

THE RATIONALITY OF EMOTIONAL CHANGE: TOWARD A PROCESS VIEW¹

Oded Na'aman
The Hebrew University of Jerusalem

This is the penultimate draft of an article that is forthcoming in *Noûs*.
Please cite published version.

Abstract

The paper argues against a widely held synchronic view of emotional rationality. I begin by considering recent philosophical literature on various backward-looking emotions, such as regret, grief, resentment, and anger. I articulate the general problem these accounts grapple with: a certain diminution in backward-looking emotions seems fitting while the reasons for these emotions seem to persist. The problem, I argue, rests on the assumption that if the facts that give reason for an emotion remain unchanged, the emotion remains fitting. However, I argue there are *rationality self-consuming attitudes*: affective attitudes that become less fitting the longer they endure while the facts that give reason for them persist. A widely held synchronic view of fitting affective attitudes denies that fittingness at a time depends on the agent's attitudes at different times and therefore denies that the fittingness of an affective attitude can depend on its duration. Once we reject the synchronic view, we may see that affective attitudes are often fitting due to the fitting processes of which they are part. These fitting processes explain the fitting diminution of backward-looking emotions as well as other diachronic aspects of the fittingness of emotions.

Like actions and choices, attitudes are subject to normative standards. From belief and desire to envy, admiration, grief, amusement, and lust—we evaluate, criticize, disown, and endorse attitudes. Though such attitudes are not voluntary, they are rational in the sense that they can succeed or fail to meet the normative standards we apply to them.² Furthermore, it is thought that the rationality of these attitudes is related to the fact that they are *about* an object or *directed* toward it. It is said that such attitudes can succeed or fail to be fitting, suitable, or appropriate with respect to their objects. They are therefore known as *fitting attitudes*.³

Recent philosophical discussion of fitting attitudes has revolved around two main issues. The first is the *explanatory issue*: whether, at the most fundamental level, the fittingness of an attitude is explained by the value of its object or, as the so-called fitting attitudes theories of value maintain, the value of an object is explained by the attitude that is a fitting response to it, or, finally, both the fittingness of an attitude *and* the value of its object are explained by one's reasons.⁴ Second, there has been much debate about a related issue, known as *the wrong kind of reason problem*: how to distinguish facts that make an attitude fitting or appropriate to its object from facts that seem to give us reason to entertain the attitude but do not make the attitude fitting.

In this essay, I wish to reinvigorate a third issue, which might as well be called a non-issue because philosophers seem to be in general agreement about it—namely, how do rational attitudes evolve over time? I will argue that the answer to this question that is commonly presupposed by philosophers is mistaken. The answer I propose instead constitutes a significant departure from the current understanding of the rationality of attitudes in general and of emotions in particular.

Consider the following scenario. Walking down a dark alley at night, I notice a shadow of a large animal appearing from around the corner. I stop dead in my tracks, paralyzed with fear. A moment later, I sigh in relief and embarrassment: the lights of a passing car made a tiny mouse cast a long, intimidating shadow. My fear rationally fades as I recognize that there was no danger. Now suppose some other shadow in fact belongs to a bear. Fear seems fitting as a response to a bear in a dark alley. And yet if the bear carries on without noticing me and disappears around the next corner, my fear may rationally diminish. In this version, my fear diminishes not because there was *never* any danger, but because there is *no longer* any danger. Finally, suppose there is a bear and he does not walk away but rather stops and looks right at me. I know that if I keep still the bear is more likely to leave me alone, but my fear makes me tremble. I *should* overcome my fear and keep still, not because my fear is not fitting—on the contrary, it is perfectly fitting given that I am engaged in a staring contest with a bear—but I have other reason not to succumb to my fear. I have reason concerning the consequences of my fear; specifically, concerning the prospects of my survival.⁵

This admittedly crude account of my fear of a bear in a dark alley is meant to illustrate three ways in which an affective attitude may rationally change or diminish. Either I realize that what seemed to call for a certain affective attitude does not in fact call for it; or the attitude is no

longer called for because its object has changed; or the attitude is called for but there are other facts that count against entertaining it. In each of the three cases, whether the attitude rationally persists or changes is independent of facts about its past existence. According to a common picture of rational change in affective attitudes, these are the only three ways in which an affective attitude can rationally change. Of this picture, I make four claims.

First, I show that this is indeed a widespread conception of rational change in affective attitudes, which guides, in particular, recent views about backward-looking emotions, such as grief, regret, resentment, and anger. Second, I argue that this common conception is mistaken: a change in the rational status of an affective attitude might depend on the attitude's duration. More specifically, I offer a counterexample to this conception of rational change by showing that some emotions are *rationally self-consuming*: the longer they endure the less rational they become. Third, I argue that the common conception of rational change in affective attitudes rests on the mistaken assumption that fittingness is synchronic, that is, that what affective attitude it is fitting for one to have at a time does not generally and directly depend on one's affective attitudes at other times. I trace the appeal of the synchronic view of fittingness to a mistaken analogy with true belief. And, fourth, I propose a diachronic view of fittingness according to which what affective attitude it is fitting for one to have at a time can generally and directly depend on what non-instrumental process it is fitting for one to undergo. The process view I propose accounts for rationally self-consuming attitudes as well as other diachronic aspects of the fittingness of emotions.

1. Backward-Looking Emotions

Emotions are commonly understood as a category of attitudes: they are affective attitudes that, unlike mere moods or feelings, are directed toward an object. This is, in any case, how I will understand the term in what follows. In recent years, philosophers have struggled to explain the rational diminution of certain backward-looking emotions, such as grief, regret, anger, and resentment. Each of these examples warrants a full paper, but here I will only mention them briefly to call attention to the common thread that connects them.

Consider the duration of grief following the death of a loved one. While psychiatrists consider prolonged grief pathological, some philosophers have recently been puzzled by the diminution of grief. Given that grief is warranted by the death of a loved one, which does not change with time, why should it be fitting to grieve less over time? How might the diminution of

grief be justified if the reason for it stays the same? Berislav Marušić argues that this puzzle eludes a solution (Marušić 2018), while Dan Moller argues that as long as we continue to love the deceased, it is never fitting to recover from grief (Moller 2017). Though there might be prudential and even moral reasons to overcome one's grief, there remains, according to Moller, a powerful reason to persist in grief. A failure to grieve is a failure to appreciate the loss of the deceased.

Regret raises similar worries. Jeff McMahan (2005), Liz Harman (2009), Jay Wallace (2013), and Kieran Setiya (2014) have all discussed cases in which there seems to be a rational lack of regret about decisions the agents themselves judge unjustified, bad, or wrong. One of the most discussed cases in this context is Derek Parfit's case of the young girl's child (Parfit 1984). A 14-year-old girl decides to have a child though she has decisive moral and prudential reason not to. And yet many think that her subsequent lack of regret is warranted. Given that the young girl's regret would be initially fitting, what might rationally justify the girl's eventual affirmation of her choice to have a child when she did?

Harman and Wallace argue that a shift in the young girl's attachments—specifically, her love for, or relationship with the child—gives her reason not to regret her unjustified decision; Setiya claims, more radically, that we have reason to prefer the existence of anyone who co-exists with us, whether or not we have a special relationship with the person; and McMahan argues that, in non-moral cases, a change in attitude can be justified by a change in the agent's personal values and commitments. It is not always clear with regard to each of these accounts whether the explanation is meant to establish that the lack of regret is fitting or merely that there are strong reasons not to regret despite the fact that regret remains fitting (I will return to this point shortly.) Furthermore, the case in question is an unusual case of regret—the regrettable decision results in a valuable upshot—and the puzzle it raises is premised on the assumption that in the normal case an unjustified or wrong decision renders regret fitting indefinitely. Against this assumption, the aforementioned accounts justify the girl's lack of regret by appealing to features peculiar to her case, namely, the normatively significant upshots of her decision. Thus, the accounts imply that in normal cases of regret, where attachments and personal values do not change and no person comes into existence, rational regret persists.⁶

Finally, Pamela Hieronymi (2001) articulates a challenge for accounts of forgiveness. Forgiveness is generally understood as the rational (if non-voluntary) forgoing of warranted

resentment. Hieronymi asks: Given that resentment is warranted by a past wrong, as well as by the moral accountability of the wrongdoer and moral standing of the victim, and given that these facts do not change, how can the diminution of warranted resentment be *rational*? Hieronymi's solution shifts the object of resentment from the past wrong itself to the threat that emerges from it. Resentment, according to Hieronymi, is a protest against a present threat created by a past wrong. Genuine forgiveness is warranted when the threat is removed—through apology, for instance—and there is nothing more to protest. Agnes Callard, by contrast, maintains that when one suffers a wrong one has “reason to be angry forever,” but may rationally cease to care about the wrong done to one once it is acknowledged as such by the wrongdoer and the moral relationship is restored (Callard 2017).⁷

These different authors seem to grapple with distinct instances of the same general problem. I want to better understand the problem and the assumption that gives rise to it. This task is complicated by the fact that the different accounts employ different normative notions or interpret the same notions in slightly different ways. Therefore, I will now introduce a unified terminology that I will then use to recast the claims reviewed above, articulate their common assumption, and describe the general problem the assumption gives rise to.

In an important paper, Justin D'Arms and Daniel Jacobson distinguish between, on the one hand, the fittingness or correctness of an emotion to its object and, on the other, its overall rational justification—whether it is what one ought to feel, all things considered (D'Arms and Jacobson 2000a, 72). While an emotion can be fitting in the sense that it properly presents its object as having certain evaluative properties, the emotion might nevertheless be unjustified overall. For example, a cruel joke might be amusing and therefore amusement would be a correct or fitting response, but there might be moral reason not to be amused and this reason can render amusement morally inappropriate or unjustified overall. It is clear why I am laughing at the joke, or what is amusing about it, but it is wrong to be amused. By contrast, the story of my sad childhood might not be amusing quite apart from any moral consideration. In this case, amusement is not (or not only) unjustified but unfitting: it does not make sense as a response to the story of my childhood. The same point is sometimes put in terms of the right and wrong kind of reason for an emotion of a certain type.⁸ I have a ‘right kind of reason’ not to be amused by a joke when the joke is not funny and a ‘wrong kind of reason’ not to be amused when the joke is offensive, though many believe

that both right and wrong kinds of reasons can be genuinely normative reasons against being amused and, as such, good reasons.⁹

Now, an account of the rational change or diminution of an emotion can appeal to *fitting reasons*, which may impact not only the overall rational justification of the emotion but its fittingness, or it might appeal to *extraneous reasons*, which do not impact the fittingness of the emotion but may impact its overall rational status, whether it is what one ought to feel.¹⁰ We might have extraneous reasons to overcome fitting grief, Dan Moller concedes, but he insists that we continue to have fitting reason to grieve as long as we love the deceased. Hieronymi, by contrast, wants to explain how resentment that was once fitting may cease to be so given a relevant change in its object. The basic distinction is between explanations of the rational change in attitude as a change in the fittingness conditions of the attitude and explanations of the rational change in attitude as a change in considerations external to the fittingness of the attitude. Only an extraneous reason explanation insists that the change in attitude, though justified overall, remains in some important sense rationally flawed because unfitting. Fittingness explanations, by contrast, attempt to justify the change in attitude by locating a shift in the facts that made the attitude initially fitting.

We can now use these distinctions to articulate the presumption common to different accounts of backward-looking emotions and the general problem this presumption gives rise to. The presumption, which I will call *the stability of fitting reasons* (SFR, for short), maintains that as long as the facts that constitute fitting reasons for an emotion persist, they *continue* to constitute fitting reasons for the emotion.¹¹ SFR might seem, at first, trivial. The past loss that gives fitting reason for grief cannot be undone and it continues to be true that it had occurred. But SFR makes the less obvious claim that the persistent loss continues *to give fitting reason for grief*.¹² This is a claim about the persistent normative status of the facts, not about the persistence of the facts themselves. If SFR is true, then a diminution in the relevant emotion would not be fitting as long as the facts that initially give us fitting reason for it persist, and in the case of backward-looking emotions these facts seem to persist.

While some philosophers concede that backward-looking emotions remain fitting and offer an extraneous reason explanation of their rational change, others insist on uncovering a change in the facts that give the fitting reason for the emotion and thereby providing a fittingness account that would explain why the backward-looking emotion not only rationally but fittingly diminishes.

Moller presupposes SFR when he maintains that our reason to grieve persists long after grief dissipates and Marušić assumes it when he claims that while grief fittingly diminishes there could be no rational explanation of its fitting diminution. SFR is also behind Hieronymi's attempt to locate a change in the facts that constitute reason for resentment in order to explain resentment's fitting diminution as well as behind Callard's idea that a reason to be angry is a reason to be angry forever. Thus, philosophers tend to assume that where the facts that give fitting reason for an emotion are stable, the fittingness of the emotion they supports is stable, too.

Notice, moreover, that the problem of backward-looking emotions as described here implies that the fittingness of an emotion at a time is determined independently of facts about its duration. More specifically, given SFR *and* a common understanding of the facts that constitute fitting reasons for backward-looking emotions, the duration of an emotion is irrelevant to its fittingness. To see this, suppose that at t_1 a set of facts, r , gives (sufficient) fitting reason to ϕ and that ϕ -ing does not occur prior to t_1 . According to SFR, the fitting duration of ϕ -ing is determined by the duration of r . So whether ϕ -ing continues to be fitting at t_2 depends primarily on whether r continues to hold at t_2 , and in cases of backward-looking emotions it normally does, or so the thought goes. However, one difference between ϕ -ing at t_1 and ϕ -ing at t_2 is that at t_1 the agent *begins to ϕ* while at t_2 the agent *continues to ϕ* . In other words, the duration of ϕ -ing changes; the emotion has a different history in each moment. But we are assuming that at t_2 r persists and continues to give (sufficient) fitting reason to ϕ , just as it did at t_1 . It follows that the fittingness of ϕ -ing at a time is independent of its duration. If the fittingness of an emotion (partly) depended on its duration, then the fitting reason for the emotion would change with the duration of the emotion. So the idea that fitting reasons persist implies that the fittingness of ϕ -ing at any time, t_n , does not directly depend on the history of this instance of ϕ -ing, i.e., on whether the agent has already been ϕ -ing at t_{n-m} .¹³ Put simply, the duration of the emotion does not directly bear on its fittingness. In the next section I claim that this is often not the case.

2. Rationally Self-Consuming Attitudes

It is a striking fact about the accounts we reviewed that none of them even considers the possibility that whether an attitude is (or continues to be) fitting might depend on its duration. This lacuna can be traced back to the paper in which D'Arms and Jacobson originally draw the distinction

between considerations of fittingness and extraneous considerations. They claim that considerations of fittingness can be divided into considerations of *shape* and considerations of *size* (2000a, 73–75). An emotional episode is unfitting in shape when it presents its object as having certain evaluative features the object in fact lacks; it is unfitting in size when it is disproportional to the evaluative features of its object. Regret, for instance, is unfitting in shape if its object is not regrettable (no reason to cry over spilt milk if the milk is still safely in the bottle,) and it is unfitting in size if it is disproportional to the significance of its object (the milk is spilt, and that is regrettable, but there is no reason to cry because *it is only milk*).¹⁴

But there is a third kind of considerations of fittingness that goes unmentioned by D’Arms and Jacobson, namely, considerations of *length*. What is the duration of a fitting emotion? Perhaps regretting the spilt milk is fitting for a moment, but surely it is not fitting to regret it all day long. A moderately funny joke warrants amusement, but it may be unfitting to be amused by it all week. Similarly, we may plausibly raise substantive questions about the fitting duration of an emotion: is it fitting to grieve longer those who were closest to us? Does it matter for fittingness whether my resentment is brief and intense rather than prolonged and moderate?

Considerations of length lend plausibility to the idea that some attitudes, and particularly some emotions, are *rationaly self-consuming*. Whether an emotion fittingly persists might not only depend on the persistence of the facts that constitute the fitting reason for it, but on the fitting evolution of the emotion over time. The fact that I have been amused by a funny joke all day long can make the dissipation of amusement fitting; the fact that I have resented an offense for many years can make a decrease in my resentment fitting; the fact that I have regretted my mistake can make a diminution in regret fitting; and, similarly, even when intense fear is appropriate as a first reaction to danger, fear may fittingly decrease to make room for other attitudes, such as a forward-looking intention to neutralize the threat. Thus, the fact that an emotion has persisted for some time might itself render its dissipation fitting and its continuation unfitting. Like fire, which can be the cause of its own expiration, it is part of the rational structure of certain attitudes that they consume themselves: the longer they endure the less fitting they become. For it is the fact that you have had the fitting response to some event for some time that makes that very response less fitting.

It might initially seem as if I am suggesting, implausibly, that the length of an emotion itself constitutes a fitting reason for its diminution. I am not. The reason for one’s resentment, whether

intense or mild, is the injustice done, not facts about the duration of one's resentment; similarly, a year after the death of one's beloved, the reason for one's lingering grief remains the death of the beloved and does not include facts about the process of grieving one has undergone (see Marušić 2018, 16). But these observations are compatible with the claim that the fittingness of an emotion can depend on its duration. Let me explain.

Fittingness explanations of rational change in attitude can focus on changes in the facts that constitute fitting reasons for the attitude or on changes in the background conditions in virtue of which these facts constitute fitting reasons for the attitude. The idea is that various facts can be part of the rational explanation of the reason to ϕ without being the reason, or part of the reason to ϕ (see Dancy 2004, 39–40; Schreoder 2007, 27; Scanlon 2014, 48). The fact that I love Helen does not itself give me a reason to grieve for her, but it is a background condition that explains why her death gives me a (fitting) reason to grieve for her. Similarly, the fact that you and I stand in a moral relationship does not give you a reason to resent me, but it is arguably a background condition that explains why the fact that I harm you without justification gives you a reason to resent me.

A change in the fittingness of an attitude might be due to a change in the fact that gives the fitting reason for the attitude or due to a change in the background conditions. If the fitting diminution of resentment is due to the neutralization of the threat that is its object, then in this case fittingness changes due to a change in the fact that constitutes the reason. Alternatively, if the fitting diminution of grief is due to the waning of one's love for the deceased, then in this case fittingness changes due to a change in the background conditions. When I use the term “fittingness conditions” or “facts that determine fittingness” I lump together the background conditions and the fact that constitutes the fitting reason, both of which play a role in determining the fittingness status of the attitude. The term “rational conditions,” on the other hand, refers to all facts that contribute to the rational status of the attitude, including extraneous facts, such as facts about the moral and prudential implications of entertaining the attitude (which might not touch on the fittingness status of the attitude.)¹⁵

I propose that in cases of rationally self-consuming emotions, the duration of the emotion is part of the background conditions that explain why the fact that made the emotion fitting in the past no longer makes it fitting, or no longer makes fitting the same intensity of emotion. That I resented the injustice for a long time might be part of the rational explanation of the fitting

diminution of my resentment without being the *reason* (or the fact that constitutes the reason) for my current resentment. The fact that constitutes the reason for whatever resentment I still harbor remains the injustice done to me; it constitutes a reason for milder resentment because I have already resented the injustice intensely in the past.

Compare explanations of rational partiality. A person might fittingly admire a great artist for her achievements, but it would not be fitting for the person to be proud of the artist unless they share some special relationship. If, per chance, the person in question becomes the artist's spouse, pride would become fitting.¹⁶ Similarly, we cite the history of a rationally self-consuming attitude to explain why the fact that once made it fitting no longer does so. In both cases a relevant change in background conditions accounts for a change in fittingness while the facts that give the fitting reason remain unchanged.

The phenomenon of rationally self-consuming attitudes is compatible with explanations of rational change in attitude that appeal to various other fittingness conditions. For example, even if resentment fittingly diminishes only once the wrongdoer asks for forgiveness, the fact that resentment is rationally self-consuming explains why it is unfitting to seek forgiveness too soon after the offense took place and why a reluctance to forgive after considerable time has passed can, at times, seem unfittingly resentful. In both cases, resentment does not run its rational, fitting course.¹⁷ Thus, the phenomenon of rationally self-consuming attitudes does not conflict with fittingness accounts that appeal to changes in facts independent of the attitude; it complements them. Moreover, positive and forward-looking emotions, too, may be rationally self-consuming. The joy we feel upon being offered the job we wanted may fittingly diminish with time while the relevant facts about the job and our judgment and expectations of it remain unchanged.¹⁸

Someone might doubt the idea of rationally self-consuming attitudes for the following reason. It might seem that what explains the diminution of an emotion such as grief, for example, is not its past existence but the time that has passed since the loss. But consider a case where one receives a letter bearing the sad news of a friend's death long after the death had occurred. In such a case, I submit, grief might fittingly diminish later in time, compared to an episode of grief that begins immediately after the loss. This shows that what matters for the fitting diminution of grief is not primarily the time that has passed since the death, but whether grief had occurred during that time and whether, once it had occurred, it evolved fittingly.¹⁹

The phenomenon of rationally self-consuming attitudes does not immediately solve the problem of backward-looking emotions, but it points us in a promising new direction. It shows that whether an attitude of an agent at a given time is fitting may depend on what attitudes the agent had at other times, and therefore that SFR should be rejected: the fittingness of an affective attitude may change even if the facts that constituted the initial fitting reason for the attitude stay the same. However, as we saw in the previous section, the conception of rational change in emotion presupposed by many philosophers precludes the possibility of such a change in fittingness. Rationally self-consuming attitudes seem to provide a powerful counterexample to the common conception of rational change because the phenomenon is pre-theoretically intuitive. But if that is true, why have philosophers overlooked this phenomenon?

3. The Synchronic View

I believe philosophers have overlooked the phenomenon of rationally self-consuming attitudes because they are attracted to the idea that fittingness is synchronically determined. *The synchronic view of fittingness* (henceforth, the synchronic view) holds that what affective attitude it is fitting for one to have at a time does not *generally* and *directly* depend on what affective attitudes one has at other times. Before considering the appeal of the synchronic view, I will explain the view and why it rules out the idea that duration bears on fittingness in the way suggested by rationally self-consuming attitudes.

The synchronic view allows for indirect dependence on attitudes at other times—i.e., it allows that facts that determine fittingness at a time themselves depend on the agent’s attitudes at other times. For instance, the fittingness of my fear of the bear in the dark alley *indirectly* depends on my intention to walk down the alley to begin with. However, the synchronic view rules out the possibility that facts about an attitude’s fittingness at a time depend on facts about the agent’s attitudes at other times *when all other fittingness-making facts are equal*. Whether I feared the bear a moment ago is not relevant to whether fear is fitting right now given that all other fittingness-making facts—such as the danger the bear poses to me—remain unchanged. Plausible as the view might seem when applied to fear, it purports to hold for all affective attitudes; it is this generalization that I wish to contest.

Furthermore, the synchronic view denies only *general* direct dependence on attitudes at other times. Specifically, it allows for direct dependence when the object of the attitude is itself an

attitude experienced at another time. I may now feel shame about the fact that in the past I envied my friend's success.²⁰ The fittingness of my present shame depends directly on my past envy. But the dependence is explained in this case by the fact that my past envy is *the object* of my shame; not by the fittingness conditions of shame in general. On the synchronic view, the fittingness of shame, like the fittingness of all other affective attitudes, does not generally and directly depend on the agent's attitudes at other times. If the fittingness of an emotion at a time directly depends on its duration, as the phenomenon of rationally self-consuming attitudes suggests, then whether the emotion has begun in the past is *generally* relevant to the fittingness of the emotion in the present, irrespective of the specific object of the emotion. This possibility is precluded by the synchronic view of fittingness.

While the synchronic view is explicitly endorsed and argued for by some authors who write about the rationality of belief, it is rarely made explicit in the literature on the rationality of affective attitudes.²¹ What might explain the implicit popularity of the synchronic view is that, given certain plausible assumptions, it follows from a more narrow view about fittingness that seems quite natural, namely, *the object view of fittingness* (henceforth, the object view). The object view holds that what affective attitude it is fitting for one to have at a time directly depends only on what the attitude is about, namely, on the properties the attitude attributes to its object.²² The view is captured by the thought that "the only ... considerations [relevant to the fittingness of an emotion] are those reasons that speak to whether an emotion correctly represents its object" because emotions "purport to be presentations of such properties as the funny, the shameful, the fearsome, the pitiable, etc." (D'Arms and Jacobson 2000a, 66). So on the object view, F is fitting in response to X iff X is E, where E is the evaluative property that F attributes to X. It follows that the fittingness of an emotion is independent of facts that are not part of the representational content of the emotion.

The view that the fittingness of an affective attitude depends only on the evaluative properties it attributes to its object does not yet imply the synchronic view and rule out rationally self-consuming attitudes. Whether the object of an affective attitude has the relevant evaluative properties can, in principle, depend on the agent's attitudes at other times, in which case the object view would imply that the fittingness of the affective attitude depends on the agent's attitudes at other times. So, for example, whether an occurrence is regrettable might partly depend on whether one has already regretted it. It would follow that the fittingness of regret about X at *t* directly

depends on the agent's past regret *because* whether X is regrettable at *t* depends on the agent's past regret.²³

This is indeed a conceptually possible version of the object view that is compatible with the existence of rationally self-consuming attitudes. However, this version of the view is, implicitly or explicitly, rejected by philosophers writing on the fittingness of affective attitudes and there are good reasons to reject it. Two examples should be sufficient to establish this. First, D'Arms and Jacobson argue that 'sentimental values'—those evaluative features properly analyzed, according to them, in terms of fitting sentiments—are “relatively stable features of the objects to which they are properly attributed” (D'Arms and Jacobson 2009, 588). However, they also claim that “our dispositions to amusement, shame, disgust, and other sentiments are prone to vicissitudes” (ibid). They attempt to defend the thesis that despite this asymmetry in stability, sentimental values depend on the associated sentiments. I will return to their argument shortly to contest the claim that due to their instability the sentiments fail to present the evaluative features we take them to track. But I agree with D'Arms and Jacobson that evaluative properties ('sentimental values') tend to remain stable through fluctuations in affective attitudes. Generally, the mere fact that a person has been amused, ashamed, or disgusted by something does not imply that the object has ceased to be amusing, shameful, or disgusting, even where one's attitude fittingly diminishes.

Consider another important view of fittingness. Conor McHugh and Jonathan Way argue that what it is for X to be good is for it to be fitting to value X (2016, 583). They maintain that a 'valuing attitude' has internal standards that specify the features that make an object fitting for it. Humility and concern for others, for instance, might be features that make someone admirable (584). Now, if the goodness of X can directly depend on whether the agent has fittingly valued X in the past, then X might cease to be good once it has been fittingly valued. But this is implausible. The same is true of the bad: that I resented you for the wrong doesn't make the wrong go away. If McHugh and Way want to avoid these results, as I believe they should, then they should also reject the idea that, as a general matter, a fitting attitude is about its own past existence.

It is therefore implausible that, as a general matter, the evaluative properties represented by an emotion depend on the emotion itself or on whether the agent has already experienced it.²⁴ When combined with the object view, this claim implies that the fittingness of an affective attitude at a time does not directly depend on one's affective attitudes at other times, which is what the

synchronic view says. However, unlike the synchronic view, the object view also rules out the possibility that fittingness at a time directly depends on *anything* other than what the attitude is about. The view therefore denies that the fittingness of an attitude directly depends on background conditions that explain the reason for the attitude.²⁵ Thus, the object view holds that at any moment t_n , the fittingness of an affective attitude F in response to X is fully determined directly by whether the properties that F attributes to X in fact obtain at t_n . It also follows from the object view that the fitting *duration* of an affective attitude F in response to X is fully determined by the persistence of the properties F attributes to X over time. For example, an object view of fear might hold that X is fittingly feared only when and as long as X is dangerous for the agent. The fittingness of the attitude is determined anew at each moment by the object it represents and the distinction between an agent's first response to the object, the agent's later response, and the agent's enduring response falls out of view.²⁶

I believe that it is because D'Arms and Jacobson presuppose the object view that they do not mention fittingness considerations of *length*. If they assume that the fittingness of an emotion at a time always and only depends on the properties it attributes to the object at that time, then they might be led to assume that the fitting length of an attitude—the fitting continuation of the attitude over time—generally and directly depends on fittingness considerations of shape and size at each moment.

The object view may seem unavoidable if we model our understanding of fitting affective attitudes on our understanding of true beliefs. Consider D'Arms and Jacobson's following explanation of fitting affective attitudes:

Emotions present things to us as having certain evaluative features. When we ask whether an emotion is fitting, in the sense relevant to whether its object is ϕ , we are asking about the correctness of these presentations. The relevant considerations, then, are just those that count as evidence for the evaluations an emotion presents to us. In this respect, the fittingness of an emotion is like the truth of a belief. (D'Arms and Jacobson 2000a, 72)²⁷

The correctness of belief is determined by its truth while the subjective justification of belief depends on the evidence and perhaps on certain practical considerations.²⁸ Whether a belief is true (and therefore correct) at a given moment, generally and directly depends only on its object—a proposition, or on what the proposition represents—and not on things other than its object, such as the agent's relation to it or the agent's attitudes at other times. Therefore, if it is now true that

the sun is out, then my belief that the sun is out is true and remains true only when and as long as the sun is out. If the model of true belief is carried over to fitting affective attitudes, then fittingness, too, would seem to directly depend only on the properties attributed to the object.

The model of true belief leads D'Arms and Jacobson to worry about cases where affective attitudes are not responsive to the relevant evaluative properties of their objects (D'Arms and Jacobson 2009). We would probably not be amused by a funny joke we heard many times before even if we agree the joke remains funny; our amusement might also be preempted by our despondent mood, by our moral judgment, by social circumstance, or by our identification with or alienation from the person who is the target of the joke. According to D'Arms and Jacobson, such obscuring factors explain what they call *the instability of affect*. As a result of this instability, they claim, it is often the case that our affective attitudes do not correctly present the value of their objects even by our own lights. On their view, amusement by X is always the *correct* or *fitting* response to the fact that X is funny just as a belief that p is always the correct or true response to the fact that p is true.

But while a belief that p retains its representational content independently of the agent's relation to p, amusement by X does not retain its representational content independently of the agent's relation to X. What amusement "says" about its object often depends on context. With regard to some jokes at least, it is fitting to be amused by them the first time one hears them but the same amusement would not be fitting upon the joke's eleventh hearing (just as it might not be fitting to be amused by a moderately funny joke for too long.) It is not simply that in these cases amusement remains fitting while extraneous reasons make it overall unjustified nor is it the case that various social and psychological phenomena interfere with fitting amusement. Rather, repeated amusement might cease to represent what is funny about the joke and therefore would not be fitting. Amusement in response to the funny is fitting—i.e., it correctly represents its object as funny—only given certain relational conditions.

D'Arms and Jacobson acknowledge in passing that fittingness might depend on the agent's relation to the object of her emotion. Fear correctly presents the tiger as fearsome, they say, but only if the tiger is near the agent and on the loose, not when the tiger is an example in a philosophy paper (2000b, 729). And I have been arguing that the duration of an attitude can be similarly relevant to whether the attitude correctly presents its object's evaluative properties. To resent a wrong for too long might reveal a failure to properly appreciate the significance of the wrong; to

grieve a friend's death and recover too quickly might cast doubt on whether one appreciated the tragedy at all.

Similarly, the history of the agent's attitudes might matter generally and directly for fittingness. Crudely put, we can say that relief is not fitting without a history of frustration or anxiety; desperation is not fitting without a history of hope; satisfaction is not fitting without a history of desire or longing. In short, it is a mistake to assume that we can have a good grasp of the evaluative content of an attitude independently of its relation to the agent's mind over time. The fittingness of an emotion at a time may be partly explained by its broader diachronic context, which includes a properly evolving response to one and the same object with one and the same relevant set of evaluative properties.

I therefore suggest that what D'Arms and Jacobson call "obscuring factors" are in some cases fittingness considerations and what D'Arms and Jacobson consider a fitting attitude in response to evaluative property E is only occasionally fitting. Accordingly, *the instability of affect*, which, on D'Arms and Jacobson's object view seems to conflict with facts about fittingness, can be seen, on a view of fittingness that takes into account the agent's attitudes at other times, as *the responsiveness of affect* to changing fittingness conditions.

To be sure, it is possible to reject the object view and still uphold the synchronic view. Even if the fittingness of an attitude can depend on facts that are not represented by it, and therefore not part of the reason for it, perhaps none of those facts involve the agent's attitudes at other times. A synchronic view of this kind is already a significant step toward resolving the problem of backward-looking emotions. In particular, the rejection of the object view implies the rejection of SFR, i.e., the assumption that as long as the facts that constitute fitting reasons for an emotion persist, they continue to constitute fitting reasons for the emotion. With the idea of background conditions in hand, we can acknowledge that even while the facts that constituted reason for an emotion persist, their status as reasons might change due to changes in background conditions.

As an example, consider the following two interpretations of Jay Wallace's account of the case of the young girl's child. Recall the crucial details of the case: though the young girl's decision to have a child at age fourteen was prudentially and morally unjustified (by hypothesis), we think she may rationally come to celebrate the decision or, at the very least, cease to regret it. Wallace claims that "to understand the evolution of the young girl's attitude without supposing either that

she changes her mind about the relevant normative questions [i.e., without supposing she now believes her decision was justified] or that she acquires responses that are inconsistent toward one and the same state of affairs” we must recognize the relevant change in the young girl’s “deliberative situation,” namely her new, loving relationship to her child (Wallace 2013, 90). This relationship gives her new reasons she did not have before. Given her new reasons, “the young girl will naturally affirm and celebrate the existence of her child, cherishing her daughter and her daughter’s role in her own life” (ibid).

On one possible reading, Wallace offers an extraneous reason account of the change in the young girl’s attitude. Though regret remains fitting, new reasons give her strong, wrong-kind reason not to regret and to affirm her past decision. So the young girl’s lack of regret is akin to one’s cold reception of a funny but morally objectionable joke: in both cases the attitude is not fitting but rationally justified overall. But the rejection of the object view makes possible another reading of Wallace’s view—one that strikes me as more plausible. On this new reading, the fittingness of the young girl’s regret changes due to a change in relevant background conditions. The new relationship with her child constitutes a significant change in fittingness-making facts, without changing the fact that gave the girl reason for regret, namely, that it was wrong or unjustified to have a child at fourteen. On this reading of Wallace, fittingness can depend on facts about things other than the object of the attitude, so the object view is false. However, the account is in line with the synchronic view because it does not appeal to the young girl’s attitudes at other times to explain the fittingness of her present affirmation of her past decision. It might now seem that we achieved what we sought: a fittingness account of rational change in a backward-looking emotion that affirms the pre-theoretical impression that the emotion is about a past occurrence, which cannot be undone.²⁹ If the mere rejection of the object view of fittingness solves the problem of backward-looking emotions, why should we also reject the synchronic view?

To begin with, notice that if the main motivation for the synchronic view is the fact that it follows from the object view, then the rejection of the object view should remove at least this obstacle: why insist on the synchronic view? Once we level the playing field, the phenomenon of rationally self-consuming attitudes seems to tilt the scale against the synchronic view. As I argued in the previous section, reflection on various cases gives us strong phenomenological evidence that the duration of an emotion is part of the background conditions for its fitting reason and therefore part of its fittingness conditions. Since we should reject the object view, we should not be surprised

that the duration of an emotion matters for its fittingness without being part of what the emotion is about.

But self-consuming attitudes are not the only reason to reject the synchronic view. Once the possibility of diachronic fittingness is articulated, it seems to explain other phenomena, too. I mentioned above that the fittingness of an attitude might (generally and directly) depend on past attitudes that are qualitatively different from it. Frustration is a predecessor of fitting relief, hope a predecessor of fitting despair, desire a predecessor of fitting satisfaction, etc. I also mentioned the idea that love is part of the background conditions of fitting grief. Note that the connection between love and grief has a temporal aspect. It is not simply that love normally accompanies fitting grief, but that it normally comes *before* it; love begins prior to when the death of the beloved gives us fitting reason for grief. It is therefore plausible that fitting grief depends generally and directly on a *history* of love. Similarly, but less dramatically, the fittingness of boredom also seems to depend on the agent's past attitudes and engagement with what bores her. In many cases, both the duration of an emotion and the agent's emotional and attitudinal history seem to play a crucial role in determining fittingness.

Moreover, even when understood according to the new reading I suggested, Wallace's account of the young girl's diminution in regret remains incomplete without a diachronic element. The fact that regret fittingly diminishes in normal cases suggests that the case of the young girl's child conflates two distinct questions. One is a question about the rational or fitting diminution of regret and another is a question about the rational impact of the child's existence on the girl's retrospective attitudes. The child's existence is part of why her retrospective attitudes rationally change, but it is not the full explanation—and certainly not the only possible explanation—of her diminution in regret.

Thus, while it is plausible that the young girl's new relationship with her child changes the rational conditions of regret, Wallace's claim that this relationship is the *only* relevant change in the young girl's "deliberative situation" is false. Another change in the girl's deliberative situation that can explain the rational diminution of her regret is the fact that she had regretted the decision in the past. Indeed, if the young girl's eventual lack of regret is not only rational overall but fitting—i.e., if it correctly presents its object—then the fact of parenthood is not sufficient to explain it. Had the girl never reflected on the flaws of her decision, her present affirmation of it would not be fitting

(though it could be justified overall by her reasons of love for her child.) Therefore, it seems to matter for the fittingness of the young girl's present affective attitude whether she had ever regretted the choice in the past and, more generally, how she has responded to her past decision until now. If this is true, then we have another indication that the synchronic view is false: what attitude is fitting at a time can generally and directly depend on one's attitudes at other times.³⁰

4. A Diachronic View

This brings me to my final claim. I have argued that the synchronic view is false; now I wish to propose a specific diachronic view according to which the fittingness of an attitude sometimes directly depends on the non-instrumental process of which it is a part. Start with *the general diachronic view of fittingness* (henceforth, the general diachronic view), according to which what attitude it is fitting for one to have at a time can generally and directly depend on what attitudes one has at other times. The general diachronic view follows from the rejection of the synchronic view; however, it allows that fittingness is sometimes synchronic, that is, that the fittingness of an attitude at a time *sometimes* does not directly depend on one's attitudes at other times. The view also allows that in those cases where the fittingness of an attitude at a time directly depends on one's attitudes at other times it *also* depends on other facts, say on facts about the object of the attitude or about one's relation to the object. As such, the general diachronic view is inclusive: it simply negates the exclusive thesis of the synchronic view, that fittingness does not generally and directly depend on attitudes at other times.

In response to the general diachronic view, one might wonder: why should the mere passing of time bear on the fittingness of an emotion? But note that the general diachronic view does not imply that fittingness is sensitive to bare temporal facts. The view appeals to the passing of time because it is in that dimension that various fittingness-making facts obtain, not because temporal facts themselves are essential to fittingness. What is essential to fittingness, on the general diachronic view, are facts about the ordering of attitudes in time. This clarification is meant to appease worries about the normative arbitrariness of temporal intervals as such. Nevertheless, it may still seem that describing diachronic phenomena in terms of rational requirements does not do enough to explain why they have the rational status I attribute to them. For instance, describing self-consuming attitudes *as* rational does not explain *why* they are rational, i.e., why the sheer

duration of an emotion bears on its fittingness. What is the *rational* explanation of the trajectory of self-consuming attitudes?

To answer this question we need a more specific and informative diachronic account. As a first step, it might help to notice that rationally self-consuming attitudes are only one primary example of rational change that is assessed by a conception of an attitude's fitting evolution. There are, in addition, *rationally self-augmenting attitudes* as well as *rational emotional transitions*. I will explain each in turn.

Rationally self-augmenting attitudes are attitudes that rationally intensify and deepen the longer they endure. It is arguable, for example, that love is such an attitude, as well as enjoyment from engagement in a valued activity and attachment to a place of residence. It is unfitting to love a person I met last week in the same way one might love a person one has known for many years. No matter the special chemistry we have or how well we understand each other, the fact that we are recent acquaintances itself seems to imply that certain emotions toward one another would not be fitting. Of course, love fittingly intensifies partly due to the accumulation of information about the beloved, but a large amount of information acquired without a history of love (such as might be acquired by stalkers, private detectives, Google, or the NSA) does not seem to render the amplification of love fitting.³¹ A history of love is essential to the fittingness of what is tellingly called, "mature love."

Similarly—although this is more speculative—it seems to me that one's growing enjoyment of a valued activity such as lap-swimming, or philosophizing, or the appreciation of art, is plausibly self-augmenting because its fittingness generally and directly depends on one's history of engagement in the activity. It is true that our past engagement in an activity often leads to improvement in our skills, tastes, and sensibilities, and these improvements are normally part of the explanation of why our growing enjoyment is fitting. But I believe that past engagement can make current enjoyment fitting even where there is little gain in skill, taste, or sensibility, and that even where such gains are accomplished the sheer history of engagement still matters to the fittingness of current enjoyment.

The same is true of the kind of attachment we have toward the place we call "home". To fittingly feel toward a place what one often feels toward one's home it is not enough to know many things about it and wish very much to return to it. I can know everything there is to know about

Tokyo, and specifically about the wonderful neighborhood NaKameguro, and wish with all my heart to visit it again, and yet longing to go there as one longs to return home would not be fitting for me. Without venturing to say what exactly makes a place into a home, I believe a certain history is required, perhaps a history of residency—of falling asleep, waking up, having friends, working, wondering around, day after day, in the same house, or neighborhood, or town. The crucial point is that the fittingness of our feelings of attachment to our home seem to generally and directly depend on our history.

Rational emotional transitions occur when one attitude is fittingly replaced by another. At first we are excited but then we might *grow* weary or bored; we view someone as no more than a friend but then *fall* in love; we are distressed by a new predicament but then *reconcile ourselves* to it; we can laugh or cry and then *calm down*. The fittingness of these transitions is not determined merely by the fittingness of the former attitude at the time of its occurrence and the fittingness of the latter attitude at the time of its occurrence. Rather, the issue partly concerns *how* one transitions from one attitude to the next. Furthermore, in these examples the latter attitude is understood and described as emerging from the attitude that preceded it; it cannot be assessed in abstraction from its history. For these reasons, the relation between the agent's attitudes at different times is essential to the fittingness of these emotional transitions.

In all the aforementioned cases—those of rationally self-consuming attitudes, rationally self-augmenting attitudes, and rational emotional transitions—we determine the fittingness of an emotion at each temporal unit by considering the fittingness of the temporal whole of which it is a part. Indeed these are examples of cases where the fittingness of an attitude at a time generally and directly depends on *the fitting process of which the attitude is a part*. Therefore, a plausible diachronic view appeals to the fittingness of processes.

Processes are often contrasted with continuants. While continuants *endure*, processes are said to *perdure*. The traditional way of drawing the distinction is by reference to whether or not the whole has temporal parts. Continuants do not have temporal parts; they exist in their entirety at each moment; they endure. Processes have temporal parts; they exist in their entirety only over time; they perdure. A different way of drawing the distinction, argued for by Hofweber and Velleman, refers to whether the identity of the whole is fully determined at every moment of its existence (Hofweber and Velleman 2010). Thus, a mental state, such as a conscious experience of

a red cube, endures because its identity is determined at every moment at which it exists. In contrast, consider the process of writing a check: “What there is of this process at a particular moment—the laying down of a particular drop [of ink]—is not sufficient to determine that a check is being written, and so it is not sufficient to determine which particular process is taking place...” (14). On either view of the distinction, we can see that self-consuming attitudes, self-augmenting attitudes, and emotional transitions, are fitting as part of diachronically fitting wholes—i.e., fitting processes.

Consider the case of anger. Agnes Callard argues that anger in response to a wrong is properly understood in the context of an impaired moral relationship. Though the wrong will remain unchanged, the moral relationship might be amended and therefore anger may rationally recede (Callard 2017). Put in my terms, though the fitting reason for anger persists, the background conditions change once the relationship is repaired and therefore anger fittingly diminishes. This strikes me as plausible (see Na’aman Forthcoming). It means, however, that the fact that anger is rationally self-consuming is only part of the explanation of its fitting diminution. Anger is properly understood as part of a fitting process of moral repair, a process that has various stages, and includes not only the attitudes, thoughts, and actions of the person who was wronged but also those of the wrongdoer, as well as interactions between the parties, such as offering and accepting an apology or asking for forgiveness and granting it. Given that it is partly constituted by sequences of attitudes and emotions, the process of moral repair also includes self-augmenting attitudes and emotional transitions. Thus, by considering the significance of processes, we arrive at a diachronic view of fittingness that goes beyond the evolution of a single emotion, and beyond individual emotional transitions, to include complex processes that involve various attitudes, feelings, thoughts, actions, and interactions between agents.³²

To be sure, the synchronic view can account for the dependence of fitting attitudes on a kind of process. On Pamela Hieronymi’s view, for example, resentment is a protest against a present threat that originated in a past wrong (Hieronymi 2001). If the relevant threat is to be disarmed so as to make forgiveness fitting, a certain process must unfold. Perhaps the person who was wronged must come to entertain a psychological state that would allow her to listen to the wrongdoer’s apology and expression of contrition. In this way, the fittingness of forgiveness in the present might depend on the past attitudes of the person who was wronged. But in this case fitting forgiveness only *indirectly* depends on the attitudes of the victim at other times. The psychological

change brings about the conditions that make forgiveness fitting. The process is not essential to fittingness, but instrumental to it. So a synchronic process view must be clearly distinguished from the kind of process view that I am proposing.

For this purpose, distinguish between instrumental and non-instrumental processes. An instrumental process aims at an end-state defined independently of the process itself. If the desired end-state is to have a cake, then the process might involve baking a cake or walking to the store to buy one, depending on which means better serves the desired end-state. By contrast, a non-instrumental process aims at an end-state defined by the process itself.³³ To be the winner of a game of basketball, to obtain a PhD, or to have a friend, is to have gone through processes of certain kinds.³⁴ Of course, historically defined end-states have a-historical features: there are various ways to get a trophy, possess a diploma, or have a drink with Charlie. But for these things to count as winning a game, having a PhD, or having a friend, they must have the relevant history. That is to say: they must have been brought about in a certain way.

To illustrate the difference between the two kinds of processes, consider grief. On an instrumental view of the fitting process of which grief is a part, the end-state of the process might be a return to the bereaved person's emotional and functional baseline (see, for example, Bonanno 2009). Grief is then construed as a step along the way to such a recovery and the fitting duration of grief is determined by its efficacy in bringing about the desired end-state. However, on a non-instrumental view of the fitting process of which grief is a part, the end-state of the process is historically defined by the relevant stages. So if one were to return to one's emotional and functional baseline without passing through the valley of grief, so to speak, then one would not arrive at the fitting end-state of the process *even if there is no other difference between one's current state and the fitting end-state of the process*. On the non-instrumental process view, we cannot skip stages in the process because the end-state is defined by the path we take to get to it. On this view, to determine the fitting duration of grief we must determine the fitting process in which grief is embedded, its different stages, and the way in which the different stages relate to each other so as to form a coherent whole.³⁵

This is not meant as a conclusive argument but as a preliminary sketch of where the rejection of the synchronic view might lead us. The idea that what attitude it is fitting for one to have at a time can generally and directly depend on one's attitudes at other times suggests that the

fittingness of an attitude at a time can generally and directly depend on the fitting process of which it is a part. Furthermore, the fact that the fittingness of one's attitude at a time can *directly* depend on one's attitudes at other times suggests that, in the cases where it does, the fitting process of which the attitude is a part is a non-instrumental process. Finally, once we consider the fitting, non-instrumental processes in which emotions are embedded, we may be able to better describe and understand the evolution of our fitting emotions over time as well as the standards by which one emotion takes the place of another while, in a crucial sense, continuing it.

References

- Bonanno, G. 2009. *The Other Side of Sadness: What the New Science of Bereavement Tells Us About Life After Loss*. New York: Basic Books.
- Brickman, P., Coates, D., and Janoff-Bulman, R. 1978. Lottery winners and accident victims: is happiness relative? *The Journal of Personality and Social Psychology* xxxvi: 917–27.
- Callard, A. 2017. The reason to be angry forever. In *The Moral Psychology of Anger*, edited by M. Cherry and O. Flanagan (London: Rowman & Littlefield International): 123–137.
- Carr, J. 2016. Don't stop believing. *Canadian Journal of Philosophy* 45, 744–766.
- Cholbi, M. 2018. Regret, resilience, and the nature of grief. *Journal of Moral Philosophy*: 1–23.
- D'Arms, J., and Jacobson, D. 2000a. The moralistic fallacy: On the 'appropriateness' of emotions. *Philosophy and Phenomenological Research* 61: 65–90.
- D'Arms, J., and Jacobson, D. 2000b. Sentiment and value. *Ethics* 11: 722–748.
- D'Arms, J., and Jacobson, D. 2009. Demystifying sensibilities: sentimental values and the instability of affect. In Goldie, P. (Ed) *The Oxford Handbook of Philosophy of Emotion*. Online publication 2010: DOI: 10.1093/oxfordhb/9780199235018.003.0027
- Dancy, J. 2004. *Ethics without Principles*. Oxford: Oxford University Press.
- Gaut, B. 2007. *Art, Emotion and Ethics*. Oxford: Oxford University Press.
- Gertken, J., and Kiesewetter, B. 2017. The right and the wrong kind of reasons. *Philosophy Compass* 12 (5): e 12412.
- Goldie, P. 2012. *The Mess Inside: Narrative, Emotion, & the Mind*. New York: Oxford University Press.
- Harman, E. 2009. "I'll be glad I did it" reasoning and the significance of future desires. *Philosophical Perspectives* 23: 177–199.
- Hedden, B. 2015a. Time-slice rationality. *Mind* 124: 449–491.
- Hedden, B. 2015b. *Reasons without Persons*. Oxford University Press.
- Hieronymi, P. 2001. Articulating uncompromising forgiveness. *Philosophy and Phenomenological Research* 62: 529–555.
- Hieronymi, P. 2005. The wrong kind of reason. *The Journal of Philosophy* 102: 437–457.
- Hlobil, U. 2015. There are diachronic norms of rationality. *Thought* 4: 38–45.
- Hofweber, T., and Velleman, D. 2010. How to endure. *The Philosophical Quarterly*: 1–21.
- Howard, C. Forthcoming. The fundamentality of fit. In *Oxford Studies in Metaethics*, edited by R. Shafer-Landau, volume 14, Oxford University Press.
- Kolodny, N. 2003. Love as valuing a relationship. *Philosophical Review* 112: 135–189.
- Kolodny, N. and Brunero, J. 2018. Instrumental rationality. *The Stanford Encyclopedia of Philosophy* (Winter 2018 Edition), Zalta, E. (ed.), URL = <<https://plato.stanford.edu/archives/win2018/entries/rationality-instrumental/>>.
- Maguire, B. 2018. There are no reasons for affective attitudes. *Mind* 127: 779–805.

- Marušić, B. 2018. Do Reasons Expire? An Essay on Grief. *Philosophers' Imprint* 18: 1–21.
- McHugh, C., and Way, J. 2016. Fittingness first. *Ethics* 126: 575–606.
- McMahan, J. 2005. Preventing the existence of people with disabilities. *Quality of Life and Human Difference*. Edited by Wasserman, D., Bickenbach, J., Wachbroit, R. Cambridge University Press.
- Moller, D. 2007. Love and death. *Journal of Philosophy* 104: 301–316.
- Moller, D. 2017. Love and the rationality of grief. In *The Oxford Handbook of Philosophy of Love*, edited by C. Grau and A. Smuts. Online publication date: April 2017. <https://doi.org/10.1093/oxfordhb/9780199395729.013.35>
- Moran, R. 1994/2017. The expression of feeling in imagination. *The Philosophical Review* 103: 75–106. Also published in his *The Philosophical Imagination: Selected Essays*. Oxford University Press, 3–25.
- Moss, S. 2015. Time-slice epistemology and action under indeterminacy. In Tamar Szabó Gendler & John Hawthorne (eds.), *Oxford Studies in Epistemology*. Oxford University Press. pp. 172–194
- Na'aman, O. Ms. Regret and repair. Unpublished manuscript.
- Na'aman, O. Forthcoming. The fitting resolution of anger. *Philosophical Studies*. <https://doi.org/10.1007/s11098-019-01317-w>
- Parfit, D. 1984. *Reasons and Persons*. Oxford: Oxford University Press.
- Podgorski, A. 2016. A reply to the synchronist. *Mind* 125: 859–871.
- Podgorski, A. 2017. Rational delay. *Philosophers' Imprint* 17: 1–19.
- Rawls, J. 1971. *A Theory of Justice*. Cambridge: Harvard University Press.
- Rinard, S. 2019. Equal treatment for belief. *Philosophical Studies* 176: 1923–1950.
- Scanlon, T. M. 1998. *What We Owe to Each Other*. Cambridge MA: Harvard University Press.
- Scanlon, T. M. 2014. *Being Realistic About Reasons*. Oxford: Oxford University Press.
- Schroeder, M. 2007. *Slaves of the Passions*. Oxford: Oxford University Press.
- Schroeder, M. 2010. Value and the right kind of reason. *Oxford Studies in Metaethics* 5: 25–55.
- Schroeder, M. 2012. The ubiquity of state-given reasons. *Ethics* 122: 457–488.
- Setiya, K. 2014. The ethics of existence. *Philosophical Perspectives* 28: 291–301.
- Skorupski, J. 2010. *The Domain of Reasons*. Oxford: Oxford University Press.
- Stroud, S. 2006. Epistemic partiality in friendship. *Ethics* 116: 498–524.
- Suh, E., Diener, E., & Fujita, F. 1996. Events and subjective well-being: only recent events matter. *Journal of Personality and Social Psychology* 70(5): 1091–1102. doi:10.1037/0022-3514.70.5.1091
- Titelbaum, M. 2015. Continuing on. *Canadian Journal of Philosophy* 45, 670–691.
- Wallace, R. J. 2013. *The View from Here: On Affirmation, Attachment, and the Limits of Regret*. Oxford:

Oxford University Press.

Wallace, R. J. Forthcoming. Trust, anger, resentment, forgiveness: on blame and its reasons. *European Journal of Philosophy*. Page numbers refer to penultimate draft.

Way, J. 2012. Transmission and the wrong kind of reason. *Ethics* 122: 489–515.

¹ In the course of writing this paper and developing its main ideas, I have learned from the insights, comments, and advice of various people, in conversation and in writing. I am indebted to Rachel Achs, Dan Baras, Selim Berker, Lindsey Chambers, Jorah Dannenberg, Jeremy Fix, Jan Gertken, Alex Gregory, David Hills, Avi Kenan, Benjamin Kiesewetter, Berislav Marušić, Conor McHugh, Tim Scanlon, Daniel Telech, Jay Wallace, Jonathan Way, Vida Yao, and Bernardo Zacka. Special thanks to Michael Bratman, Colin Chamberlain, Byron Davies, David Enoch, Johann Frick, Barry Maguire, Aleksy Tarasenko-Struc, and two anonymous referees for *Noûs*—all of whom provided incisive written comments on various drafts of the paper. Earlier versions of the paper were presented at the 2018 Humboldt-Southampton Normativity Conference at Humboldt University in Berlin, at the 2018 ‘Bechtel Reunion’ at Northwestern University, and at the Law and Philosophy Workshop at the Hebrew University of Jerusalem. I thank the participants for valuable discussion. The article was written with the support of the Martin Buber Society of Fellows at the Hebrew University of Jerusalem.

² A note about my use of the term “rational”. T. M. Scanlon distinguishes between an attitude being “irrational in the narrow sense” and it being “vulnerable to rational criticism” (Scanlon 1998, 25–30.) A person’s attitude is *irrational in the narrow sense* when it fails to relate appropriately to the person’s other attitudes, for instance, when the attitude fails to cohere with other attitudes. By contrast, a person’s attitude is *vulnerable to rational criticism* when it fails to adhere to objective reasons or to some substantive normative standard. I might coherently pursue an end that I have objective and decisive reason not to pursue, in which case I would arguably be rational in the narrow sense but still vulnerable to rational criticism. Niko Kolodny and John Brunero draw what is essentially the same distinction, between ‘rational coherence’ and ‘reason’, see Kolodny and Brunero 2018. In this paper, I am concerned with the structure of objective standards in light of which emotional change is rationally criticizable. I thank an anonymous referee for urging me to clarify this point.

³ A clarification about the use of the term ‘fitting attitudes’. ‘Fitting’ here is used to mark a type of attitude rather than an evaluation of a token attitude: fitting attitudes in this sense are of the type that can succeed or fail to be fitting, whereas other attitudes cannot even be assessed for fittingness (assessing them for fittingness would be a category mistake.) This usage is borrowed from the common usage of the term ‘rational’: by ‘rational attitudes’ we often mean those attitudes that can succeed or fail to be rational, as opposed to those attitudes with regard to which the question of rationality does not arise. Similarly, by ‘rational agents’ we often mean those agents who can succeed or fail to be rational, by contrast to agents who are not appropriate targets of rational evaluation.

⁴ So in addition to the question of whether value is analyzable in terms of fittingness, there is a question about whether fittingness is analyzable in terms of reasons (see Schroder 2010; McHugh and Way 2016; Howard forthcoming).

⁵ Note, with regard to this last scenario, that it is not merely the case that I have reason to suppress *the expression* of my fear; I have reason to quell some of fear’s constitutive components, such as the urge to flee. This is why I seem to have a prudential reason not to fear the bear. I thank Barry Maguire for urging me to clarify this point.

⁶ I am indebted to an anonymous referee for pressing me to clarify why the accounts of the girl’s lack of regret seem to assume that normally regret rationally persists.

⁷ For a fuller discussion of Hieronymi and Callard on anger, see Na’aman Forthcoming.

⁸ It is a substantive claim that the distinction between right and wrong kind of reasons aligns with the distinction between reasons that make an attitude fitting and reasons that can only justify it overall, but this alignment is widely accepted and it strikes me as very plausible. For more about why it is the case that the two distinctions align in this way, see (Howard Forthcoming).

⁹ Some have argued that moral flaws in jokes and comedies are inevitably comic flaws (Gaut 2007), but the distinction between right and wrong kinds of reasons stands. For instance, even if a morally inappropriate joke is not funny because morally inappropriate, the health-related benefits of laughter would provide another example of a wrong kind of reason to laugh at the joke (though, again, that laughter is healthy may be a good wrong kind of reason to laugh, one that genuinely counts toward the justification of laughter). Nevertheless, some philosophers are wrong-kind reason skeptics, i.e., they deny that wrong kinds of reasons can be normative reasons (Hieronymi 2005; Skorupski 2010; Way

2012). For them it is even more pressing to find an account of fitting change in attitude because on their view there cannot be a rational change in attitude that is not a change in fittingness.

¹⁰ Maguire (2018) argues that fitting reasons are importantly different from reasons for action. I will not rely on these differences in what follows.

¹¹ I discuss the role of this assumption in the debate about the rational resolution of anger and resentment in Na'aman Forthcoming.

¹² Sometimes I will say that a fact *constitutes* a reason and other times that it *gives* a reason. Nothing substantive is implied by alternating between the two expressions. The only reason for doing so is keeping the sentences clear and easy to read. The main substantive point is the distinction between the natural fact and its normative status (as *a reason* or as *giving a reason*.)

¹³ The assumption that the fittingness of ϕ -ing does not *directly* depend on its duration still allows for *indirect dependence*, such as when the duration of ϕ -ing makes it the case that *r* changes. For instance, my fear made me run away from the danger and consequently it is no longer fitting. But it is not the duration of my fear *per se* that changes its fittingness, it is the fact that my fear made me act in a way that removed the reason for it. I am concerned with cases of direct dependence, where the sheer fact that the emotion has persisted changes the fittingness of its continuation.

¹⁴ The example is taken from Richard Moran. Considering everyday emotional responses to the “nonactual”, Moran writes:

Relief, regret, remorse, and nostalgia are, after all, among the paradigm cases of emotional response; and although they are essentially backward-looking, they are not commonly thought to present any special puzzle among the emotions. But at the same time, from within everyday psychology we have also to confront the proverbial injunction against crying over spilt milk, which can itself seem very puzzling. After all, if we can't cry after the milk is spilled, when *can* we cry? Presumably not upon confronting the milk still safely in the bottle. (Moran 1994/2017, 78–79/5–6)

Moran plausibly interprets the injunction as targeting the *shape* of the response to the spilt milk—i.e. no point in crying about what's already done and cannot be changed—but it is worth noting that the proverbial injunction might draw its persuasive force from the *size* of the response. The plausible impression that it is not fitting to be anguished by a minor misfortune as spilt milk lends apparent plausibility to the suggestion that it is not fitting to be concerned *at all* about a past misfortune. Replace “milk” with something more serious and the plausibility goes away: surely there is reason to cry over spilt *blood*. Indeed, there is even reason to cry over spilt milk if one is dying of thirst.

¹⁵ The above two paragraphs are drawn from Na'aman Forthcoming.

¹⁶ I owe the example to Johann Frick.

¹⁷ Pamela Hieronymi writes about the process of forgiveness: “the one wronged might face two distinct tasks: first, achieving the stance of readiness-to-forgive, and second, actually forgiving. Demanding the second before it's *appropriate* can prevent a person from achieving the first” (Hieronymi 2001, 554, my italics). Hieronymi's choice of words here is significant. It is not merely that demanding forgiveness before it is *possible* can be counterproductive, but that demanding forgiveness before resentment has run its *appropriate* course can be counterproductive.

¹⁸ This is compatible with empirical research, cited in Moller 2007, that shows that both good and bad life events tend to have a short-term impact on our subjective well-being. See Brickman, Coates, and Janoff-Bulman 1978, and Suh, Diener, and Fujita 1996.

¹⁹ For discussion of this example I am, again, indebted to Johann Frick.

²⁰ I thank an anonymous referee for raising this type of cases and pointing out the importance of this qualification.

²¹ Hedden (2015a, 2015b) and Moss (2015) present accounts of “time-slice rationality” in the theoretical realm. Against Hedden and Moss, Hlobil (2015), Titelbaum (2015), Carr (2016), and Podgorski (2016, 2017) have argued for fundamentally diachronic norms of belief formation. One important difference between the synchronic view in the theoretical realm and its counterpart in the emotional realm, is that the former is concerned with *subjective justification* while the latter is concerned with *correct representation*. The theoretical synchronist, who holds that what an agent ought to believe at a time does not directly depend on the agent's beliefs at other times, affirms a thesis about subjective norms for belief, taking for granted that a belief ‘p’ is correct iff p is true. Thus, a correct belief can be unjustified and an incorrect belief can be justified, depending on one's subjective reason for belief. The focus on subjective justification in the theoretical realm allows the synchronist to hold that what I (subjectively) ought to believe at t2 about my belief at t1 is not directly determined by what I believed at t1 but by my evidence at t2 about what I believed at t1 (Hedden 2015b, 8). This idea does not carry over well to a synchronic view about the fittingness of affective attitudes. It is not plausible that my past envy determines the fittingness of my present shame by shaping my present evidence about my past envy. Rather, what makes my shame about my past envy fitting *is* my past envy, just as what makes my belief that p correct is that p is true. In light of this, and as I explain in the main text, the synchronic view of fittingness insists that while, in this instance, the fittingness of my shame directly depends on my past envy because my past envy is its object,

fittingness does not *generally* depend on the agent's attitudes at other times. I thank Brian Hedden and an anonymous referee for discussion.

²² Note that the object view of fittingness is about *all* fittingness-making facts, not only fitting reasons. It therefore differs significantly from the view that right-kind reasons (which, I am assuming, are equivalent to fitting reasons) are "object-given." The view that right-kind reasons for an attitude are object-given reasons—i.e., they are reasons that bear on the object of the attitude—leaves room for the possibility that the background conditions that enable the object-given reasons do *not* bear on the object. This possibility is ruled out by the object view of fittingness. So the object view entails the view that right-kind reasons are object-given, but the latter does not entail the former. For a description and criticism of the view that right-kind reasons are object-given, see Schroeder 2012. For a helpful overview of the different attempts to give a theory of the right/wrong kind reason distinction, see Gertken and Kiesewetter 2017.

²³ I thank an anonymous referee for raising this possibility and urging me to address it.

²⁴ Another possibility, which I do not consider in the text, is that the content of the attitude includes, in addition to the relevant evaluative property, also relational properties that appeal to the agent's past attitudes. This view can preserve the object view while avoiding both the implication that the evaluative property depends on the past existence of the emotion and the implication that fittingness is synchronic. This version of the object view effectively builds the background conditions for the fitting reason into the reason. I am not sure it is plausible to think of fitting reason in this way, but insofar as it is, the view does not seem significantly different from the view I am arguing for in this paper. It simply invokes an unusual understanding of fitting reason for emotions. I thank Alex Gregory for raising this possibility and discussing it with me.

²⁵ The implication follows on the assumption that the attitude is about the facts that constitute the reason for it *and not* about the background conditions that enable these facts to constitute reason for the attitude. I do not argue for it here, but I think this assumption is plausible. The fact that we are friends is arguably a background condition that explains why your absence from my birthday party is a reason for fitting insult and resentment. But what I am insulted *by* and what I resent is your absence from my birthday party, not our friendship. So our friendship is not the object of my insult and resentment, but it is part of the rational explanation of these attitudes. I thank an anonymous referee for noting this assumption.

²⁶ In light of this formulation of the object view, it might seem that the view cannot accommodate a wide range of emotions that represent nonactual properties, such as regret and anger about events that ended long ago, admiration of people who no longer exist, etc. Indeed, if only actual properties can be the objects of fitting emotions, it seems the problem of backward-looking emotion does not even get off the ground. But I think the correct understanding of the object view "actualizes" the relevant properties. For example, if resentment is about a wrong done in the past, then the fact that the past action was wrong still obtains long after the action had been completed; if grief is about the loss of a loved one, then the fact that the loss occurred still obtains long after the person died; similarly, a person can remain admirable long after her death. Fear, by contrast, is about *present* danger, and that is why it is no longer fitting once the danger dissipates. This way of talking about the relevant evaluative properties blurs the distinction between *facts* and *true propositions*, a move which might be objectionable for independent reasons (do emotions represent true propositions?), but I believe this is how the object view should be understood. This interpretation makes sense in light of the claim I make below, namely, that the object view is modeled on true belief. I thank Dan Baras for discussion.

²⁷ Notice that D'Arms and Jacobson conflate here the correct representation of evaluative properties with the subjective justification of an emotion given one's evidence. Whether regret is fitting depends on whether its object is regrettable, but the object might be regrettable whether or not one has sufficient evidence for this fact. I understand fittingness as correct representation. I thank Avi Kenan for bringing this conflation to my attention.

²⁸ For an argument that friendship can give reason for belief, see Stroud 2006. For an argument that justification of belief is in principle the same as justification of action, see Rinard 2019.

²⁹ There is good reason to think Wallace in fact has the first reading in mind. In a forthcoming paper, Wallace says the following about the rationality of forgiveness: "How can it ever be reasonable for us to forgive the wrongdoer if the reasons to which angry disapprobation directly responds remain indefinitely in place? Answers to these questions will remain elusive so long as we attend solely to considerations of aptness or warrant for such reactive attitudes as resentment" (Wallace Forthcoming, 23). He goes on to say: "To forgive another, it is natural to suppose, is to forswear the resentment that would be fitting or warranted under the circumstances, and this is a matter of managing the attitude in a certain way" (23). On Wallace's view, resentment remains fitting while there are other, extraneous reasons that explain the change in the overall rational status of the attitude. This suggests that his account of the young girl's diminishing regret is also meant as an account of the extraneous reasons that change the rational status of her regret, while allowing that regret about her decision remains fitting.

³⁰ I elaborate on regret in general and on Wallace's account in particular in Na'aman Ms.

³¹ I thank David Enoch and an anonymous referee for raising this point and urging me to address it.

³² One might wonder at this point about my claim that fittingness should be understood as correct representation of evaluative properties. Given the complexity of fitting processes, it is not obvious in what sense do they *represent* the properties of the object. Such processes seem to involve various attitudes, thoughts, and actions that are about things other than the object in response to which the process is fitting. To insist that the process *as a whole* represents the evaluative properties of the object seems to stretch the notion of representation to its limit. My response is concessive: if processes cannot be understood as representing the evaluative properties of the objects to which they are a fitting response, then so much the worse for the notion of fittingness as correct representation. A different notion of fittingness must then be employed in order to account for fitting processes. Conversely, I take my argument to show that if fittingness *is* correct representation, then processes can correctly represent the evaluative properties of the objects to which they are a fitting response. I thank Rachel Achs, Conor McHugh, and Jonathan Way for discussion of this point.

³³ We find a useful analogy to the distinction between instrumental and non-instrumental processes in John Rawls' distinction between, on the one hand, perfect and imperfect procedural justice and, on the other hand, pure procedural justice. The required outcome of perfect and imperfect procedural justice can be specified independently of the procedure that brings it about. Perfect procedures are *guaranteed* to bring about the required outcome while imperfect procedures are only *likely* to bring about the required outcome. For example, a procedure of criminal justice, which aims to convict the guilty and only the guilty, cannot guarantee success but might be sufficiently reliable and therefore count as a procedure of imperfect procedural justice. However, pure procedural justice obtains when there is no criterion for the right outcome other than the execution of a correct or fair procedure (Rawls 1971, 86). For example, when applied fairly, a gambling procedure renders its consequences legitimate, whatever they are.

³⁴ Although some processes are more clearly defined than others, such as getting a PhD in comparison to forging a friendship. For the view that friendship entails a relationship with a certain kind of history, see Kolodny 2003.

³⁵ Peter Goldie has argued that grief itself is best understood as a process, and in particular a narrative process. Grief, Goldie says, is

a complex pattern of activity and passivity, inner and outer, that unfolds over time, and the unfolding pattern over time is explanatorily prior to what is the case at any particular moment, and moreover, explanatorily prior to any particular mental state or event at any particular moment that is part of the process. This process needs a history, and more specifically it needs a narrative account ... because narrative accounts have such powerful explanatory, revelatory, and expressive powers. (Goldie 2012, 69)

Such a view of grief can fit nicely with the process view of fittingness I am suggesting. I have been focusing on the diachronic structure of fittingness and assuming for simplicity that emotions *endure*—that their identity is fully determined at every moment at which they exist. But it might very well be that many emotions *perdure*—their identity at every moment depends on their identity over time. If the latter is the case, then not only the fittingness of emotions is diachronic but also the emotions themselves. For another recent process account of grief, see Cholbi 2018.