Synthese (2023) 201:165
https://doi.org/10.1007/s11229-023-04165-8

ORIGINAL RESEARCH

Stubborn emotions, stubborn beliefs

Kate Pendoley[1]

Received: 6 October 2022 / Accepted: 17 April 2023 / Published online: 1 May 2023

Abstract
Recalcitrant emotions, such as fear of flying, are emotions that persist even though they are in tension with the emoter's considered belief. A widely accepted argument against cognitivist theories of emotion holds that recalcitrant emotions show that emotions are more like sensory states than like thoughts or beliefs. I show that this argument does not succeed: Emotions are usually sensitive to our changing beliefs in a way that is more akin to cognitive states than to sensations. Moreover, empirical evidence strongly suggests that beliefs sometimes behave much like recalcitrant emotions do.

Keywords Emotion · Encapsulation · Recalcitrance · Cognitivism

1 Introduction

Imagine that you, a person in possession of a normal degree of rationality, see your neighbor walking their dog, Fido, on the other side of the street. You know Fido and believe him to be very gentle. His owner has given assurances that he is quite friendly, and Fido has never demonstrated any aggression in your past meetings. Your considered belief—the belief you could provide reasons for and which you would stand behind—is that Fido is not dangerous. Nevertheless, upon seeing the dog you are frightened. You are so afraid that you begin to tremble and even cross the street to avoid him. This is the familiar phenomenon of stubborn or 'recalcitrant' emotion.

Recalcitrant emotions highlight the ways that emotions seem to be beyond our rational control, come unbidden, have us at their mercy. You are afraid of the dog—afraid that the dog will bite you, perhaps—*while you simultaneously believe* that the dog is gentle. These two mental states, your emotion and your belief, seem to present the world in two incompatible ways. And your thought does not change your fear, which persists even in the face of your belief that the dog is gentle and nonthreatening. Given the rational nature of belief, if you can think that the dog is not dangerous while feeling

✉ Kate Pendoley
katependoley@gmail.com

[1] Vassar College, Poughkeepsie, USA

afraid of the dog, then emotions and beliefs must be two completely different types of mental states.

It is compelling and intuitive to think that recalcitrant emotions drive a wedge between emotion and cognitive states and show that emotions are not even partly composed of thoughts or beliefs. This is important, because it might otherwise seem attractive to explain the intentionality of emotion by saying that emotions at least partly consist in thoughts or beliefs. Like these paradigmatically cognitive states, emotions are about objects and states of affairs that may be distant in both time and place from the emoter: You are happy that you are visiting friends, sad that your favorite bookstore has closed. So, emotions look like they are ways of relating to a structured content, just like mental attitudes such as belief are. Cognitivist views of emotion can account for this easily: According to these views, emotions either are judgements or beliefs, or else have these cognitive states as constituent parts.[1] However, recalcitrant emotions that persist in spite of your considered beliefs seem to reveal that emotions do not even have beliefs as constituent parts, since this recalcitrance seems to many to be incompatible with belief. Some version of this argument is widely endorsed as decisive against cognitivist views of emotion.[2]

In spite of its intuitive power, the argument from recalcitrance has proven difficult to pin down precisely.[3] One important and promising formulation of this argument draws an analogy with perceptual illusions.[4] However, a more nuanced picture of belief has emerged from empirical work, revealing that many of our intuitions about rationality and belief are likely idealized and mistaken. Drawing on this work, I argue that the formulation of the argument from recalcitrance which draws a close analogy with perceptual illusion fails because it relies on an incorrect picture of belief. Emotions

[1] Here, I count as cognitivist views which would reduce emotions to (or identify them with) cognitive states such as judgment or belief (Nussbaum, 2004; Solomon, 1993), as well as views which claim merely that a cognitive state such as belief or judgment is a constituent part of emotion, such as belief-desire accounts (Gordon, 1994; Green, 1992; Reisenzein, 2009). Some of what are termed 'Neo-Cognitivist' views (which often position themselves between cognitivist theories and perceptualist theories that seek to align emotions with perceptual states) will fall in this second camp. However, some Neo-Cognitivist views which hold that emotions have propositional content might not be vulnerable to the formulations of the argument from recalcitrance I discuss here. For instance, attitudinal accounts often allow that emotions have propositional content, but whether attitudinal views are vulnerable to these arguments from recalcitrance will depend on the way they specify the nature of the attitude in question (see Helm, forthcoming, for helpful discussion). According to Deonna and Teroni (2012, 2017), the attitude in question is the emoter's experience of their body as ready for a particular action. This kind of attitude is unlikely to be vulnerable to the argument from recalcitrance (though note that Deonna and Teroni do not spell out the content of emotion propositionally, either). See D'Arms and Jacobson (2003), as well, for a discussion of what they dub 'quasi-judgmentalist' theories. See Grzankowski (2020) for helpful discussion and taxonomy of cognitivist, neo-cognitivist, and perceptualist views in the literature.

[2] See, for instance, Deonna and Teroni (2012, pp. 54–55), D'Arms and Jacobson (2003, p. 129), Benbaji (2013), Döring (2014), Naar (2018), Grzankowski (2017), Greenspan (1981), Helm (2015), Prinz (2004), Brady (2009), Tappolet (2012, 2016).

[3] See Grzankowski (2017) and Naar (2018) for helpful overviews.

[4] There is another, slightly different formulation advanced in Grzankowski (2017), which improves on Greenspan's (1981) formulation. For brevity I will not include a separate discussion of this formulation, since it relies on the same idealized notion of belief, and hence fails for the same reasons as do the two formulations of the argument from recalcitrance that I discuss in Sects. 2 and 5.

may turn out to be among, or to have as constituent parts, the most paradigmatically cognitive mental states we have, after all.

The paper is organized as follows. Section 2 sets out the first formulation of the argument from recalcitrance via analogy with perceptual illusion. Section 3 draws upon empirical work and recent discussion thereof to show that beliefs can exhibit recalcitrance, too, and hence that the encapsulation formulation of the argument from recalcitrance fails. Section 4 considers and responds to two potential objections to the claim that beliefs are sometimes recalcitrant. In Sect. 5, I consider a different sort of argument that recalcitrance poses a problem for cognitivism, before raising several considerations against that line.

2 Recalcitrance as evidence of encapsulation

Recalcitrant emotions look a good deal like perceptual illusions. When you are driving on the highway and seem to see a large puddle in the distance in the middle of the road—even when you know that it is an optical illusion due to heat rising off the pavement—your belief that the puddle is illusory does not make the visual appearance of the puddle go away. It still *looks* to you as if there is a puddle there, even though you know better. Recalcitrant emotions seem very much like your visual perception of the illusory puddle: In recalcitrant emotion, as in known perceptual illusions, one of your mental states continues to present the world one way in spite of conflicting with another mental state that presents the world differently. And so, it may seem natural that analogies with recalcitrant perceptual illusions such as the classic Müller-Lyer illusion are everywhere in discussions of emotional recalcitrance.[5]

Persistent perceptual illusions may be a reason for thinking that perception and cognition are distinct mental systems. The illusion above persists even when you believe that the puddle doesn't exist or that the lines in the figure are the same length, and so the perception must not be able to access the information that belief is drawing upon: your perceptual state is *informationally encapsulated* from your belief to the contrary. If one processes is encapsulated from another, Process A cannot access information stored in Process B. Informational encapsulation has seemed to some a good reason for positing that two processes belong to distinct mental systems (Fodor, 1983, p. 69).

It appears that recalcitrant emotions likewise show that emotion is or belongs to a distinct mental system from cognition: The persistence of some emotions in the face of our considered beliefs to the contrary looks like evidence that emotions, like perceptions, are encapsulated from cognition. Several authors in the literature appear to have an argument along these lines in mind when they mention persistent perceptual illusions in connection with recalcitrant emotions, even though most do not spell it out.[6]

Accordingly, this formulation would be: Recalcitrant emotions reveal that emotions are informationally encapsulated from cognition; mental state A's being encapsulated

[5] See Döring (2014, p. 132), for an example.

[6] See, for example, Döring (2014), as well as Tappolet (2012, 2016), Prinz (2006).

from mental state B provides a good reason for positing that A and B belong to distinct systems; so recalcitrant emotions give us a good reason to think that emotion is a distinct system from cognition, and thus that emotions do not have beliefs as even necessary parts. Each form of cognitivism about emotion holds that belief is at least a necessary part of emotion, and so cognitivism is false.

Let's spell this formulation out explicitly:

The Encapsulation Formulation of the Argument from Recalcitrance:

1. Recalcitrant emotions show that emotions are encapsulated from cognition (information available to cognition is not available to emotion).
2. If a mental state or process is encapsulated from cognition, that mental state is not cognitive.
3. (2) is incompatible with the claim that beliefs are at least necessary parts of emotions.
4. Thus, cognitivism about emotions is false.

This formulation has several advantages over other formulations of the argument that focus on the *degree* to which having recalcitrant emotions is irrational. First, the encapsulation formulation makes the notion of recalcitrance precise. Moreover, this formulation makes it much clearer what is meant by 'cognitive:' It is a minimal criterion for a state to be cognitive that it is *not* encapsulated from cognition.[7]

This argument faces two problems. First, not all token emotions are recalcitrant; occasional recalcitrance doesn't show that emotions as a system are encapsulated from cognition. Second, beliefs exhibit stubbornness, too. If stubbornness is evidence of encapsulation, then at least some beliefs are encapsulated. I explain these points in the following section.

3 Two problems for recalcitrance as evidence of encapsulation

3.1 Many instances of emotions are sensitive to changing beliefs

Emotions are sometimes recalcitrant,[8] but they aren't *always* recalcitrant. We would be very different beings if all—or if even most—of our emotions were recalcitrant. Instead, emotions often change in light of changing beliefs in a way that suggests that many instances of emotion are sensitive to belief. If every instance of emotion were recalcitrant, you would not ever transition from fear to relief upon learning that you passed a crucial exam that you had antecedently believed you had failed. As Solomon notes, if I'm angry that you've stolen my car but then learn that my car has not been stolen after all, my anger will sometimes persist, but in many cases will dissipate

[7] This formulation of the argument has the added bonus of engaging closely with Feldman-Barrett's (2017) challenge that, contra basic emotions theorists, emotions are not modular. However, whether encapsulation implies modularity is debated and nothing I say here takes a stance on this issue.

[8] Note that even recalcitrant emotions seem in some instances laudatory: Emotions sometimes urge an emoter to act or to adjust their considered beliefs—and in some cases this seems to be a good thing. Recall the oft-cited example of Huck Finn, who hides Jim out of emotion but in spite of Huck's considered judgment. This case is meant to show that not all recalcitrant emotions should yield to an emoter's standing, considered beliefs and is discussed, for instance, in Döring (2014), and Benbaji (2013).

(1976, p. 119). It might even be replaced with relief if, for instance, my discovery of the car's location is the way I learn that my car hasn't been stolen. If all occurrences of emotions were recalcitrant, it would be difficult to make sense of the way emotions are often sensitive to changing beliefs.

In any case, it is not obvious that emotions are recalcitrant *so often* that we should consider them to be encapsulated from cognition.[9] Occasional instances of recalcitrant emotion are not on their own enough to show that emotions as a system must be encapsulated from cognition. However, even if the advocate of the argument from recalcitrance grants this point, they might persist by pointing out that beliefs do not *ever* exhibit recalcitrance, and hence that even occasional recalcitrance of emotion is a reason for rejecting cognitivism and perhaps for drilling down into the claim that emotions are encapsulated from cognition. However, as I'll show in the following section, this response won't go through. Beliefs are sometimes stubborn in much the way emotions are.

3.2 Empirical evidence for stubborn beliefs

An extensive body of empirical work demonstrates that we do not always update our beliefs in accordance with classical ideals of rationality.[10] People often persist in their standing beliefs in the face of good evidence to the contrary (Anderson, 1983; Anderson & Sechler, 1986; Anderson et al., 1980; Slusher & Anderson, 1989); and equivocal evidence may even strengthen, rather than moderate, preexisting beliefs.,[11,12] Additionally, people tend to selectively expose themselves to congenial over uncongenial data, vigorously argue against evidence that challenges their existing beliefs, and judge evidence that supports their preexisting views as stronger than evidence that challenges them.[13]

More specifically, robust and well replicated empirical work shows that beliefs are sometimes stubborn in much the way that emotions sometimes are: Some beliefs persist in spite of countervailing evidence. Take, for instance, the phenomenon of

[9] It is important to note here that whether and the extent to which encapsulation can divide perception from cognition is not settled. The extent to which perception is encapsulated from cognition is hotly debated, especially in light of cognitive penetration (the top-down influence cognition often has on perception). See e.g. Prinz (2006); Clark (2015). Hence, the role that encapsulation ought to play in drawing the border between perception and cognition is controversial, never mind how well this approach applies to questions about the divisions between emotion and cognition. Some writers in the literature on the perception-cognition border have suggested that encapsulation may come in degrees; perhaps this provides a more promising way to use encapsulation to divide perception from cognition. See Pacherie (2008); Shea (2015); Quilty-Dunn (2017).

[10] See Kunda (1990) and Kuhn and Lao (1996) for an overview.

[11] See, for example, Lord et al. (1979); Nyhan and Reifler (2010); and Nyhan et al. (2013).

[12] Eric Mandelbaum's work first drew attention to these examples, as well as those discussed later in this section, and to the pressure they put on the view that beliefs update reliably in an ideally rational way. Lord et al. (1979) was first discussed in the philosophical literature on belief by Mandelbaum (2013); Johnson and Seifert (1994) by Mandelbaum and Quilty-Dunn (2015); and Anderson (1983), Anderson et al. (1980), Anderson and Sechler (1986), Batson (1975), Lord et al. (1979), Slusher and Anderson (1989), and Taber and Lodge (2006) by Mandelbaum (2019). See also discussion of many of these studies in Bendaña (2021) and Bendaña and Mandelbaum (2021).

[13] See Taber and Lodge (2006); and Tappin et al. (2021).

misinformation persistence. People are evidently troublingly poor at updating their beliefs even once they learn that the information on which those beliefs are based is false.

As Wilkes and Leatherbarrow's (1988) experimental paradigm for misinformation persistence shows, people who are otherwise typical reasoners often fail to revise their beliefs even once they learn that the information which they had previously accepted is false. For instance, in one iteration of this well replicated paradigm, Johnson and Seifert (1994) gave study participants a series of false news articles about a warehouse fire. The fifth article in the series stated that there were cans of oil paint and gas cylinders stored in the room where the fire started. In one of the experimental conditions, participants were informed immediately after they read this fifth article that there were in fact no cans of oil paint or gas cylinders stored in the room where the fire started. In the other condition, subjects didn't receive this corrective until they had read the entire sequence of fake news articles about the warehouse fire. The participants completed a brief distractor task and were then given a questionnaire that was designed to determine what they believed was the cause of the fire. Over ninety percent of subjects in both misinformation conditions remembered the correction and reported that they accepted this correction, but nevertheless ninety percent used the misinformation in their reasoning as they answered the questionnaire.[14]

The subject in these trials gives us good reason to believe that they hold incompatible beliefs: both the correction they accept, as well as its contradiction, which they use in their reasoning in answering the questionnaire. The subject doesn't have any other properties by which to pick out the gas cylinders, etc., outside of the details the experimenter has provided, and presumably isn't thinking of two different sets of gas cylinders. So, the subject looks like they maintain two contradictory beliefs. This study is especially nice because it is not straightforwardly a case of emotionally-laden motivated reasoning. It doesn't bear on participants' political or personal beliefs, nor even many of their more mundane beliefs. Study participants don't come into the study with antecedent beliefs about the case, and they likely don't have much of an antecedent bias toward the misinformation.

Does this amount to evidence that ordinary reasoners have recalcitrant beliefs? A critic might object that misinformation persistence is importantly weaker than recalcitrant emotion: In cases of misinformation persistence the subject may not be aware of both beliefs. If a subject isn't aware of two beliefs, it makes sense to think they won't be aware that the two beliefs are incompatible, either. Meanwhile, on a cognitivist reading of recalcitrant emotion, the subject not only has two incompatible beliefs, but is aware of them both, and is aware of the tension between the two states.

If this critique of the misinformation persistence findings holds up, then a more precise characterization of recalcitrance emerges. In order to count as recalcitrant, a belief will not only need to be (a) incompatible with another belief the subject holds, but it will also have to be the case that (b) the subject is aware that the beliefs are incompatible.[15] Even if this is enough to reject that the misinformation studies are

[14] As discussed in Bendaña (2021) and Bendaña and Mandelbaum (2021); see for more detailed discussion.

[15] For the purposes of this paper, in (b) 'awareness' is meant to be first-personal awareness. I take it that many people will think that we are under less rational pressure to revise incompatible beliefs if we learn

evidence of recalcitrant belief, this potential objection guides us nicely toward a more precise characterization of recalcitrance. The task is now to show that beliefs exhibit (a) and (b).

Indeed, one form of belief perseverance shows just that. In the backfire effect—sometimes termed the backlash or boomerang effect—beliefs get stronger in the face of disconfirming evidence, rather than weakening as one would expect. The effect was recently found by Nyhan and Reifler (2010), among others, and is nicely illustrated in the following experiment.

In Batson's (1975) experiment, participants were divided into those who believed that Jesus was the son of God and those who believed that he was not. Both groups were then given what they were told was an unpublished article that was "denied publication in the New York Times at the request of the World Council of Churches because of the obvious crushing effect it would have on the entire Christian world" (Batson, 1975, p. 180). According to the unpublished article, archeologists had brought new, decisive evidence to light that proved that the authors of the New Testament themselves knew that Jesus wasn't the son of God.

After reading the article, the participants reported whether they believed its claims. They then took another test to track whether their belief that Jesus was the son of God had changed. Unsurprisingly, participants who antecedently did not believe that Jesus was the son of God and reported that they believed the claims of the article maintained and strengthened their belief that Jesus was not the son of God. Similarly, those who did antecedently believe that Jesus was the son of God and did not believe the article's claims about the authors of the New Testament did not change their belief: This group continued to believe that Jesus was the son of God. But there was a group of participants who antecedently believed that Jesus was the son of God who also reported that they believed the article's report to be true. For all of the participants in this last group, their belief that Jesus was the son of God was *strengthened* even though they also reported that they accepted the claims of the article stating that Jesus was not the son of God. They reported that they had just received good evidence that not-P and accepted that not-P, and yet the strength of their belief in P increased in spite of the very evidence they accepted and in spite of their new reported belief that not-P.[16]

The subjects reported both beliefs in quick succession. Participants in this last group have (a) incompatible beliefs and (b) seem to be aware of that incompatibility. One might even think that what's happening in the backfire effect is *more irrational* than what we see in recalcitrant emotion: In the backfire effect, the subject not only has two incompatible beliefs and is aware of their incompatibility, but one of the beliefs is *strengthened* by the acquisition of a contradictory belief.

This looks to be a compelling case of recalcitrant belief. If this is right, then some token beliefs are recalcitrant. However, this doesn't seem to be a good reason to think that beliefs are encapsulated *as a system* from cognition—they are beliefs, after all. The recalcitrance of some token beliefs does not establish that beliefs as a system

Footnote 15 continued
of the incompatibility third-personal means (by, for example, taking an implicit association test, contested though they may be) than if we learn of it first-personally.

[16] As discussed in Bendaña (2021).

are encapsulated from cognition. Or else, if stubbornness implies encapsulation, then at least some beliefs are encapsulated and encapsulation on its own is insufficient to establish that the state in question isn't cognitive.

At this point, a reader might worry that these cases of stubborn belief all look to be 'hot,' motivated by emotional concerns, and hence insufficiently strong as evidence of stubborn belief. However, there is reason to think that even 'cool' beliefs exhibit this stubbornness. Jern et al. (2014) find that belief polarization can occur even for 'cool' rather than only for emotionally charged 'hot' beliefs, where the effect is usually observed.[17]

Evidence of belief perseverance gives us good reason to accept that the some of the token beliefs of ordinary, actual reasoners are stubborn in a way that's closely analogous to the way that some episodes of emotions are. If this is right, then the stubbornness of a token mental state is not enough to show that this state cannot have a cognitive state even as a mere constituent part. Beliefs, stubborn or not, are cognitive states if anything is, and so the recalcitrance of token instances of a mental state type cannot show on their own that the state in question is not cognitive.[18]

There is compelling evidence that, from time to time, the beliefs of ordinary, actual human reasoners exhibit stubbornness akin to the stubbornness of emotion. If this is right, then the potential objection I considered at the end of the previous section does not go through: Appealing to recalcitrance of some instances of emotion is not enough on its own to show that emotion exhibits a feature that belief never does. Occasional recalcitrance of emotion is not enough on its own to tell against cognitivist views of emotion, and the encapsulation formulation of the argument from recalcitrance fails.[19]

[17] Though see Mandelbaum (2019) and Williams (2018) for critiques of these models.

[18] One might also worry that recalcitrance reveals that emotions must have non-propositional or non-conceptual content. The idea would be that states with propositional content don't exhibit resistance to rational revision, so a state that exhibits recalcitrance can't have propositional content. However, belief arguably has propositional content if any state does. If beliefs are sometimes stubborn, then stubbornness of token mental states does not reveal that the type of mental state to which the token belongs must have non-propositional or non-conceptual content. This is already acknowledged in the literature on the perception-cognition border, since architectural approaches allow that perception and cognition may turn out to use the same kinds of representations; their difference is to be drawn according to the ways the different types of states are structured and interact with one another, rather than according to their representational features. See, e.g., Pylyshyn (2003). For a nice discussion of perceptual pluralism, propositional content, and perceptual views of emotion, see Wringe (2015).

[19] There are other ways a cognitivist might avoid the problems purportedly associated with recalcitrant emotion than this paper considers. Nussbaum, for instance, argues that that in recalcitrant emotions a subject does not have incompatible judgments simultaneously, but rather toggles between them (e.g. Nussbaum, 2009, pp. 383–85). If this paper's argument goes through, however, no such toggling is required. Since this paper is concerned neither with evaluating nor with advancing a positive view of emotion, a full treatment of Nussbaum's view as well as of her solution to the problem of recalcitrance is beyond the scope of this paper.

4 Recalcitrance, belief, & rationality

4.1 Recalcitrance and the nature of belief

The success of Sect. 3.2 depends on the claim that the relevant mental states in the studies are actually beliefs. Some may want to resist this claim on the grounds that (a) having incompatible beliefs where (b) one is aware of having incompatible beliefs would violate the nature of belief.[20] It might intuitively seem that you can't have incompatible beliefs. If beliefs by their nature just aren't the sorts of things that can be recalcitrant, then the examples of what I have termed recalcitrant beliefs in the previous sections couldn't be cases of beliefs.

That line assumes that a token mental state must meet the ideal criteria of belief to count as a token belief, and that those criteria preclude recalcitrance. For instance, Gendler (2008) and Levy (2015) seem to agree that token beliefs must satisfy the following criteria:

i. Feature in normatively respectable transitions; and
ii. Be consistent with other beliefs.[21]

Perhaps (i) and (ii) reflect at least some norms of ideal rationality that govern belief and according to which we evaluate belief. And yet, that beliefs are governed by these norms doesn't require that each token belief satisfies (i) and (ii). Many have conflated the governing norms of belief with definitional criteria for counting as belief.[22] This is a mistake.

We rarely make it criterial for each token of a mental state to meet the normative ideals of its type. You can be in a perceptual state that is more or less accurate or precise. Suppose you have poor vision and aren't wearing any kind of corrective lens, for instance. You'll still have a visual experience, even if it's very blurry. Or try imagining you're sitting at your kitchen table over a cup of coffee. You might imagine some aspects of the scenario in great detail, and others vaguely. You might even just imagine the scenario vaguely, in very little detail at all, because your mind is wandering elsewhere (or perhaps because you're reading an academic paper). Your mental state need not represent all details of the imagined scenario in great detail in order for the state to be one of imagining. Similarly, when remembering, one can make lots of mistakes that fall short of confabulation before falling short of remembering at all. So even if a token state is subject to norms that govern its state type, the token can fall short of the norm without failing to be a token of the type altogether. Moreover,

[20] This assumption about the nature of belief is central to yet another formulation of the argument from recalcitrance. See, for instance, Greenspan (1980), Helm (2015), Benbaji (2013), Döring (2014), and Naar (2018).

[21] As nicely noted and discussed in Bendaña (2021).Even more permissive views, such as McLaughlin's (1988), don't allow for recalcitrant beliefs. Though McLaughlin allows that a subject can have incompatible beliefs, he denies that a subject could be *aware* of both incompatible beliefs. Instead, he has it that at least one of the beliefs must be "inaccessible" to the subject in cases of incompatibility.

[22] Criterialists bear the additional burden of positing new types of mental states to account for what I've been calling stubborn beliefs. Some posit 'aliefs' (Gendler, 2008) or merely very vivid imaginings. See Mandelbaum (2013) and Viedge (2018) for detailed treatments of the problems they see plaguing this kind of approach.

though many would agree that belief is governed by the norm of aiming at the truth, even the criterialists wouldn't make *arriving* at the truth criterial of every token belief. A token belief can be governed by the norm of aiming at the truth even when the belief is itself false.

The question is not whether a token state *meets* the governing criteria to count as belief, but whether it is part of a system that is answerable to those normative criteria.[23] Accordingly, we ought to allow that some token beliefs may fall short of ideal rationality. I have cited evidence that we have incompatible and even recalcitrant beliefs, in spite of what many intuitions about belief would like to admit.[24] If this is right, then there is evidence that some beliefs *do* in fact fall short of normative ideals such as (i) and (ii).

4.2 Recalcitrance and the nature of the evidence

Some might accept that some beliefs fall short of (i) and (ii), and also accept that we seem to have evidence that beliefs do not behave in the optimally rational way that intuition suggests, but nevertheless deny that the studies cited above show that beliefs rise to the level of recalcitrance. Current critical debate merits some discussion here. There isn't yet consensus about the extent of the backfire effect.[25] Some recent studies have failed to replicate it, and others have found the effect to be highly contextually constrained.[26] However, backfire has been observed even in even 'cool' and unemotional cognition, suggesting that the effect is in fact *not* especially constrained.[27] Moreover, some of the criticism lodged at the robustness of the backfire effect focuses on the fact that the effect is more likely to be observed in behavior rather than in self-report. But behavior is arguably a more reliable measure of belief in this domain. The cost of reporting one's beliefs is lower than the cost of changing one's other behavior, after all, not to mention that self-report is notoriously fraught.

Note, too, that the backfire effect is more extreme than recalcitrant belief would seem to require: The subject's antecedent belief does not just persist but is *strengthened* by evidence to the contrary. Beliefs need not meet the backfire threshold in order to count as stubborn. Hence, even if backfire is mooted, we still have good evidence of stubborn beliefs.

Most importantly, the backfire effect is just one species of belief perseverance. Other perseverance effects, such as confirmation bias in the evaluation of evidence,

[23] c.f. Huddleston's (2012) argument that 'naughty' beliefs are still beliefs.

[24] Some people who take the line that beliefs could never be recalcitrant might still be comfortable admitting that these studies show that beliefs are sometimes evidence-resistant. Perhaps the critic of would insist that nevertheless the intuitions about the *coherence* of belief must be right, even if beliefs are not always as evidence-responsive as many intuitively take them to be.There are thorny issues about the relationship between evidence-responsiveness and coherence. Though a treatment of this issue is beyond the scope of this paper, some argue that coherence is at bottom a matter of evidence-responsiveness. See for instance Kolodny (2007, 2008) and Broome (2013).

[25] Thanks to an anonymous reviewer for pressing me on this issue.

[26] See Tappin and Gadsby (2019) for review and discussion.

[27] See Jern et al. (2014).

are robustly replicated.[28] People are more likely to rate information that confirms, or is consistent with, their prior beliefs as stronger than information that contradicts those prior beliefs.[29]

Insofar as perseverance is a form of stubbornness—your belief persists in the face of countervailing evidence and so is not evidence-sensitive in the way we expect beliefs to be—the work on belief perseverance still suggests that beliefs are sometimes stubborn in much the way that emotions sometimes are. As long as you accept that the states in these studies are beliefs, you should accept that beliefs turn out to behave in a way that falls far short of the intuitive notion of rational belief. If that is right, then we should rely on empirical study rather than intuitions to ask whether beliefs sometimes also satisfy (a) and (b), above, and rise to the level of recalcitrance.

It is also worth noting that there is also active debate over whether and the degree to which belief perseverance poses a problem for a Bayesian picture of belief.[30] Some argue that Bayesian principles are consistent with, or even entail, belief perseverance.[31] Jern et al. (2014) model, for instance, entails biased assimilation effects.

However, my point does not hang on how the debate over Bayesianism about belief resolves. Both sides of that debate grant that beliefs exhibit perseverance. They differ only on whether a Bayesian picture of belief can satisfactorily account for belief perseverance. Nevertheless, those arguing that belief perseverance is rational in the sense that it is consistent with Bayes actually provide support for my position. Here's why: Some may think that stubborn beliefs are always the result of 'hot' cognition about emotionally-charged issues, so one might worry that belief on its own doesn't exhibit stubbornness but does so only under the influence of emotion. But if belief perseverance is consistent with Bayesian principles, and if even emotionally 'cool' beliefs exhibit stubbornness, then this worry is blocked: Stubborn beliefs are not to be explained exclusively by the interference of emotion.

None of this is to deny that there is rational failure in cases of stubborn belief. Things certainly seem to have gone wrong from a rational point of view when a person has a stubborn belief. But without careful attention to the ways that beliefs actually behave, we won't have a full grip on the limits on belief, given that beliefs turn out to violate several stated intuitions about how they behave.

5 Norms of rationality, belief, and emotion

A different formulation of the argument from recalcitrance focuses on a more straightforwardly normative concern, beginning with the observation that we usually don't judge a person in the grip of a recalcitrant emotion to be seriously irrational, while we do ordinarily judge a person who has conflicting beliefs to be *very* irrational. Many

[28] Tappin and Gadsby themselves note that this effect is "extremely robust" (2019, p. 110).

[29] See, for example Lord et al. (1979), Taber and Lodge (2006), Tappin et al. (2021).

[30] Some of this debate also concerns the degree to which belief perseverance might be rational. This is largely beside the point for what I argue here; for brevity, I leave it aside.

[31] See Tappin and Gadsby (2019) and Baron and Jost (2019) for recent reviews in support of their claim that many of cases of purportedly biased or motivated cognition can be accommodated by a Bayesian model, given appropriate right priors and likelihoods.

people are familiar with feeling afraid of something that they believe is safe, or guilty about doing something that they do not really believe was wrong, and so it is very understandable to have recalcitrant emotions. And, the argument goes, people typically find it more irrational to have stubborn beliefs than to have recalcitrant emotions: If you have had plenty of evidence that *p*, but you still believe that *-p*, or if you have beliefs that are mutually incompatible, then intuitively you are being very irrational. So, the thinking goes, people tend to judge those in the grip of recalcitrant emotions as less irrational than they would judge people who were in the grip of similarly stubborn beliefs. But if cognitivism is right and emotions are beliefs, are reducible to beliefs, or at least have beliefs as constituent parts, then a person who has a recalcitrant emotion would, counterintuitively, be *just as irrational* as someone who has stubborn beliefs. On this line, the primary problem with cognitivism's handling of recalcitrant emotions is that cognitivism makes the emoter out to be more irrational than intuition says they are.[32]

A complete treatment of this formulation is beyond the scope of this paper, but I will point toward some potential ways of responding, here. First, there is at least one instance of conflicting beliefs that it seems eminently rational to hold. Second, at least some instances of recalcitrant emotions resemble this case of rational but conflicting beliefs. Third, I think we should be circumspect about how much everyday judgments of comparative irrationality reveal about the structures of the states in question. There will not be sufficient space to do justice to this final point, but I will give some initial considerations in support of it.

But first, an instance of rationally holding conflicting beliefs. There is one classic case of belief that clearly rises to the level of recalcitrance, but which nevertheless appears eminently reasonable. In the Preface Paradox,[33] it seems perfectly rational to hold two mutually incompatible beliefs, even though one knows that the beliefs are incompatible. Suppose you write a very long book. You are a diligent and conscientious researcher and writer, so you've double-checked each reference carefully and cannot find an assertion you don't believe to be true. At the same time, you are intellectually humble and know that in a book this long, no matter how careful your fact-checking, it is unlikely that *every* claim is correct. Thus, as you write the preface to your book it seems very rational to say that you believe of every individual assertion that it's true, but that you do not believe the conjunction of the book's assertions to be true. This is a case where many would be happy to say that the believer has two conflicting beliefs but is nevertheless quite rational. So, there is at least one instance where most would be happy to say that a person is not irrational in holding conflicting beliefs.

Moreover, at least some cases which we might be inclined to deem recalcitrant emotion resemble the Preface Paradox more closely than they resemble the stubborn beliefs discussed in Sect. 3. Cases of recalcitrant fear tend to dominate the discussion of recalcitrant emotion, and fear tends to be forward-looking in Gordon's (1994) sense: Fear is usually about states of affairs that have not yet come to pass, or which have passed but about which the emoter is not certain. Even when fear seems to be directed

[32] See Brady (2007, 2009), Helm (2001, pp. 41–46, 2015), Benbaji (2013), Döring (2014), Naar (2018), Grzankowski (2020); for discussion of this formulation of the argument from recalcitrance.

[33] First presented by Makinson (1965).

at an entity or person instead—as when someone says they are afraid of dogs, of planes, etc.—usually the fear is about what the entity in question might do or might have done (bite, crash, etc.).[34] In some instances, an emoter will have good reason to believe that a fearful outcome is unlikely, but will also have some reason to believe that the fearful outcome might occur. For instance, suppose you've just checked that you've turned off all the knobs on the stove. This gives you good reason to think that the burners are not emitting gas and that the stove is safe. However, you also believe that over the many occasions you've checked the knobs, it's unlikely that you'll never err. Suppose that, further, you've also gone through the very same procedure in the past only to return home and realize that you left the very last knob turned on and that the stove has been emitting gas.[35] In this case, it seems you have reason both to believe that the stove is off and to fear that it is on; there doesn't seem to be much in the way of rational failure. This parallel is not restricted to fear but should hold instead for any cases where you have considerations counting both for and against the relevant belief.[36]

Finally, what reasons do we have for judging that a person in the grip of recalcitrant emotion is less irrational than a person who has conflicting beliefs? Perhaps a person who has recalcitrant emotions trusts or endorses their belief but not their emotion, and hence is less irrational: They disavow the content of their emotion while endorsing the content of their belief, and this is less irrational than either endorsing two incompatible beliefs or endorsing a belief which persists in spite of evidence to the contrary.[37]

The assumption that people usually don't endorse their recalcitrant emotions requires careful treatment. While it may be true that people will usually verbally disavow their recalcitrant emotions and endorse the conflicting beliefs, it is a mistake to take this verbal behavior as decisive. This tendency to report trusting one's own beliefs over the emotions that conflict with them may itself reveal more about *attitudes* toward emotion than it does about one's actual degree of trust in their emotions versus trust in beliefs: Many people generally think reason and considered belief ought to outweigh their emotions, even—or perhaps especially—when the two states are in conflict. One may report endorsing their considered beliefs over their recalcitrant emotions even when their actions suggest that the emotion wins out.

Given that the behavior of some instances of belief turns out to violate many familiar intuitions about belief, as I discussed in Sect. 3.2, theorists should be circumspect about relying on intuitions about the comparative rationality of having recalcitrant emotions compared with holding conflicting beliefs. Moreover, these intuitions about comparative rationality might reflect attitudes toward emotion and belief, as well as

[34] With the possible exception of phobias, a full consideration of which is beyond the scope of this paper.

[35] And, finally, suppose that this case is not pathological.

[36] Additionally, if a possible outcome is sufficiently bad, perhaps it is in fact most rational to avoid that event if the cost of doing so isn't too high, even if the emoter believes the likelihood of the dangerous outcome is small. Perhaps for fear of flying, your belief that flying is relatively safe overrides your fear, maybe in part because you estimate that the costs of *never* flying are too high.

[37] Döring (2014) makes this suggestion, writing that when a person is in the grip of a recalcitrant emotion, "The subject does not contradict himself because he only regards his judgement's content as true, whilst the content of his emotion merely appears to be true to him." (p. 134).

practices of privileging belief over emotion, rather than revealing much about either the structure of the states themselves or which norms govern the respective states.

6 Conclusion

The paper does corrective work to show how we should go about investigating the place of emotions in the mind. We want to know which features of emotion can point us toward their place in the mind. Much of the literature has treated recalcitrance as if it is up to the task, especially in light of concerns over encapsulation and intuitions about the rationality of belief. I argue that this is a mistake. Advocates of the argument from recalcitrance have relied on recalcitrance and an analogy with persistent perceptual illusions to drive a wedge between emotions and belief, and to align emotions with perception over belief. I have argued that not all instances of emotion are recalcitrant: Unlike perceptions, many instances of emotion change in response to changing beliefs. Moreover, I have drawn attention to empirical evidence that suggests that some beliefs exhibit the very stubbornness that seems to be characteristic of recalcitrant emotion. Hence, it is unlikely that emotion can be shown not to be cognitive by appeal to recalcitrant emotion.

At the very least, this paper shows that the problem of emotional recalcitrance needs to be spelled out much more carefully and with greater sensitivity to the variety of irrationality we see in actual beliefs than it has been so far. Finally, I have drawn an analogy with a clear case of recalcitrant but rational belief—the preface paradox—to suggest that a second formulation of the argument from recalcitrance is also unlikely to succeed. This suggestion is compatible with Cognitivist views that say emotions are beliefs, are reducible to beliefs, or have beliefs as necessary parts.

If my arguments go through, then emotions may well involve beliefs—the most cognitive sorts mental states we have.

Acknowledgements This paper is especially indebted to Joseph Bendaña and Eric Mandelbaum. Conversations with Joseph on an earlier version of this paper greatly improved it and helped develop the paper to bring it into closer conversation with the empirical literature. This paper follows Joseph's own work which discusses that literature in the context of implicit bias and belief. The paper not only benefitted from Eric Mandelbaum's guidance and feedback, but also depends on Eric's prior work drawing attention to the body of empirical studies marshalled here as telling against an idealized notion of belief. I also thank Simon Brown, Nemira Gasiunas-Kopp, Matthew Heeney, Eric Bayruns García, Santiago Echeverri, David Rosenthal, Michelle Montague, Jeff Seidman, and participants in the New York Philosophy of Psychology workshop and the Cognitive Science Symposium at the CUNY Graduate Center for their feedback and conversation. Helpful comments from several anonymous reviewers also significantly improved the paper.

Declarations

Conflict of interest The author declares that there is no conflict of interest.

References

Anderson, C. A. (1983). Abstract and concrete data in the perseverance of social theories: When weak data lead to unshakeable beliefs. *Journal of Experimental Social Psychology, 19*(2), 93–108.

Anderson, C. A., Lepper, M. R., & Ross, L. (1980). Perseverance of social theories: The role of explanation in the persistence of discredited information. *Journal of Personality and Social Psychology, 39*(6), 1037–1049.
Anderson, C. A., & Sechler, E. S. (1986). Effects of explanation and counterexplanation on the development and use of social theories. *Journal of Personality and Social Psychology, 50*(1), 24–34.
Baron, J., & Jost, J. T. (2019). False equivalence: Are liberals and conservatives in the United States equally biased? *Perspectives on Psychological Science*, *14*(2), 292–303. https://doi.org/10.1177/1745691618788876
Barrett, L. F. (2017). The theory of constructed emotion: An active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience, 12*(1), 1–23.
Batson, C. D. (1975). Rational processing or rationalization? The effect of disconfirming information on a stated religious belief. *Journal of Personality and Social Psychology, 32*(1), 176–184.
Benbaji, H. (2013). How is recalcitrant emotion possible? *Australasian Journal of Philosophy, 91*(3), 577–599.
Bendaña, J. (2021). Implicit attitudes are (probably) beliefs. In A. Kinderman & O. C. Borgoni (Eds.), *The fragmented mind.* Berlin: Oxford University Press.
Bendaña, J., & Mandelbaum, E. (2021). The Fragmentation of belief. In A. Kinderman & O. C. Borgoni (Eds.), *The fragmented mind.* Oxford: Oxford University Press.
Brady, M. S. (2007). Recalcitrant emotions and visual illusions. *American Philosophical Quarterly, 44*(3), 273–284.
Brady, M. S. (2009). The irrationality of recalcitrant emotions. *Philosophical Studies, 145*, 413–430.
Broome, J. (2013). *Rationality through reasoning*. Wiley.
Clark, A. (2015). *Surfing uncertainty: Prediction, action, and the embodied mind*. Oxford University Press.
D'Arms, J., & Jacobson, D. (2003). The significance of recalcitrant emotion (or, anti-quasijudgmentalism). *Royal Institute of Philosophy Supplements, 52*, 127–145.
Deonna, J., & Teroni, F. (2012). *The emotions: A philosophical introduction*. Berlin: Routledge.
Deonna, J., & Teroni, F. (2017). Getting bodily feelings into emotional experience in the right way. *Emotion Review, 9*(1), 55–63.
Döring, S. A. (2014). Why recalcitrant emotions are not irrational. In S. Roeser & C. Todd (Eds.), *Emotion and value* (p. 124). Oxford University Press.
Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology*. MIT Press.
Gendler, T. S. (2008). Alief and belief. *Journal of Philosophy, 105*(10), 634–663.
Gordon, R. M. (1994). *The structure of emotions: Investigations in cognitive philosophy (Reprinted)*. Cambridge University Press.
Green, O. H. (1992). *The emotions: A philosophical theory*. Springer.
Greenspan, P. (1980). A case of mixed feelings: Ambivalence and the logic of emotion. In A. O. Rorty (Ed.), explaining emotions (pp. 223–250). University of California Press.
Greenspan, P. S. (1981). Emotions as evaluations. *Pacific Philosophical Quarterly, 62*(2), 158–169.
Grzankowski, A. (2017). The real trouble with recalcitrant emotions. *Erkenntnis, 82*(3), 641–651.
Grzankowski, A. (2020). Navigating recalcitrant emotions. *The Journal of Philosophy, 117*(9), 501–519.
Helm, B. W. (2001). *Emotional reason: Deliberation, motivation, and the nature of value*. Cambridge University Press.
Helm, B. W. (2015). Emotions and recalcitrance: Reevaluating the perceptual model. *Dialectica, 69*(3), 417–433.
Helm, B. W. (forthcoming). Cognitivist theories of emotions. In A. Scarantino (Ed.) *The Routledge Handbook of Emotion Theory*. Routledge.
Huddleston, A. (2012). Naughty beliefs. *Philosophical Studies, 160*(2), 209–222.
Jern, A., Chang, K. M. K., & Kemp, C. (2014). Belief polarization is not always irrational. *Psychological Review, 121*(2), 206–224.
Johnson, H. M., & Seifert, C. M. (1994). Sources of the continued influence effect: When misinformation in memory affects later inferences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(6), 1420–1436.
Kolodny, N. (2007). How does coherence matter? *Proceedings of the Aristotelian Society, 107*(1.3), 229–263.
Kolodny, N. (2008). Why be disposed to be coherent? *Ethics, 118*(3), 437–463.
Kuhn, D., & Lao, J. (1996). Effects of evidence on attitudes: Is polarization the norm? *Psychological Science, 7*(2), 115–120.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin, 108*(3), 480–498.
Levy, N. (2015). Neither fish nor fowl: Implicit attitudes as patchy endorsements. *Noûs, 49*, 800–823.
Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology, 37*(11), 2098–2109.
Makinson, D. C. (1965). The paradox of the preface. *Analysis, 25*(6), 205–207.
Mandelbaum, E. (2013). Against alief. *Philosophical Studies, 165*(1), 197–211.
Mandelbaum, E. (2019). Troubles with Bayesianism: An introduction to the psychological immune system. *Mind & Language, 34*, 141.
Mandelbaum, E., & Quilty-Dunn, J. (2015). Believing without reason: Or: Why liberals shouldn't watch fox news. *The Harvard Review of Philosophy, 22*, 42–52.
McLaughlin, B. P. (1988). Exploring the possibility of self-deception in belief. In B. P. McLaughlin & A. O. Rorty (Eds.), Perspectives on Self-Deception (p. 36). University of California Press.
Naar, H. (2018). Sentiments. In H. Naar & F. Teroni (Eds.), *The ontology of emotions.* Cambridge University Press.
Nussbaum, M. C. (2004). *Upheavals of thought: The intelligence of emotions*. Cambridge University Press.
MC, Nussbaum. (2009). *The stoics on the extirpation of the passions. In the therapy of desire* (pp. 359–401). Princeton University Press.
Nyhan, B., & Reifler, J. (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior, 32*(2), 303–330.
Nyhan, B., Reifler, J., & Ubel, P. A. (2013). The hazards of correcting myths about health care reform. *Medical Care, 51*, 127–132.
Pacherie, E. (2008). The phenomenology of action: A conceptual framework. *Cognition, 107*(1), 179–217.
Prinz, J. J. (2004). *Gut reactions: A perceptual theory of emotion*. Oxford University Press.
Prinz, J. J. (2006). Is the mind really modular? In R. J. Stainton (Ed.), *Contemporary debates in cognitive science* (pp. 22–36). Blackwell.
Pylyshyn, Z. W. (2003). *Seeing and visualizing: It's not what you think*. MIT Press.
Quilty-Dunn, J. (2017). *Syntax and semantics of perceptual representation*. Ph.D. dissertation, The Graduate Center, City University of New York.
Reisenzein, R. (2009). Emotional experience in the computational belief–desire theory of emotion. *Emotion Review, 1*(3), 214–222.
Shea, N. (2015). Distinguishing top-down from bottom-up effects. In D. Stokes, M. Matthen, & S. Biggs (Eds.), *Perception and its modalities* (pp. 73–91). Oxford University Press.
Slusher, M. P., & Anderson, C. A. (1989). Belief perseverance and self-defeating behavior. In R. C. Curtis (Ed.), *Self-defeating behaviors: Experimental research, clinical impressions, and practical implications* (pp. 11–40). Springer.
Solomon, R. C. (1993). *The Passions: Emotions and the Meaning of Life*. Berlin: Hackett Publishing.
Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science, 50*, 755–769.
Tappin, B. M., & Gadsby, S. (2019). Biased belief in the Bayesian brain: A deeper look at the evidence. *Consciousness and Cognition, 68*, 107–114.
Tappin, B. M., Pennycook, G., & Rand, D. G. (2021). Rethinking the link between cognitive sophistication and politically motivated reasoning. *Journal of Experimental Psychology: General*, *150*(6), 1095.
Tappolet, C. (2012). Emotions, perceptions, and emotional illusions. In C. Calabi (Ed.), *Perceptual Illusions* (pp. 205–222). Springer.
Tappolet, C. (2016). *Emotions, Values, and Agency*. Oxford University Press.
Viedge, N. (2018). Defending evidence-resistant beliefs. *Pacific Philosophical Quarterly, 99*(3), 517–537.
Wilkes, A. L., & Leatherbarrow, M. (1988). Editing episodic memory following the identification of error. *The Quarterly Journal of Experimental Psychology a: Human Experimental Psychology, 40A*(2), 361–387.
Williams, D. (2018). Hierarchical Bayesian models of delusion. *Consciousness and Cognition, 61*, 129–147.
Wringe, B. (2015). The contents of perception and the contents of emotion. *Noûs, 49*(2), 275–297.