has to be a difference because one sometimes encounters people who think that it is analytic that the original metre bar is one metre long, but when correcting them one doesn't want them to believe that it *isn't* one metre long. If there is a difference between revising the belief that p and revising the belief that q, then there is plausibly also a difference between taking p to be revisable and taking q to be revisable.

One plausible response for the holist at this point is that whilst the above account of analyticity may be tenable, it breaks the link between analyticity and revisability. Even if I hold that 'All vixens are female' is analytic, and I am thus committed to not revising it, I may still revise my belief that all vixens are female by first revising that commitment. So it seems that analytic truths turn out to be revisable, and thus that it is not after all crucial to holism (or at least epistemological holism) to reject analyticity. This way of looking at things founders, however, once we contrast judgements of analyticity with UR* itself. If the judgement that a non-normative sentence is analytic involves a commitment to not revising one's assent to that sentence, then it must be incompatible with the judgement that all non-normative beliefs are revisable. For whatever precisely 'revisable' means here, holding that a belief is revisable surely requires not being committed to not revising it.[8] If holding that 'Vixens are female' is analytic is inconsistent with accepting UR* (and indeed any similar principle), then holism does seem to have to provide arguments against analyticity.[9]

If Quine's argument against analyticity now appears weak in the light of a strong normative/non-normative distinction, then a holism involving UR* is not a stable refuge. For some analytic truths would be non-normative, and thus would be counterexamples to UR*. And without an argument against analyticity it is unclear why the holist is so insistent that all non-normative beliefs are unrevisable, since she now admits that some beliefs (like ER) are revisable. This is not to deny that there are better formulations of holism than UR* (my reason for focussing on UR* is simply that it is proposed in the literature).[10] But it is worth bearing in mind that it is not much use having a consistent revisability principle if it undermines the motivation for holism, and that is the problem with UR*. Will other attempts do better?

4. Conclusion

Katz's attempt to produce a paradox by diagonalisation on UR fails, as Colyvan shows. But there is a paradox of self-reference lurking. Quineans must admit that ER is not revisable, notwithstanding Colyvan's points about the methodology of such disputes. They are thus committed to a restriction of UR that goes beyond excluding logical laws from its scope; for ER is not a logical law. The obvious restriction is to exclude normative beliefs. But if a normative/non-normative distinction can help Quine out, it can also undercut his argument against analyticity, and thus a Quinean argument for UR*. Katz was on to something: holists need to watch out for revisability paradoxes.[11]

REFERENCES

Adler, J. (2003). "The Revisability Paradox," Philosophical Forum **XXXIV**, pp. 383–389.
Ahmed, A. (2000). "Hale on Some Arguments for the Necessity of Necessity," Mind **109**, pp. 81–92.
Colyvan, M. (2006). Naturalism and the Paradox of Revisability," Pacific Philosophical Quarterly **87**, pp. 1–11.
Katz, J. J. (1998). Realistic Rationalism. Cambridge, MA: MIT Press.
Quine, W. V. O. (1960). Word and Object. Cambridge, MA: MIT Press.
Quine, W. V. (1980). "Two Dogmas of Empiricism," in From a Logical Point of View, second edition. Cambridge, MA.: Harvard University Press, pp. 20–46.
Resnik, M. D. and Orlandi, N. (2003). "Holistic Realism: A Response to Katz on Holism and Intuition," Philosophical Forum **XXXIV**, pp. 301–315.
Wright, C. (1986). "Inventing Logical Necessity," in Language, Mind and Logic, ed. J. Butterfield. Cambridge: Cambridge University Press, pp. 187-209.

---

[1] I do not mean that whenever I notice that I have made a mistake in reasoning I must abandon the belief formed on the basis of the bad reasoning; perhaps there is some independent justification for that belief which I noticed only after adopting it. But the test of whether this justification is genuinely independent is whether it was in principle available even before the belief was adopted. Insofar as beliefs can provide justification for themselves, given Colyvan's point that there is a valid argument from a belief to itself, this justification will obviously not pass that test of independence, and so will not help to maintain equilibrium in the case given in the text.

[2] It may seem from my argument here that I should simply endorse Katz's original argument against Colyvan's objection. But there are still two reasons for doubting Katz's argument and preferring the new paradox. One is that given a belief-revision logic it isn't clear that the revision of N to ¬N can't be justified, since (applying Resnik and Orlandi's criterion) there is nothing in ¬N to prohibit revision of N. This contrasts with the case of the revision of ER. The second point is that, as I have argued above, a reflective equilibrium approach is at least as plausible as a belief-revision logic as a way of adjudicating the question of whether UR produces a paradox, and the new paradox works much better than the old in the former framework.

[3] Or to achieve the same result by restricting the domain of quantification so that not all beliefs are within it, or by operating with a restricted understanding of 'belief'.

[4] Ahmed (2000) suggests a response to an argument like Wright's. Although it is always in some sense possible for the holist to revise logic so as to render any experience non-recalcitrant, this will not always be a real option. In other words, the holist can take a rule of inference to be up for grabs sometimes, but not always, and thus avoid collapsing empiricism. This position is delicate because it relies on there being some rule for determining when it is an option to revise a rule of inference. Suppose that faced with an apparently recalcitrant experience this second-order rule makes it not an option to revise one's first-order rules of inference so that the experience is not recalcitrant. Why can't we revise the second-order rule itself, so that it is then an option to revise the first-order rule which makes the experience recalcitrant? Presumably because there is a third-order rule which makes it not an option to revise the second-order rule in those circumstances. But we are now threatened with an infinite regress of rules, which would involve an implausible infinitist account of our grasp of logical rules. Hope of avoiding this regress perhaps lies in rules which apply to themselves.

[5] Both Resnik and Orlandi (2003) and Adler (2003) suggest (in response to Katz's alleged paradox) using the normative/non-normative distinction to restrict UR.

[6] As argued above, ER is unrevisable, but it may not seem to be analytic. If it is not, then the definition of 'analytic' given in the text would have to be modified, but I think the basic point would survive.

[7] This objection was suggested by an anonymous referee.

[8] What this reveals is that if we model the revisability of a belief in modal logic as the accessibility of a web of beliefs not containing that belief from the present web, then the relevant accessibility relation is not transitive. I suspect that the incorrect intuition (that the suggested view of analyticity collapses) rests on an assumption that the accessibility relation must be transitive.

[9] As I was reminded by an anonymous referee, Quine (1960) does accept a limited class of analytic truths: the one-criterion word ones. But the strategy I have outlined could I think be used to explicate all claims of analyticity.

[10] An anonymous referee suggests that the various restricted versions of UR expressible whilst doing without an unrestricted truth predicate will likely be immune to revisability paradoxes.

[11] I am grateful to Arif Ahmed, David Liggins, Florian Steinberger, Tim Storer and two anonymous referees for help with this paper, and to an audience at the 2006 Joint Session of the Mind Association and the Aristotelian Society, where I presented an earlier version.