

Counterfactuals, Blame, and Counterfactual Blame

Gus Turyn

13 November 2025

The puzzle

Explaining away
the puzzle

Standing to blame

Encroachment

My proposal

Counterfactual
blame

The puzzle

The puzzle

Explaining away
the puzzle

Standing to blame

Encroachment

My proposal

Counterfactual
blame

Suppose that someone said the following:

- (1) If I had lived in the US in the 1940s, I would have watched Citizen Kane in theaters.

Ordinarily, we would defer to someone who said (1); this is part of a broader pattern of deference to others' *first-person authority* about their own intentional actions.

The puzzle

Suppose that someone said the following:

- (2) If I had lived in Nazi-occupied Poland in the 1940s, I would have hidden my persecuted neighbors from the SS.

Ordinarily, we wouldn't defer to someone who asserted (2). But curiously, it seems like we have roughly the same evidence about the two counterfactuals.

Real cases

Counterfactuals,
Blame, and
Counterfactual
Blame

Gus Turyn

The puzzle

Explaining away
the puzzle

Standing to blame

Encroachment

My proposal

Counterfactual
blame

Wahlberg on 9/11:

If I was on that plane with my kids, it wouldn't have went down like it did. There would have been a lot of blood in that first-class cabin and then me saying, 'Okay, we're going to land somewhere safely, don't worry.'

Trump on Parkland:

You don't know until you test it, but I think—I really believe—I'd run in there even if I didn't have a weapon.

The puzzle

The puzzle

Explaining away
the puzzle

Standing to blame

Encroachment

My proposal

Counterfactual
blame

This isn't just about the fact that (2) has moral content and (1) doesn't:

(3) If I'd lived in occupied Poland in the 1940s, I would've turned my persecuted neighbors over to the SS.

We would react negatively to an assertion of (3), but not skeptically.

The puzzle

Counterfactuals,
Blame, and
Counterfactual
Blame

Gus Turyn

The puzzle

Explaining away
the puzzle

Standing to blame

Encroachment

My proposal

Counterfactual
blame

My question: if we have the same kind of evidence about (1), (2), and (3), why are we more skeptical of (2)?

Clarifications

I'm interested in counterfactuals about which we deliberate *practically* rather than *theoretically*.

- (4) An oracle recently gave me some excellent news: if I'd lived in Nazi-occupied Poland in the 1940s, I would have hidden my persecuted neighbors from the SS.
- (5) If I had lived in Poland in the 1940s, I wouldn't have hidden my persecuted neighbors from the SS. I would certainly have wanted to, but I would have crumbled under the pressures that the Poles who did so faced.

Clarifications

This also isn't just about counterfactuals that concern historical events:

- (6) If I lived in Russia, I would dedicate my life to visiting as many art museums and historical sites as I could.
- (7) If I lived in Russia, I would dedicate my life to exposing Putin's corruption and misdeeds.
- (8) If I lived in Russia, I would spend my life helping Putin remain in power.

Explaining away the puzzle

Counterfactuals,
Blame, and
Counterfactual
Blame

Gus Turyn

Maybe asserting (3) reveals something about the speaker's character?

Maybe we should assign different prior probabilities based on the proportion of people who ϕ ed?

Maybe it has to do with the difficulty or costliness of the actions described in sentences like (2)?

Maybe it's easier to know facts about people's casual preferences than their deeply held values?

Maybe it's about speech acts or conversational dynamics?

The puzzle

Explaining away
the puzzle

Standing to blame

Encroachment

My proposal

Counterfactual
blame

Standing to blame

Counterfactuals,
Blame, and
Counterfactual
Blame

Gus Turyn

The puzzle

Explaining away
the puzzle

Standing to blame

Encroachment

My proposal

Counterfactual
blame

Much recent work on blame has to do with whether a blamer has the moral standing to blame another person for some wrong action. A number of conditions on standing have been proposed; I'll focus here on the anti-hypocrisy condition.

It should be clear that counterfactuals like (2) are relevant to our standing to blame: they're relevant to whether we're sufficiently committed to the relevant values at hand (Todd) and they're relevant to whether we're open to *tu quoque* charges (Lippert-Rasmussen).

Standing to blame

Counterfactuals,
Blame, and
Counterfactual
Blame

Gus Turyn

The puzzle

Explaining away
the puzzle

Standing to blame

Encroachment

My proposal

Counterfactual
blame

Markovits (ms.) argues that the truth of counterfactuals like (2) directly determines whether we have the standing to blame.

Whether counterfactuals like (2) are true is relevant to whether we have standing to blame people who didn't hide their neighbors from the Nazis.

The train cases

Counterfactuals,
Blame, and
Counterfactual
Blame

Gus Turyn

The puzzle

Explaining away
the puzzle

Standing to blame

Encroachment

My proposal

Counterfactual
blame

Low Stakes: Anya is in Boston waiting for a train to Providence. She overhears the following conversation between two passengers waiting on the platform:

Ben: Does this train stop in Foxboro?

Carol: Yes, it does. They told me so when I bought the ticket.

It doesn't matter much to Anya whether the train stops in Foxboro; she's heading to Providence, after all. She forms the belief that the train will stop in Foxboro.

The train cases

Counterfactuals,
Blame, and
Counterfactual
Blame

Gus Turyn

The puzzle

Explaining away
the puzzle

Standing to blame

Encroachment

My proposal

Counterfactual
blame

High Stakes: Dennis is in Boston waiting for a train to Foxboro. It's imperative that he arrives there before an important job interview at noon—his career depends on it—and he knows that if he boards the wrong train, he won't arrive on time. He overhears Ben and Carol's conversation and thinks to himself: *She might be misremembering or the salesperson could have been misinformed. I'd better ask the conductor before I board the train.*

My proposal

Counterfactuals,
Blame, and
Counterfactual
Blame

Gus Turyn

These two facts together can explain why we believe (1) and (3) but not (2), even when we have the same evidence about all (3): the moral stakes are higher because (2) positions its speaker to have standing to blame others.

My proposal is not that this is done *knowingly* or *intentionally*; I am not making a psychological conjecture about why people assert things like (2).

It's clear why counterfactuals like (1) aren't high-stakes, but why aren't counterfactuals like (3)?

When one asserts (3), one does not thereby position oneself to blame those who turned their neighbors over to the Nazis.

The puzzle

Explaining away
the puzzle

Standing to blame

Encroachment

My proposal

Counterfactual
blame

Why like my proposal?

Counterfactuals,
Blame, and
Counterfactual
Blame

Gus Turyn

It lets us have our cake and eat it too.

If you like any of the explanations I argued against before, you can still appeal to those to explain particular cases.

At a certain level of description, we'll appeal to worries about difficulty or costliness; at another, we'll appeal to the conjunction of encroachment and the facts about standing to blame.

There's a parallel here with the train cases: we can appeal to the possibility that the person on the platform gave the wrong answer at one level and encroachment at another level.

The puzzle

Explaining away
the puzzle

Standing to blame

Encroachment

My proposal

Counterfactual
blame

Counterfactual blame

We would respond negatively (though non-skeptically) to ordinary assertions of (3). Is this negative reaction a form of blame?

Shoemaker (2024) identifies a form of blame that he calls *counterfactual blame*:

(9) That's something *you* would do!

This isn't blame for a person's character or attitudes; it's blame for how they would have acted under those circumstances.

The puzzle

Explaining away
the puzzle

Standing to blame

Encroachment

My proposal

Counterfactual
blame

Counterfactual blame

Counterfactuals,
Blame, and
Counterfactual
Blame

Gus Turyn

The puzzle

Explaining away
the puzzle

Standing to blame

Encroachment

My proposal

Counterfactual
blame

My suggestion: it's theoretically fruitful to think of our responses towards counterfactuals like (3) as instances of counterfactual blame.

This is true despite the fact that there are some differences between these cases and Shoemaker's.

Recall (5): it's not just about the counterfactual state of affairs, but also about the ways in which we reason about those states of affairs.

Thank you!

Counterfactuals,
Blame, and
Counterfactual
Blame

Gus Turyn

The puzzle

Explaining away
the puzzle

Standing to blame

Encroachment

My proposal

Counterfactual
blame

Thanks to Ryan Bollier, Jaime Andres Fernandez, Julia Markovits, Shaun Nichols, Joseph Orttung, Carlotta Pavese, David Shoemaker, Geoffrey Weiss, and Hannah Winckler-Olick.