Self-Excluding Strengthenings and the Limits of Representation

Abstract

Some distributions of representational states look locally intelligible but cannot be globally realized. My earlier paradox involving occurrent belief brings this out: from an apparently possible allocation of attitudes, contradiction follows. Priest has offered a dialetheic treatment of the paradox, and Field's recent work on propositions, properties, and paradox indicates a wider non-classical setting in which cases of this sort may be handled. This paper develops a classical alternative, closer in methodological spirit to Williamson's response to Prior–Kaplan paradoxes. The key claim is not that logic must be weakened, but that the relevant representational distribution is not genuinely possible. I argue that the impossible feature is that both sides of a contingent divide are exhausted by self-excluding strengthenings. A self-excluding strengthening of P, relative to a truth-assessable attitude or representational relation O, is, roughly, a strengthening of P that also says that no one ever truly O's any strengthening of P. Such a content cannot be truly represented. If a total distribution exhausts the P-side of a contingent divide in that way, P is false. Hence no possible distribution of truth-assessable attitudes can have both sides of a contingent divide exhausted by self-excluding strengthenings. This covers both the original two-sided case $\{P, \neg P\}$ and the one-person case generated by $P \vee \neg P$. The paradox therefore reveals a general limit on truth-assessable representation, not a failure of classical logic.

1. Introduction

In earlier work, I presented a paradox involving occurrent belief and other representational states and activities (Myers, 2019). The paradox begins with a simple observation: there are certain contingently true propositions that no one can truly occurrently believe. It then turns that observation into a contradiction by embedding such propositions inside an apparently possible total distribution of occurrent beliefs. I also argued there that the same general pattern extends beyond occurrent belief to assertion, utterance, writing, imagining, hoping, and related representational states and activities.

That paradox naturally invites a non-classical response. Priest (2021) offers a dialetheic treatment. He reconstructs the paradox with an intensional operator, read for convenience as "says that," while stressing that the same formal role could be played by occurrent belief, assertion, writing, or other representational relations. His conclusion is that paraconsistent resources can block the original derivation and, under suitable conditions, accommodate the paradoxical upshot without triviality.

Field's recent work on propositions, properties, and paradox suggests a different kind of non-classical background (Field, 2026). He argues against views on which the basic paradoxes arise merely because certain sentences fail to express propositions, express different propositions from the ones they appear to express, or express multiple harmless propositions. Instead, he argues that the paradoxes arise already at the level of propositions and properties themselves, or alternatively at the level of sentential and predicate quantification. Although he notes that such paradoxes may be handled classically or non-classically, his own discussion focuses mainly on non-classical options.

This paper develops a classical alternative. I do not argue that Priest's response is incoherent, nor that Field-style non-classical resources are unavailable here. My aim is different. I ask what a classical theorist should say if the original paradox is not taken to motivate a revision of logic. The minimal classical answer is that the relevant world-description is impossible. But that answer is not enough. A better classical treatment should say what kind of distribution is impossible, and why.

The proposal of this paper is that the paradox reveals a structural prohibition on representation. The underlying idea is that no distribution of truth-assessable attitudes can have both sides of a contingent divide exhausted by self-excluding strengthenings.

Here is the intuitive picture. Let O be any truth-assessable attitude or representational relation: occurrently believing that, asserting that, writing that, uttering that, and so on. Let a proposition M strengthen P just in case M entails P, but P does not entail M. Now consider a content that not only strengthens P, but also says that no one ever truly bears O to any strengthening of P. Such a content excludes its own truth if it is ever truly represented. It is therefore natural to call it a self-excluding strengthening of P.

The core thought is then simple. If the P-side of a representational distribution is exhausted by a self-excluding strengthening, the P-side collapses: P must be false. If the same thing happens on the $\neg P$-side, contradiction follows. And if the whole divide is represented in compressed form by $P \vee \neg P$, then exhausting that single proposition by a self-excluding strengthening yields contradiction at once.

This is the sense in which the paradox concerns a contingent divide. A contingent proposition P determines two sides: P and $\neg P$. Sometimes those sides are displayed separately, as in the original two-person version of the paradox. Sometimes they are represented together, as in a one-person case involving $P \vee \neg P$. The present proposal treats both as instances of the same underlying impossibility. The split presentation tracks the two sides individually; the compressed presentation closes both sides at once.

The methodological stance is close to Williamson's response to Prior–Kaplan paradoxes (Williamson, 2018). Williamson resists taking such paradoxes as evidence that classical logic must be weakened. He prefers to preserve a strong logical framework and to look instead for a more structural diagnosis of what has gone wrong. The present paper adopts a similar orientation. Namely, rather than revealing problems with classical logic, it is taken to show a hidden constraint on representational possibility.

2. The occurrent belief case

I begin with the original belief version of the paradox. Let W be a contingent proposition. In the original presentation, W is the proposition that some narwhal weighs more than 1.5 tons. Now consider the proposition:

$E_B(W)$ = no one at any time truly occurrently believes a proposition that strengthens W.

Then consider the proposition:

$S_B(W) = W \wedge E_B(W)$.

$S_B(W)$ is a strengthening of W, and it is self-excluding in the relevant sense. If someone truly occurrently believes $S_B(W)$, then someone truly occurrently believes a strengthening of W, namely $S_B(W)$ itself. But that is exactly what $E_B(W)$ rules out. So $S_B(W)$ cannot be truly occurrently believed. The same pattern applies on the other side. Let

$E_B(\neg W)$ = no one at any time truly occurrently believes a proposition that strengthens $\neg W$,

and let

$S_B(\neg W) = \neg W \wedge E_B(\neg W)$.

Again, $S_B(\neg W)$ is a self-excluding strengthening, now of $\neg W$. If someone truly occurrently believes it, then someone truly occurrently believes a strengthening of $\neg W$, contrary to what the content itself says. So $S_B(\neg W)$ cannot be truly occurrently believed either.

The original paradox arises when these two one-sided results are inserted into a single total description of a world. Suppose that Omar occurrently believes $S_B(W)$, Sarah occurrently believes $S_B(\neg W)$, Omar's content is the only occurrently believed strengthening of W, and Sarah's content is the only occurrently believed strengthening of $\neg W$.

Thus, Omar's belief is not true. Since Omar's content is the only believed strengthening of W, it follows that no strengthening of W is truly occurrently believed. So $E_B(W)$ is true. But Omar still occurrently believes $W \wedge E_B(W)$. Since that content is false while its second conjunct is true, W must be false. Hence $\neg W$.

Now run the same reasoning on Sarah's side. Sarah's belief is not true. Since her content is the only believed strengthening of $\neg W$, it follows that no strengthening of $\neg W$ is truly occurrently believed. So $E_B(\neg W)$ is true. But Sarah occurrently believes $\neg W \wedge E_B(\neg W)$. Since that content is false while its second conjunct is true, $\neg W$ must be false. Hence W. Therefore both W and $\neg W$ are true. Contradiction.

A non-classical theorist may treat that outcome in a variety of ways. The classical theorist, however, should reject the initial total world-description. But this raises the question: what, exactly, is wrong with it? The answer I want to develop is that the description makes both sides of the $W/\neg W$ divide exhausted by self-excluding strengthenings. That is the real impossibility.

3. From occurrent belief to arbitrary truth-assessable attitudes

The phenomenon is not peculiar to occurrent belief. Let O be any truth-assessable attitude or representational relation. To say that someone truly O's P is just to say that the person bears O to P, and that P is true.

Now, for any proposition P, define:

$E_O(P)$ = no one at any time truly O's a proposition that strengthens P.

Then define the associated self-excluding strengthening of P:

$S_O(P) = P \wedge E_O(P)$.

This strengthens P, but if anyone truly O's it, then someone truly O's a strengthening of P, which makes $E_O(P)$ false. We can now state the basic lemma.

Lemma 1. If O is truth-assessable, no one can truly $O\ S_O(P)$.

Proof. Suppose, for reductio, that someone truly O's $S_O(P)$. Since $S_O(P)$ strengthens P, it follows that someone truly O's a strengthening of P. So $E_O(P)$ is false. But $E_O(P)$ is a conjunct of $S_O(P)$, and the content was assumed to be truly O'd. Contradiction. Therefore, no one can truly $O\ S_O(P)$.

The next step is to shift from an isolated content to a total distribution of O-states. A total O-distribution at a world is simply the complete pattern, at that world, of who bears O to which propositions. The point of speaking this way is to emphasize that the paradox is not generated by a single representational event in abstraction. It depends on a whole distribution having a certain structure.

I will say that a total O-distribution **exhausts the P-side by a self-excluding strengthening** iff:

1. someone O's $S_O(P)$, and
2. every proposition that anyone O's and that strengthens P is equivalent to $S_O(P)$.

Thus, the entire represented strengthening space on the P-side involves the self-excluding strengthening of P. There is no other represented strengthening on that side to make $E_O(P)$ false while leaving P true. This yields the first main result.

Theorem 1. If a total O-distribution exhausts the P-side by a self-excluding strengthening, then $\neg P$.

Proof. Assume the P-side is exhausted in the stated way. By Lemma 1, no one truly O's $S_O(P)$. By the exhaustion condition, every O-content that strengthens P is equivalent to $S_O(P)$. So, no strengthening of P is truly O'd. Hence $E_O(P)$ is true. But someone does $O\ S_O(P)$, and $S_O(P)$ is $P \wedge E_O(P)$. Since $S_O(P)$ is not true while $E_O(P)$ is true, P must be false. Therefore $\neg P$.

This is the one-sided core of the paradox. Exhausting the P-side by a self-excluding strengthening forces the falsity of P. If P is contingent, that need not yet be a contradiction; it merely forces things to the other side of the divide. Contradiction arises when the other side is exhausted as well, or when the entire divide is represented in one compressed proposition.

4. Contingent divides

4.1. Split and compressed presentations

The original paradox is most naturally presented by starting with a contingent proposition P and then considering the two opposed sides P and $\neg P$. I will call that opposition the **contingent divide generated by** P.

There are two natural ways to represent that divide. The first is the **split presentation**. Here the divide is displayed explicitly by the pair:

$\{P, \neg P\}$.

The second is the **compressed presentation**. Here the same divide is represented by the single proposition:

$P \vee \neg P$.

The compressed form matters because it shows that the deeper structure is not essentially tied to two people or two separately represented sides. A single proposition can already span the whole divide. If that single proposition is exhausted by its self-excluding strengthening, contradiction follows in one step.

Thus, when I say that both sides of a contingent divide are exhausted, I mean to cover both of the following cases: either the two sides are separately exhausted in split form, or the entire divide is exhausted in compressed form.

4.2. Exhausting both sides

We can now define the general pattern. A total O-distribution **has both sides of the contingent divide generated by** P **exhausted by self-excluding strengthenings** iff either:

1. it exhausts the P-side by $S_O(P)$ and exhausts the $\neg P$-side by $S_O(\neg P)$; or
2. it exhausts the $(P \vee \neg P)$-side by $S_O(P \vee \neg P)$.

These are not two unrelated cases. They are two presentations of the same structural idea. In the split case, the two sides are excluded separately. In the compressed case, the single disjunction already ranges over both sides at once. This is what allows the classical principle to be stated in this form: no distribution of truth-assessable attitudes can have both sides of a contingent divide exhausted by self-excluding strengthenings.

4.3. The main theorem

Now, we can state the general classical result.

Theorem 2. For any truth-assessable attitude or representational relation O, and any contingent proposition P, no possible total O-distribution can have both sides of the contingent divide generated by P exhausted by self-excluding strengthenings.

Proof. First consider the split case. Suppose the distribution exhausts the P-side by $S_O(P)$ and the $\neg P$-side by $S_O(\neg P)$. By Theorem 1, exhaustion of the P-side yields $\neg P$. By the same

theorem applied to $\neg P$, exhaustion of the $\neg P$-side yields $\neg\neg P$, and hence P. So, both P and $\neg P$. Contradiction. Now consider the compressed case. Suppose the distribution exhausts the $(P \vee \neg P)$-side by $S_O(P \vee \neg P)$. By Theorem 1, it follows that: $\neg (P \vee \neg P)$. But $P \vee \neg P$ is classically necessary. Contradiction. So, in neither presentation is such a distribution possible.

4.4. Two immediate consequences

The original two-person paradox is now just a special case. Let O be occurrent belief and let $P = W$. Then a world in which Omar's belief exhausts the W-side and Sarah's belief exhausts the $\neg W$-side is impossible, precisely because both sides of the $W/\neg W$ divide are exhausted by self-excluding strengthenings.

The one-person case is a second special case. Let O again be occurrent belief, and let one subject occurrently believe:

$(P \vee \neg P) \wedge E_O (P \vee \neg P)$,

with that content being the only occurrently believed strengthening of $P \vee \neg P$. Then, the whole divide is exhausted in compressed form, and contradiction follows immediately.

This is why the classical diagnosis should be stated at the level of contingent divides rather than at the narrower level of the original two-person arrangement. The original paradox is one instance of a broader structural impossibility. What fails is not merely that Omar and Sarah cannot jointly occupy those states. What fails is that no total distribution can close off both sides of a contingent divide by self-excluding strengthenings.

5. Why this is explanatory

A natural worry is that the preceding theorem does no more than redescribe the paradox at a higher level of abstraction. The original world-description yields contradiction; the theorem says that no possible total distribution has that form. Why regard that as explanatory progress? I think it is progress for three reasons.

5.1. It locates the real source of the impossibility

The central point is that the paradox is not generated by self-reference alone, nor even by self-exclusion taken in isolation. The one-sided result already shows something important: if the P-side of a total O-distribution is exhausted by a self-excluding strengthening of P, then P is false. But that by itself does not always yield contradiction. Contradiction emerges only when the other side of the same contingent divide is exhausted as well, or when the whole divide is represented in compressed form by $P \vee \neg P$.

That matters because it yields a more exact diagnosis than the tempting but rough claim that "self-referential representational contents generate paradox." They do not, by themselves, do that. Nor is it enough merely that a content excludes its own true representation. What matters is a further organizational feature: the relevant same-side strengthening space is closed off by the very content whose truth would rule out any true occupation of that space. And when that happens on both sides of a contingent divide, or when it happens to the disjunctive compression of the whole divide, contradiction follows.

The upshot is that the impossibility is neither merely local nor merely lexical. It belongs to a certain pattern of global representational distribution. That is why the theorem has more explanatory force than the bare claim that one specific paradox-world is impossible.

5.2. It explains nearby harmless variants

The present diagnosis also explains why nearby cases fail to be paradoxical. Suppose, for example, that in addition to Omar's bearing O to $S_O(P)$, some third subject truly bears O to an ordinary strengthening of P that is not equivalent to $S_O(P)$. Then the P-side is no longer exhausted by the self-excluding strengthening. In that case $E_O(P)$ is false, and the route from the total distribution to $\neg P$ is blocked.

Likewise on the split version of the contingent divide, one side may be exhausted without the other being so. Then only the corresponding one-sided consequence follows. If the P-side is exhausted, we get $\neg P$. But unless the $\neg P$-side is also exhausted, or unless the whole divide has been compressed into a necessary disjunction and exhausted there, contradiction does not result.

This is explanatory because it shows which feature matters and which do not. The paradox is not triggered by the mere occurrence of Omar and Sarah, nor by the mere existence of contents like $P \wedge E_O(P)$. What matters is the role those contents are made to play inside a total representational arrangement.

5.3. It reveals a structural law

Most importantly, the theorem is not just a redescription of the original narwhal example. It states a general law:

No distribution of truth-assessable attitudes can have both sides of a contingent divide exhausted by self-excluding strengthenings.

That law does not mention narwhals, Omar, Sarah, or even belief. It applies to any truth-assessable representational relation O, and to any contingent proposition whose divide is presented either in split form $\{P, \neg P\}$ or in compressed form $P \vee \neg P$. If the law is right, then representational possibility is more tightly structured than a naive recombinational picture would suggest. Some total arrangements of thought, assertion, writing, and the like are impossible, not because any one local component is obviously incoherent, but because of the way those components are globally organized.

6. Relation to Priest, Field, and Williamson

The classical account developed here stands between two alternatives.

6.1. Priest's dialetheic treatment

Priest offers a dialetheic treatment of the paradox (Priest, 2021). He reconstructs the case using an intensional operator, read for convenience as "says that," while emphasizing that the same formal role may be played by occurrent belief, assertion, writing, and related representational relations. His proposal is twofold. First, in a paraconsistent setting the second stage of the original argument fails, because the crucial inferences are not valid there. Second, under suitable assumptions, the resulting conclusion can be accommodated without collapse into trivialism.

That is a genuine solution. But it is not the one pursued here. The present paper asks a different question: what should a classical theorist say if the paradox is not taken to motivate a revision of logic? My answer is that the paradox should be read as revealing a structural impossibility theorem about total representational distributions. The contradiction is not embraced; rather, the relevant world-description is rejected as an impossible arrangement of truth-assessable attitudes.

6.2. Field's broader paradoxical framework

Field's recent discussion of propositions, properties, and paradox supplies a broader setting in which cases like this can belong (Field, 2026). He argues against approaches that explain away the basic paradoxes by claiming that the troublesome sentences fail to express propositions, express only harmless propositions, or express several non-paradoxical propositions at once. Instead, he argues that the basic paradoxes arise already at the level of propositions and properties themselves, or alternatively at the level of sentential and predicate quantification.

That broader orientation is congenial to the present paper. My own proposal likewise treats the paradox as a paradox of contents and their distribution, not as a superficial pathology of linguistic expression. Where I differ is in the lesson drawn from that fact. Field's framework leaves room for classical treatment, but it primarily develops non-classical options. My aim here is to show that one may instead preserve classical logic while still extracting a substantive theoretical lesson: some total distributions of truth-assessable representation are impossible because they attempt to close both sides of a contingent divide with self-excluding strengthenings.

6.3. A classical method in Williamson's spirit

The closest methodological analogue is Williamson's published response to Prior–Kaplan paradoxes (Williamson, 2018). Williamson resists taking those paradoxes to show that classical logic must be weakened. He emphasizes that the difficulty is robust across different ways of individuating content, and he is skeptical of strategies that restrict the underlying logical machinery at too high a theoretical cost. The better strategy, on his view, is to preserve a strong logical framework and to locate the failure elsewhere.

The present paper adopts that same general stance. The paradox is not dismissed as illusory. Nor is it taken to reveal a defect in classical logic. Rather, it reveals a hidden restriction on which total distributions of truth-assessable attitudes are genuinely possible. In that sense, the proposal is classical in content and Williamsonian in method.

7. Clarifications

The central claim of this paper is that no possible total distribution of a truth-assessable attitude can have both sides of a contingent divide exhausted by self-excluding strengthenings. It is worth clarifying some points about what this does and does not establish.

7.1. Why speak of a contingent divide?

It may be asked why the paper is framed in terms of a contingent divide rather than in the language of partitions. The reason is that the paradox is fundamentally organized around the opposed ways a single contingent matter can go. In the split presentation, that opposition is

explicit in the pair $\{P, \neg P\}$. In the compressed presentation, the same opposition is represented by the single proposition $P \vee \neg P$.

The language of "partition" is too strong for this purpose. What matters here is not just any jointly exhaustive family of propositions. What matters is a divide generated by a single contingent proposition, together with the possibility of representing that divide either side-by-side or in compressed form. Speaking this way makes it clearer why the original two-person case and the one-person disjunctive case are not two unrelated curiosities, but two versions of the same structural phenomenon.

7.2. The result is not merely a dressed-up contradiction claim

Another possible objection is that the theorem amounts to little more than the classical observation that contradictory assumptions cannot all be satisfied. But that understates what has been achieved. There are indefinitely many inconsistent descriptions of possible worlds. The interest of the present theorem lies in isolating a distinctive family of them: those in which each relevant side of a contingent divide is occupied entirely by a content that denies the possibility of any true same-side strengthening.

That is not a trivial matter of notation. It identifies a specific mode of representational failure. The contradiction is real, but its importance lies in the fact that it emerges from a pattern that is general, repeatable, and philosophically interpretable.

7.3. The account does not deny the existence of the troublesome contents

The classical treatment offered here does not require denying that contents such as $S_O(P)$ exist, are meaningful, or are proposition-like. One need not say that they fail to express anything, or that they are excluded from the ontology of contents. The claim is instead that there are limits on how such contents can be distributed across a total representational arrangement.

This matters dialectically. Field argues that the basic paradoxes should not simply be explained away by a narrow appeal to semantic defectiveness. The present account is consistent with that moral. It locates the impossibility not in the mere existence of paradox-prone contents, but in the attempt to let such contents exhaust both sides of a contingent divide.

7.4. Representational possibility is global, not merely local

Finally, representational possibility here is global rather than merely local. A given subject-content pairing may seem intelligible in isolation. Omar's bearing O to $S_O(P)$ does not, by itself, obviously look impossible. Nor does Sarah's bearing O to $S_O(\neg P)$. The problem appears only when we specify the whole distribution and assign those contents a certain exhaustive role.

That point matters independently of the original paradox. It shows that there is no simple recombinational principle for truth-assessable attitudes according to which any locally intelligible representational item may be freely combined with any others. Some total combinations fail, not because they already contain an explicit contradiction, but because their global organization closes off the representational space in an impossible way.

8. The classical moral

The difference between the present proposal and its main alternatives should now be clear. Priest shows how a dialetheist can respond: adopt paraconsistent logic, block the crucial

inferences, and under suitable conditions treat the paradoxical outcome without triviality. That is a coherent and sophisticated response.

Field's broader framework points toward a different family of options: retain a broad conception of propositions and properties and revise the logic or related principles that govern paradoxical contents. That too is a serious path.

The present paper takes a third route. It leaves classical logic in place and asks what the paradox teaches on that assumption. The answer is that the paradox reveals a restriction on representational possibility:

No distribution of truth-assessable attitudes can have both sides of a contingent divide exhausted by self-excluding strengthenings.

That principle applies equally to the split case $\{P, \neg P\}$ and to the compressed case $P \vee \neg P$. On this view, the original paradox is not evidence that classical logic needs revision. It is evidence that some total arrangements of thought, assertion, utterance, and related truth-assessable activities are impossible.

That is why the lesson is philosophical rather than merely defensive. The paradox does not merely force the classical theorist to reject one awkward world-description. It reveals a general structural law governing the organization of representation.

9. Conclusion

The paradox with which I began can be read in at least three ways. It can be given a dialetheic treatment, along Priest's lines. It can be placed within a broader non-classical theory of propositions and properties, along lines suggested by Field. Or it can be given a classical treatment.

I have argued for the third option. But the classical lesson is not merely that the original world-description is impossible. The deeper lesson is that the paradox instantiates a more general prohibited structure. No total distribution of a truth-assessable attitude can have both sides of a contingent divide exhausted by self-excluding strengthenings. The original two-sided belief case is one instance of that law. The one-person disjunctive case is another.

If that is right, then the philosophical significance of the paradox extends beyond the original derivation. It shows that representational possibility is globally constrained. Some total arrangements of truth-assessable attitudes are impossible even though their local parts appear intelligible. The lesson is therefore not that classical logic fails. It is that representation has limits.

References

Field, H. (2026). Propositions, properties and paradox. *The Australasian Journal of Logic*, *23*(1), 1-26.

Myers, B. (2019). A Paradox Involving Representational States and Activities. *Thought: A Journal of Philosophy*, *8*(2), 96-100.

Priest, G. (2021). Myers' paradox. *Thought: A Journal of Philosophy*, *10*(2), 147-154.

Williamson, T. (2018). Reply to Bacon, Hawthorne and Uzquiano. In *Williamson on Modality* (pp. 90-95). Routledge.