Article:

An Unacceptability Paradox and an "Unacceptable" Equivocation

Abstract: This paper presents, and solves, a new paradox of unacceptability. Inspired by a recent argument by Bradley Armour-Garb and James Woodbridge, my generalisation of their reasoning, if valid, shows that anyone who takes any sentence whatsoever to be unacceptable is committed to contradiction. I then show how to solve the paradox, by arguing that the reasoning in question equivocates on the word 'unacceptable'. Those who think my solution works will learn something about how (not) to reason about acceptability, with important consequences for our evaluation of arguments in the philosophy of truth. Those who disagree will have a new paradox to grapple with.

This paper presents, and solves, a new paradox of unacceptability. The paradox is inspired by a recent argument by Bradley Armour-Garb and James Woodbridge (2024), hereafter AGW. My generalisation of their reasoning, if valid, shows that anyone who takes any sentence whatsoever to be unacceptable is committed to contradiction. I call this, the Unacceptability Paradox. This is a perfectly general problem that, to my knowledge, has not been identified in the literature before (not even by AGW, who do not intend to present a paradox, but rather to refute a niche view in the philosophy of truth). I then show how to solve the paradox, by arguing that the reasoning in question equivocates on the word 'unacceptable'. If the reader thinks my solution works, they learn something about how (not) to reason about acceptability (with consequences, for example, for our evaluation of arguments in the philosophy of truth). If the reader finds fault with my solution, they have a new paradox to grapple with.

First, AGW's original argument. AGW's target is alethic nihilism, the theory that nothing is true, which David Liggins and I have independently defended as a response to the alethic paradoxes.[1] AGW's goal is to identify a 'revenge' problem for nihilism. To do so, they focus on (R):

(R)   If (R) is acceptable, then (R) is true.

Two things are crucial to the argument. First, AGW's notion of *acceptability*: 'a sentence is acceptable for a speaker provided it should be accepted by that speaker' (2024: 688). With AGW, I will usually suppress the relativisation to a speaker. They argue that the following principles are 'rational constraints on making "acceptability evaluations"' (2024: 688):[2]

(A)   If '*p*' is acceptable, then ''*p*' is acceptable' is acceptable.

(B)   If ''*p*' is acceptable' is acceptable, then '*p*' is acceptable.

Second, a 'classical conception of the acceptability and unacceptability of conditionals' (2024: 689). AGW do not explicitly state what they take this conception to consist in, but it is clear from the discussion that they take it to involve at least the following two principles, which are sufficient to state the objection:

(C)   If '*p*' is unacceptable, then 'if *p* then *q*' is acceptable.

(D)   If '*p*' is acceptable and '*q*' is unacceptable, then 'if *p* then *q*' is unacceptable.

---

[1] See Liggins 2019, 2024 and Gamester 2023. AGW suggest the problem is specifically for my 'more permissive' variety of nihilism, but it seems to apply equally to Liggins.

[2] These might be better expressed as inference rules rather than conditionals, but nothing turns on this presentational choice. AGW motivate two further rational constraints, but these are unnecessary for the problem.

With (A)-(D) in place, AGW argue that alethic nihilists are committed to contradictory evaluations of (R). In reconstructing the argument, it will be helpful to divide it into four steps.

In Step I, AGW argue that the antecedent of (R) cannot be evaluated as unacceptable on pain of contradiction:

(1)	The antecedent of (R) is unacceptable.	*Assumption for reductio*.
(2)	(R) is acceptable.	*(1), (C).*
(3)	'(R) is acceptable' is acceptable.	*(2), (A).*
(4)	The antecedent of (R) = '(R) is acceptable'.	*Premiss*.
(5)	The antecedent of (R) is acceptable.	*(3), (4).*
(6)	*Contradiction*.	*(1), (5).*

In Step II, AGW conclude from this that we 'can evaluate the antecedent of (R) only as acceptable' and so 'must evaluate the antecedent of (R) as acceptable' (2024: 689):

(7)	The antecedent of (R) is acceptable.	*From (1)-(6).*

Since the antecedent of (R) is '(R) is acceptable', Step III moves from (7) to (R) being acceptable:

(8)	'(R) is acceptable' is acceptable.	*(4), (7).*
(9)	(R) is acceptable.	*(8), (B).*

And since nihilists are committed to the consequent of (R) being unacceptable,[3] Step IV argues from (7) to (R) being unacceptable to nihilists:

---

[3] For AGW, to say that '*p*' is acceptable for a speaker is to say that they *should* accept it. There are many senses of 'should': moral, prudential, legal, rational, etc. AGW do not say what sense of 'should' they have in mind.

(10)	'(R) is true' is unacceptable.	*Commitment of nihilism.*
(11)	The consequent of (R) = '(R) is true'.	*Premiss.*
(12)	The consequent of (R) is unacceptable.	*(10), (11).*
(13)	(R) is unacceptable.	*(7), (12), (D).*

From (9) and (13), we see that nihilists are committed to contradictory evaluations of (R).

So presented, this looks like bad news for the alethic nihilist, but of little concern to the rest of us. However, it's straightforward to verify that we can apply reasoning parallel to that just given to show that anyone who takes any sentence whatsoever to be unacceptable is committed to contradiction. Suppose, for instance, we replace '(R) is true' in (R) with something obviously unacceptable, like 'genocide is morally good':

(R2)   If (R2) is acceptable, then genocide is morally good.

Steps I-III can be used to show that (R2) is acceptable, then Step IV to show that anyone who thinks that 'genocide is morally good' is unacceptable is committed to (R2) being unacceptable. In general, Steps I-III can be used to show that *any* conditional of the following form is acceptable:

(R3)   If (R3) is acceptable, then *p*.

This matters. The nihilist might think that (R) is not true while nonetheless thinking that it is, say, prudent or rational for a particular individual to accept that it is true. For the sake of argument, however, I suppose that there is *some* relevant sense of 'should' in which thinking that not-*p* commits you to '*p*' being (strongly) unacceptable.

We can then use Step IV to show that anyone committed to the unacceptability of '*p*' will be committed to contradictory evaluations of the relevant instance of (R3). In short, this reasoning, if valid, would be problematic for *everyone*. This is the Unacceptability Paradox.

Rather than identifying a 'revenge' problem for alethic nihilism in particular, then, AGW's argument exemplifies a perfectly general problem that threatens us all – like an arrow that misses its target and instead flies straight into the launch button for a nuclear bomb.[4] Now we all have a problem on our hands. We're all in need of a solution. To my knowledge, the Unacceptability Paradox has not been previously identified in the literature before – and, in any case, my interest here concerns whether we can solve it. We might see it as a *reductio* of principles (A)-(D). We might notice its structural similarity to the Curry Paradox and hope that our preferred solution carries over. But such responses, I will now argue, are unnecessary. In the rest of this paper, I defuse the bomb by showing that the above argument equivocates on the word 'unacceptable'.

Recall that '*p*' is *acceptable* for a speaker *S* provided:

(a)   *S* should accept '*p*'.

What, then, does it mean to say that '*p*' is *not acceptable* or *unacceptable*? We should be careful to distinguish the following:

(b)   It is not the case that *S* should accept '*p*'.

(c)   *S* should not accept '*p*'.

---

[4] As an objection to alethic nihilism, AGW's argument only seems to show that it does not solve *all* paradoxes, which is not very surprising. Liggins and I intend for it to solve the alethic paradoxes: I focus on the Liar and Curry paradoxes; Liggins also includes Yablo's paradox and the truth-teller. This problem, it should now be clear, has nothing in particular to do with truth. It is thus no more an objection to alethic nihilism that it does not solve the Unacceptability Paradox than that it does not solve the Sorites Paradox.

Assuming that one should never both do and not do something, 'S should not φ' entails 'It is not the case that S should φ'; but the converse entailment does not hold. For example, if S should paint her bedroom, but may paint it either blue or red, then it's *not* the case that she *should* paint it red (since she can paint it blue, if she wants), but obviously that does not mean that she *should not* paint it red. So, while (c) entails (b), (b) does not entail (c) unless we assume that S should accept '*p*' or S should not accept '*p*'. If it is permissible for S to either accept '*p*' or not accept '*p*', this will not hold.

Both (b) and (c) should, of course, be kept distinct from:

(d)   S should accept 'not-*p*'.

Assuming that one should never accept both a sentence and its negation, (d) entails (c) – and thus (b) – but the converse entailment only holds if we assume that S should accept '*p*' or S should accept 'not-*p*'. There are three ways this could fail to hold: (i) S should accept neither; (ii) it is permissible for S to accept either; or (iii) S should not accept one of them, but may either accept or not accept the other.

Call the notion of unacceptability captured by (b), *weak unacceptability*; the notion captured by (c), *moderate unacceptability*; and that captured by (d), *strong unacceptability*. Importantly, if 'unacceptable' is understood in the moderate or strong sense, then it is logically possible for a sentence to be neither acceptable nor unacceptable. It is only if it is understood in the weak sense a sentence must be one if it's not the other.[5]

---

[5] AGW's use of the term 'acceptable' is thus potentially misleading. The suffix '-able' often has the modal force of *may* or *can*, rather than *should*: to say that something is *doable* is to say that it can be done or may be done, not that it should be done. This makes it tempting to read ''*p*' is acceptable' as saying that it is *permissible* to accept '*p*', and thus tempting to read ''*p*' is not acceptable' or ''*p*' is unacceptable' as saying that it is *impermissible* to accept '*p*'. On this reading, two things hold: (i) '*p*' must be either acceptable or unacceptable; and (ii) ''*p*' is unacceptable' implies that S should not accept '*p*'. But that is only because 'acceptable' is being read in a weaker sense than intended. On the stipulated meaning for 'acceptable,' we either get (i) or (ii), *but not both*.

Armed with these distinctions, let's return to our paradox. Recall that Step I argues that the antecedent of (R) cannot be evaluated as 'unacceptable' on pain of contradiction, and Step II concludes from this that we must evaluate it as acceptable. As we've just seen, if 'unacceptable' is understood in the moderate or strong sense, then Step II is invalid, since the antecedent of (R) might be neither acceptable nor unacceptable. But, as I will now argue, if 'unacceptable' is understood in the weak sense, then Step I is invalid. So there is no understanding of 'unacceptable' on which Steps I and II of the argument are both valid.

Recall that the reasoning of Step I starts as follows:

(1)   The antecedent of (R) is unacceptable.   *Assumption for reductio*.

(2)   (R) is acceptable.   *(1), (C).*

As noted, this relies on a 'classical conception of the acceptability and unacceptability of conditionals' and particularly the principle:

(C)   If '*p*' is unacceptable, then 'if *p* then *q*' is acceptable.

We can distinguish three versions of this principle depending on how we understand 'unacceptable':

(C1)   If it is not the case that *S* should accept '*p*', then *S* should accept 'if *p* then *q*'.

(C2)   If *S* should not accept '*p*', then *S* should accept 'if *p* then *q*'.

(C3)   If *S* should accept 'not-*p*', then *S* should accept 'if *p* then *q*'.

If 'unacceptable' in (1) means *weakly* unacceptable, as the paradox requires, the inference from (1) to (2) requires (C1). But it is *at most* (C3) that can plausibly be said to be part of 'a classical conception of the acceptability and unacceptability of conditionals'. The idea, I take it, is that on this conception 'not-

p' entails 'if *p* then *q*,' so one might argue that any good reason to accept the former is a good reason to accept the latter, and hereby motivate (C3). This reasoning is questionable, but set that aside: the important point is that no such reasoning can be used to motivate (C2), let alone (C1). The mere fact that one *should not* accept '*p*' does not mean that one *should* accept 'not-*p*' – perhaps one should accept neither – and so gives one no reason to accept 'if *p* then *q*'.[6]

The upshot is that (C), and thus Step I of the argument, requires that 'unacceptability' in assumption (1) be *strong* unacceptability, otherwise the inference from (1) to (2) is invalid. But this scuppers Step II of the argument, which requires it to be weak unacceptability: Step I can still be used to show that the antecedent of (R) cannot be evaluated as strongly unacceptable, but it does not follow that it can only be evaluated as acceptable, since it can still be evaluated as weakly or even moderately unacceptable. So, when 'unacceptability' is understood in the strong sense, Step I is valid, but Step II is not. When 'unacceptability' is understood in the weak sense, Step II is valid, but Step I is not. And when 'unacceptability' is understood in the moderate sense, neither step is valid.

It is thus perfectly coherent to think that the antecedent of an instance of (R3) is weakly or even moderately unacceptable – that is, that it's not the case that we should accept it or even that we should not accept it – without thinking that the conditional as a whole is acceptable. We are therefore not forced into contradictory evaluations of the sentences in question by virtue of thinking that the consequent is unacceptable (in any sense). While the Unacceptability Paradox is a perfectly general problem that has hitherto gone unnoticed in the literature, it is a problem that we can solve.[7]

---

[6] AGW say that 'Gamester […] indicates that he endorses' (2024: 689) the classical conception. This is because I think that alethic nihilists should accept (TC): 'If there is such a property as truth, then '*p*' is true iff *p*.' But this is not merely because it's not the case that the nihilist should accept its antecedent, nor even because the nihilist should not accept its antecedent, but because the nihilist should accept the negation of its antecedent. My discussion thus indicates, at most, an endorsement of (C3), rather than (C1) or (C2) – though this is weak evidence even for that, since it only amounts to endorsing one instance of the generalisation.

[7] Thanks to Bradley Armour-Garb, David Liggins, Robbie Williams, James Woodbridge and to the audience at the University of Leeds's Centre for Theoretical Philosophy for helpful feedback on this material.

University of Leeds, UK, w.gamester@leeds.ac.uk

References

Armour-Garb, B. and J. Woodbridge. 2024. Revenge for alethic nihilism. *Journal of Philosophy* 121: 686–97.

Gamester, W. 2023. Nothing is true. *Journal of Philosophy* 120: 314–38.

Liggins, D. 2019. In defence of radical restrictionism. *Philosophy and Phenomenological Research* 98: 3 –25.

Liggins, D. 2024. *Truth Without Truths*. Oxford: Oxford University Press.