Emotions and Process Rationality

Oded Na'aman

Abstract
Some epistemologists hold that all rational norms are fundamentally concerned with the agent's *states or attitudes at an individual time* [Hedden 2015, 2016; Moss 2015]; others argue that all rational norms are fundamentally concerned with *processes* [Podgorski 2017]. This distinction is not drawn in discussions of emotional rationality. As a result, a widely held assumption in the literature on emotional rationality has gone unexamined. I employ Abelard Podgorski's argument from rational delay to argue that many emotional norms are fundamentally concerned with *emotional processes*. I also claim that the main response available to the synchronist about belief is not available to the synchronist about emotions and, therefore, fundamental process norms are more plausible than epistemologists tend to believe.

Keywords: diachronic norms; rationality of emotions; fittingness; processes; reasons

0. Introduction

Process rationality is the view that rational norms are fundamentally process-oriented, not state-oriented. There are at least two possible versions of process rationality. *Strong process rationality* is the view that *all* rational norms are fundamentally process-oriented. I shall neither argue for this version of process rationality nor contest it. It is, however, endorsed by Abelard Podgorski, to whom we will return later. *Weak process rationality* is the view that *some* rational norms are fundamentally process-oriented. Brian Hedden, to whom we will also return later, *explicitly rejects* weak process rationality: he argues there are *no* fundamental process-oriented norms. Hedden has many hidden allies. Most contemporary philosophers writing about norms for emotions unwittingly agree with him, which is to say that they *implicitly reject* weak process rationality. I will argue that Hedden and the prominent implicit view in the literature on emotions are mistaken: once the question is raised with regard to emotions, weak process rationality appears highly plausible.

1. The Common View of Emotions

Following D'Arms and Jacobson's work [2000a, 2000b], the normativity of emotions is often described in terms of *fittingness*. According to one interpretation, an emotion is fitting to its

object (or, the object *merits* the emotion) iff the emotion properly represents the relevant evaluative properties of the object. Shame is fitting to what is shameful, disgust to what is disgusting, fear to what is fearsome, amusement to what is amusing, etc. Fittingness norms for emotions, then, are norms for correct representation [Rosen 2015; Howard 2018: 6]. On another interpretation, fittingness norms are internal to—or constitutive of—the attitude in question: they are essential to what the attitude is [McHugh and Way 2016: 583–84; Sharadin 2015: 387]. Some maintain that fittingness is best explained in terms of value or reasons while others hold that fittingness is a fundamental normative relation that cannot be further analyzed in normative terms [Chappell 2012; McHugh and Way 2016; Howard 2019]. I do not wish to take a stand on these issues here, but simply employ the notion of fittingness to circumscribe the relevant kind of assessment that is my subject in this paper.

Though norms of fittingness are especially important for appraising emotions, D'Arms and Jacobson argue they do not *exhaust* the norms for appraising emotions. Fitting envy, for instance, might be morally or prudentially objectionable [D'Arms and Jacobson 2000b: 71–2]. This is because some reasons not to envy have nothing to do with whether the object of envy is in fact *enviable* or whether it *merits envy*. So the moral, prudential, or overall normative status of an attitude are distinct from its fittingness status. Accordingly, reasons for an emotion that do not bear on its fittingness are sometimes called *wrong-kind reasons*, whereas fit-related reasons are called *right-kind reasons*.

D'Arms and Jacobson also distinguish between an attitude being *fitting* and it being *warranted* [D'Arms and Jacobson 2000a: 745]. Even if the object of my envy merits envy, I might lack sufficient evidence for this fact, in which case my envy, though arguably fitting, would not be *warranted.* Alternatively, I might have sufficient evidence that I'm in danger but in fact not be. In this case, fear is warranted but not fitting. Therefore, that an attitude is warranted means it is *rationally justified as fitting*, but it does not entail that *it is in fact fitting*, and vice versa. This distinction hasn't been properly appreciated in the literature on fitting emotions, where it is often assumed that if an attitude is rationally justified as fitting, then it is in fact fitting. [1]

---

[1] I thank Avi Kenan for discussion. He expands on this distinction and on the failure to appreciate it in the literature on fittingness in Kenan [ms].

The distinction between *warrant* and *fit* is an instance of a more general distinction between *what is rational* and *what is supported by objective normative reason*. The fact that the lemonade is arsenic-laced is a conclusive reason for you not to drink it. This fact remains a conclusive reason for you not to drink the lemonade even if you are in no position to see that the lemonade is arsenic-laced. But since you are in no position to see that the lemonade is arsenic-laced, you would not be irrational in drinking it.[2] So meta-ethicists distinguish rational φ-ing from φ-ing that is supported by objective normative reason. Rationality requires responding to *apparent* reasons. Correspondingly, 'warrant' denotes a rational assessment and 'fit' denotes an assessment of objective normativity. An attitude is warranted when it is a response to *apparent fit-related reasons*. If facing danger makes fear fitting (i.e., if danger is a fit-related reason for fear), then fear is *warranted* for an agent if the agent is facing *apparent* danger. Norms of fittingness inform norms of warrant.

All this is to say that fittingness is a distinctive normative status that does not settle on its own whether one ought to feel the emotion all things considered, nor whether the emotion is morally or prudentially advisable, nor whether the emotion is warranted. In this essay I put to one side questions about moral, prudential, and overall justifications of emotions; I focus instead on fittingness and, by extension, warrant.

The primary question at hand is whether all norms of fittingness fundamentally apply to an agent's emotional state at a time irrespective of the agent's attitudes at other times. Since this question has not been taken up in the literature, my first task is to demonstrate that a positive answer to this question is often presupposed. I will consider different discussions and argue that they are constrained by the implicit assumption that fittingness norms always and fundamentally apply to an agent's emotional state at an individual time independently of the agent's attitudes at other times.

First example: the instability of affect. Justin D'Arms and Daniel Jacobson maintain that emotional responses are 'notoriously fickle' and yet the values they track are relatively stable [D'Arms and Jacobson 2009]. This disparity, they say, gives rise to an apparent puzzle. If the joke remains funny but our amusement fades, in what sense does amusement track funniness? To maintain that our emotions are fitting to the evaluative properties of their objects, we must be able

---

[2] The example is from Sylvan 2015.

to make sense of the normal divergence between the instability of emotions and the stability of the properties they purportedly track.[3]

D'Arms and Jacobson consider several possible replies to the puzzle. The first denies that the evaluative properties are stable; the second denies that emotions are the relevant evaluative response to evaluative properties; and the third concedes that our emotions are very often not fitting but offers a systematic explanation of the divergence.

Consider the first response. Perhaps the joke ceases to amuse because it is no longer funny. This response denies the stability of the funny: a joke is funny *for you* only when it amuses you. Or perhaps the joke is funny for you only when you hear it for the first time. But, D'Arms and Jacobson note, we often judge that a joke is funny even when it no longer amuses us [D'Arms and Jacobson 2009: 591]. Without going into the details of their arguments, it seems plausible that our emotional responses often fluctuate even when the relevant evaluative facts remain unchanged.

The second response to the puzzle identifies our sense of humor with our sincere judgment that the joke is funny rather than with the pattern of our amusement. Amusement fades, but one's judgment that the joke is funny is as stable as the funniness of the joke. The instability of affect is compatible with stable judgment, which is the relevant fitting response to the stable evaluative property. This response will not do either, D'Arms and Jacobson claim, because our judgments can fail to reflect our underlying sense of humor. For example, my despondent mood can distort my sense of humor and make me judge a comedy that would otherwise amuse me dull. In such a case, my judgment fails to adhere to my sense of humor. Since my *sense* of the evaluative property can diverge from my *judgment* of it, the two must be distinct.

Finally, D'Arms and Jacobson's own solution is to admit the disparity between emotions and the values they track while offering a systematic account of it. They posit 'underlying sensibilities' that explain 'the coherent patterns of reaction to which people are prone in virtue of which they can be understood as having a distinctive perspective on various sentimental values, rather than just a welter of emotional responses' [ibid.: 593]. On their understanding, sensibilities issue in affective responses, and by looking at the right responses we discern the underlying

---

[3] To be sure, D'Arms and Jacobson are particularly concerned with the implications of this disparity for a metaethical thesis they favor, namely, that so-called 'sentimental values' depend essentially on their associated evaluative responses. But even putting this thesis aside, we may still wonder whether the instability of emotions coupled with the stability of the evaluative properties that make emotions fitting implies that our emotional responses are very often not fitting.

sensibility. In the normal case, one's evaluative judgment arises from one's sensibility by way of sentimental response [ibid.: 596]. But our sentimental responses are also influenced by other factors, unrelated to the relevant sensibility. For instance, D'Arms and Jacobson say that if overexposure to a joke impacts one's amusement, then repetition is 'an obscuring factor with respect to a joke's funniness' [ibid.: 597] because it affects the sentimental response without revealing the underlying sensibility. They further argue that there are various obscuring factors that explain why affective responses often fail to be fitting by the agent's own lights, that is, according to the agent's own sensibility [D'Arms and Jacobson 2009, 2014]. So there is a disparity between the value and the emotional response, but it can be accounted for in a way that preserves the idea that emotions track (sentimental) values.

However, the three responses described above are not exhaustive. It is also possible that some fluctuations in our emotional responses are part of a fitting emotional response to stable values. To appreciate this possibility it would help to appreciate why it is overlooked. D'Arms and Jacobson assume that if amusement in response to a funny joke is fitting, then it remains fitting independently of how many times one hears the joke (as long as the joke remains funny). What is the basis of this assumption? We get a clue from the comparison they draw with perception. They quote Elijah Milgram's observation that 'the scarf still looks red the hundredth time you see it, whereas the joke no longer gets a laugh the hundredth (or usually even the second) time through' [Millgram 1999: 253]. The perceptual state of seeing red obtains whenever redness is encountered whereas amusement does not obtain whenever the funny joke is encountered. The comparison seems to give rise to the puzzle about the instability of affect, but only on the assumption that amusement is fitting in response to the funny in the same way that the perceptual state of seeing red is 'correct' in response to the red scarf. Specifically, both kinds of responses are assumed to be mental states that are merited independently of the agent's mental states at other times. The assumption is crucial in order to explain why the fact that one had already been amused by the funny joke is not relevant to the fittingness of future amusement. That I already saw the scarf as red before is not relevant to whether I correctly see it as red next time; that I laughed at the joke before is not relevant to whether I fittingly laugh at it next time.

But the fittingness of amusement might be different from correctly seeing the scarf as red. It is possible that unlike correctly seeing red, the fittingness of amusement at a time depends not only on its object at the time but on the agent's attitudes over time. Consider a different analogy.

Applause in response to a musical performance are normally taken to represent the audience's appreciation of the value of the performance as such. But when the audience's enthusiastic applause finally wind down, this is no indication that the audience has changed its minds about the performance, nor that there was some rational or psychological interference. As if in the absence of interfering considerations and concerns, the audience would have continued to fittingly applaud indefinitely. Similarly, were the performer to repeat the very same number in the very same way, the audience would likely applaud her with less verve, or for a shorter duration of time, and this, too, need not indicate that the audience judges the second performance less laudable than the first or that it falls short of a fittingly enthusiastic response. The reason is that the 'correct' audience response to a wonderful musical performance is temporally extended and sensitive to diachronic context: the fact the audience has already applauded the performance is relevant to fittingness of its continued applause.

Like the audience's applause, amusement might be fitting as a first response to a funny joke, but then, depending on the funniness of the joke, by the fifth hearing amusement might not be fitting anymore. Similarly, amusement tends to fade, and it can do so fittingly or not. It might not be fitting to be amused by a moderately funny joke for a whole week, while the gradual diminishment of amusement might be proportional to the funniness of the joke. These are existential, not general claims: if the fitting response to *some* jokes is temporally extended, then it is a mistake to assume that if a joke remains funny, amusement remains fitting.

D'Arms and Jacobson assume without argument that all norms of fitting emotions are fundamentally about the agent's emotional state considered at an individual time irrespective of her attitudes at other times. If this view is false, fluctuations in affect might themselves be part of a fitting response to a stable sentimental value. Of course, the mere conceptual possibility that the instability of affect is fittingly responsive to stable value is only the very beginning of a response to the puzzle. But all I wish to establish at this stage is that this possibility has been overlooked.

Second example: the diminution of backward-looking emotions. 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 merited 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.

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: 358-59]. 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 fitting. Given that the young girl's regret would be initially fitting, what might rationally merit the girl's eventual affirmation of her choice to have a child when she did? The puzzle such a case raises is premised on the assumption that an unjustified or wrong decision normally renders regret fitting *indefinitely*. Thus, the accounts offered to explain the case imply that in normal cases of regret, where attachments and personal values do not change and no person comes into existence, fitting regret fittingly persists.

Finally, Pamela Hieronymi [2001] articulates a challenge for accounts of forgiveness. Forgiveness is generally understood as the rational forgoing of fitting resentment. Hieronymi asks: Given that resentment is merited by a past wrong, as well as by the moral accountability of the wrongdoer and the moral standing of the victim, and given that these facts do not change, how can the diminution of fitting resentment be *fitting*? 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 fitting when the threat is removed—through apology, for instance—and there is nothing more to protest.[4]

The problem of the diminution of backward-looking emotions can be seen as an instance of D'Arms and Jacobson's puzzle about the instability of affect. The value of the object is stable because frozen in the past, while the emotion that constitutes a fitting evaluative response to the value keeps evolving. The various responses to the problem of backward-looking emotions also take similar avenues to the ones D'Arms and Jacobson lay out with regard to the instability of

---

[4] For a fuller discussion of Hieronymi's account of forgiveness see Na'aman [2020].

affect. Either the object of the emotion changes despite first appearances and thereby merits the corresponding diminution in emotion (e.g. Hieronymi [2001]); or the initial emotion is not essential to a fitting response and therefore its diminution is not indicative of a failure of fittingness (e.g. Bittner [1992]); or, finally, there is a failure of fittingness, but also a systematic normative explanation for it (for example, prudential and moral reasons might justify letting go of resentment, or grief, even if these emotions remain fitting, see Moller [2017] and Wallace [2019]).

Like D'Arms and Jacobson, philosophers writing on the problem of backward-looking emotions fail to consider the possibility that the diminution of the emotion is itself part of a fitting response to a stable value. For example, it is arguable that resentment in response to a slight can fittingly diminish with time, just as a sudden burst of joy upon receiving good news might fittingly decrease. The trajectory of such emotions can be fitting to their objects without any change to the object itself. Again, this possibility does not immediately solve the problem of backward-looking emotions, but the fact it has not been considered is telling. Thus, in different discussions of emotional normativity there is an implicit assumption that all fundamental norms of fittingness apply to an agent's emotional state at an individual time independently of the agent's attitudes at other times.[5]

2. The Argument from Rational Delay

Abelard Podgorski claims that 'most recent discussion of rational requirements on mental behavior concerns, explicitly or implicitly, norms that are *synchronic*, taking agents to be rational or irrational in virtue of states they are in at individual times' [Podgorski 2017: 1]. Similarly, in the first section of this paper, I argued that most recent discussion of normative requirements on *emotional* behavior concerns, implicitly, norms that are synchronic, taking agents' emotions to be fitting or not in virtue of states they are in at individual times.

However, while I have been focusing on norms of *fittingness*, the epistemological literature is concerned with norms of *rationality*. To clarify the relation between the two kinds of norms, recall the distinction between fittingness and warrant. If I am not in a position to learn about the dire consequences of my decision, my lack of regret might be *rationally justified as fitting* (warranted), but the decision is in fact regrettable due to its consequences, so regret is (objectively) *fitting*.

---

[5] I discuss backward-looking emotions extensively and argue in detail for the claims I make here in Na'aman [forthcoming].

Accordingly, an attitude is warranted due to *apparent* fit-related reasons and it is fitting due to *objective* fit-related reasons.

So norms of warrant are norms of fittingness qualified by what is apparent to the agent. If being wronged makes a certain process of resentment *fitting*, then that process of resentment is *warranted* for an agent if the agent was *apparently* wronged. Just as the fitting diminution of resentment might not entail a change in the evaluative property that merits it, the warranted diminution of resentment might not entail a change in the *apparent* evaluative property that merits it.[6] If the synchronic view of norms of fittingness is false, then so is the synchronic view of norms of warrant. This connects the discussion of emotional normativity in the previous section to the discussion of rational norms in this section.

Now consider a taxonomy of views about rational norms, which I adopt from Podgorski [2017]. *A synchronic, state-oriented view of rationality* maintains that rationality fundamentally governs the relation between one's attitudes and reasons at individual times, and all rational norms that govern behavior over time can be derived from these more fundamental synchronic norms. This picture of rationality is not only synchronic but also state-oriented, because it maintains that rationality fundamentally governs an agent's mental states. A paradigmatic example of a synchronic, state-oriented norm is the norm of *Belief Consistency*, which requires that, at each time, an agent not hold beliefs that are inconsistent with each other.

A diachronic, state-oriented view of rationality maintains that some rational norms fundamentally govern the relation between an agent's states or attitudes at two or more times. A paradigmatic example of a diachronic, state-oriented norm is *Conditionalization*, which holds that an agent's credal state at one time must stand in the right relation to her credal state at earlier times (in particular, that her credence in P, once she has evidence E, is equal to her prior conditional credence in P/E) [Podgorski 2017: 2].

Process rationality maintains that rational norms apply directly to an agent's processes and not to the mere relation between the agent's states and attitudes. This view is like the second view in being diachronic, for it governs the agent's mind over time; but it is unlike both other views in

---

[6] There are ongoing debates about what apparent (or 'subjective') reasons amount to, but for the purposes of this essay I need not take a stand on this issue. See, for example [Schroeder 2004, 2007, 2009; Setiya 2004; Way 2009, 2012; Markovits 2014; Whiting 2014; and Sylvan 2015; Kiesewetter 2017; Lord 2018].

being process-oriented. As an example of a process-oriented norm, Podgorski gives *Consideration*: 'When considering whether P, one is rationally required to take into account only the evidence for and against P, and conclude by believing P only if the evidence considered sufficiently favors P' [ibid.: 2].

We therefore have a twofold distinction. A distinction between synchronic and diachronic norms, and a distinction between state-oriented and process-oriented norms. All synchronic norms are state-oriented, but diachronic norms might be either state- or process-oriented. Furthermore, different possible views of rationality are shaped by whether the view admits the existence of only one, or two, or all three types of norms, and which types of norms the view denies.

As I mentioned at the outset, Podgorski [2017] endorses strong process rationality, the view that *all* rational norms are fundamentally process-oriented. Titelbaum [2015] and Carr [2016] endorse weak diachronic, state-oriented views of rationality, conceding the existence of synchronic norms alongside diachronic, state-oriented norms. Hedden [2015] and Moss [2015] subscribe to strong synchronic views of rationality, denying the existence of diachronic norms, whether state- or process-oriented. I wish to argue for weak process rationality: the view that some rational norms are fundamentally process-oriented.[7] My view is compatible with strong process rationality as well as with weak diachronic, state-oriented rationality. However, my view is incompatible with the strong synchronic view. (see Figure 1 below)

View of rationality → Type of norm ↓	Strong synchronic view (Hedden, Moss, and rationality of emotion literature)	Weak diachronic state view (Carr, Titelbaum)	Weak process rationality (my view)	Strong process rationality (Podgorski)
Synchronic, state-oriented norms	Yes	Yes	No commitment	No
Diachronic, state-oriented norms	No	Yes	No commitment	No
Diachronic, process-oriented norms	No	No commitment	Yes	Yes

(Figure 1)

---

[7] Ulf Hlobil [2015] argues there are diachronic norms that govern rational transitions between belief states. Although Hlobil does not distinguish between diachronic state-oriented norms and process-oriented norms, his argument supports the existence of the latter kind of norms. However, unlike Podgorski, Hlobil allows for the existence of synchronic state-oriented norms. His view is therefore similar to mine: he subscribes to weak process rationality.

To argue for weak process rationality, I will adapt an argument that Podgorski offers for strong process rationality—the argument from rational delay—and consider its relevance to emotional norms. Though I suspect that Podgorski's argument falls short of establishing strong process rationality, I believe it offers support for weak process rationality and that, as such, it is even more powerful when applied to norms for emotions.

Podgosrki begins by considering the following norm:

Synchronic State-Oriented Reasons (SSR): If an agent's reasons at t favor attitude X, she is rationally required to have X at t. If her reasons at t disfavor attitude X, she is rationally required not to have X at t. [Podgorski 2017: 3][8]

This is a state-oriented norm because it governs the states of the agent in light of her (apparent) reasons, rather than governing a process or activity that unfolds over time. It is a synchronic norm because it appeals to the agent's attitudes at a single time. Podgorksi focuses on a version of this norm that takes belief as the relevant attitude, but he suggests that it might be applied to reasons for intention, desire, gratitude, etc. [ibid.: 4]. In the next section I will take up this suggestion, applying SSR to emotion.

Podgorski argues that SSR has an implausible implication. Upon encountering new reasons, it takes time to update our attitudes. Facts that constitute reason to change our beliefs must be taken into consideration and processed in order to yield the required revision. But SSR entails that *any* update time is rationally flawed. For in the time it takes one to update one's beliefs, it is already true that one has reason to believe otherwise. Any time an agent has reasons without having the attitudes the reasons favor is a time at which the agent is failing to satisfy SSR.

Podgorski argues that humans do not have the psychological ability to satisfy SSR by updating concurrently with their reasons. But even if we had the psychological ability to do so, he points out, insofar as our attitudes are formed *in response* to the considerations we have received,

---

[8] I made small changes to Podgorski's formulations in order to avoid ambiguities that might cause confusion.

they always follow whatever they are responding to. The very idea that rational norms can *guide* or *govern* behavior implies that we *follow* rational norms.[9] But no matter how fast we update, as long as updating takes any time at all, we fail to satisfy the norm. To avoid irrationality we must change our attitudes at the moment we gain our reasons. Podgorski claims that a creature whose attitudes do not *respond* to its reasons but simply covary with them, is 'a creature whose cognitive behavior is no longer recognizably human, and it seems perverse to treat its behavior as a constraint on us' [ibid.: 6]. In short, the phenomenon of *rational delay* is entailed by the idea that we can rationally *follow* norms and *respond* to our reasons—an idea normally taken as a truism. So *if* all synchronic norms are like SSR in being incompatible with rational delay, the synchronic view comes at a hefty theoretical price. And Podgorski argues that, indeed, all synchronic norms are incompatible with rational delay. Insofar as a synchronic norm prescribes an attitude when a certain fact obtains, it leaves no room for *responding* to the relevant fact and therefore no room for rational delay.

In response, it might seem natural to endorse diachronic state-oriented norms. These are norms that govern the agent's states and attitudes at different times. Consider, for example, a diachronic version of SSR, DSR:

Diachronic State-Oriented Reasons (DSR): An agent is rationally required, if her reasons at t1 favor attitude X, to have attitude X at t2 (some appropriate time after t1). [ibid.: 10]

This formulation seems to allow for rational delay by referring to the agent's attitudes at different times. The problem is to account for the 'appropriate time' clause. If the gap between t1 and t2 is too small, the principle fails to avoid the argument against synchronic norms and precludes rational delay, but if the gap is too large, the principle allows for *irrational delay*—a slow, rationally flawed update. Call this 'the specification problem'.

It might seem possible to specify the appropriate time gap by appealing to what is psychologically possible [Kiesewetter 2017: 65–8]. But Podgorski argues that this is far from sufficient. For any threshold we specify, we can imagine contingent circumstances that undermine it. A person who is knocked unconscious before she finishes updating is not thereby made irrational.

---

[9] The point turns on taking 'following a norm' to mean more than 'behaving in accordance with a norm'. 'Following' here implies behaving in accordance with the norm *because of* the norm. And it is also crucial that one can only behave 'in accordance with the norm because of the norm' *after* the norm applies. So if one only possesses a reason once one is able to behave in accordance with it, it follows that one can behave in accordance with one's reason concurrently with possessing it. However, for the rational delay objection to come through, it is enough that in a significant number of cases we have reason for attitude X before we can be reasonably expected to form it. I thank Aaron Elliott for valuable comments on this point.

Other factors related to the agent's particular capacities and circumstances also seem relevant for determining the appropriate update time. Indeed, sometimes an agent might be barred from ever completing her update through no fault of her own. Podgorski argues that once we factor all these variables into a diachronic norm, the norm ceases to be informative. And while such a norm might be true, it seems more plausible as a corollary of more fundamental process norms.

There is more to say, of course, about the prospects of diachronic state norms, but for my purposes this is sufficient. Whether or not the appropriate update time can be usefully specified, Podgorski claims there is a way of avoiding the problem altogether: by positing process-oriented norms. Processes are not combinations of attitudes at different times, but 'causally continuous activities—ways of reflecting upon and questioning attitudes or performing other cognitive labor' [Podgorski 2017: 13]. Explicit reasoning is, according to Podgorski, a paradigm example of a process governed by rational norms.

State-oriented views can grant the importance of processes in allowing us to satisfy rational norms, but they still maintain that the most fundamental norms govern mental states. As a result, they run into the problem of rational delay. Process norms, by contrast, do not govern mental states directly; they govern only mental activities, which, in turn, produce attitudes. Podgorski sums the point nicely: 'If fully rational agents, in virtue of their rationality, manifest some regularity in their attitudes, it is because those patterns are the shadows cast by properly functioning processes' [ibid.: 14].

Thus, Podgorski argues (in more detail than I can do justice to here), that process norms avoid the problem of rational delay by avoiding the prescription of states altogether. Plausible process norms tell us what steps to take in deliberation given various circumstances, but to do so they need not require any combination of attitudes at individual times nor tell us how long each step should take. Nevertheless, process norms can still condemn an update that takes too long, only indirectly. What explains the rational flaw in a slow update is not *that* it is slow, but *why* it is slow. For instance, a slow update might be irrational because the process was inappropriately interrupted or carried out incorrectly. Rational and irrational delays are shadows of processes properly and improperly carried out.[10]

---

[10] It is not clear, however, that an indirect explanation can be provided for *any* temporally flawed update. Moreover, process norms might have their own share of specification problems. Even if process norms do not give rise to

Again, much more needs to be said to defend Podgorski's solution to the problem of rational delay and his view that *all* rational norms are process norms.[11] But next I argue that the normative and rational primacy of processes over states is even more obvious in the case of affective mental activity than it is in the case of doxastic mental activity.

3. Emotions and Rational Delay

As D'Arms and Jacobson alarmingly acknowledge, emotional states are especially fickle and fleeting [D'Arms and Jacobson 2009: 588]. Even when an emotional state endures, it rarely endures without change and normally evolves according to a familiar trajectory: amusement gradually subsides, love intensifies (though it might also dwindle), frustration grows, etc. Some emotions *burst*, others *crop up*; some *evaporate abruptly*, others *fade*. These affective movements are themselves subject to normative and rational evaluation and they are not few and far between; indeed, my point is that such emotional movements are the primary subject of emotional norms. My main claim in this section is that we cannot plausibly abstract away from the movement of emotion and determine whether an agent's emotion is fitting at a time only on the basis of her emotional states considered at individual times. Thus, norms for emotions are most plausibly construed as process norms from which norms about emotional states at individual times can be derived.

Consider again Podgorski's example of SSR. Suppose that attitude X is an emotion. Fit-related reasons for an emotion are those evaluative facts in response to which the emotion is fitting. When those evaluative facts are apparent to an agent, the emotion is warranted, i.e., it is rationally justified as fitting. When applied to fitting amusement, for instance, SSR maintains that if a funny joke is an apparent fit-related reason at t in favor of amusement, then amusement is rational and lack of amusement is irrational at t.[12] It follows that if it takes me any time at all to 'get the joke'

---

ambiguities about time gaps between attitudinal states, there would likely be other ambiguities about their proper application. These doubts notwithstanding, what I want to take from the argument is the idea that often process norms are more illuminating than diachronic state-norms. I thank John Brunero for pressing me on this point.

[11] Podgorski himself admits that there might be state norms that do not govern *the relation* between states at all, such as the norm that precludes having the belief "P & ~P" [Podgorski 2017: 12]; he thereby qualifies the generality of his view.

[12] Here and in what follows I am using 'rationality' to refer only to rational norms of *warrant*. In principle, there might also be rational norms that pertain to the agent's responsiveness to wrong-kind reasons, but I leave those aside.

and be amused by it, then my response is *rationally flawed*. For according to SSR, in the time it takes me to get the joke amusement is already rational but I am not (yet) amused.

As we saw, Podgorski argues that adopting a belief in the very moment one has reason for it is (1) psychologically impossible, (2) incompatible with the idea of responding to reasons, and (3) that creatures who can do this are not recognizably human. These claims are also true of being amused by a funny joke in the very moment one encounters it. Instantaneous amusement is psychologically impossible, incompatible with the idea of responding to the joke, and creatures who can be instantaneously amused are not recognizably human. But in addition to these three claims, the idea of adopting an emotion about an object concurrently with encountering the object is barely intelligible. How can instantaneous amusement be *about* the joke (or about the funniness of the joke) if it is not a response to it? For what it's worth, I can only imagine instantaneous amusement as very creepy and fake, as if it had been decided upon even before the joke was told and delivered nervously regardless of whether the joke is funny. I therefore suggest that amusement that comes too early fails to represent the comic value of the joke and is thereby *not fitting*. Since instantaneous amusement is not fitting, it is also not rational in response to an apparently funny joke.

Before I consider other emotions, note that if instantaneous emotions—emotions that occur at the same time as their object appears—are not only impossible and bizarre but flawed as evaluative responses, then a stronger version of the rational delay argument is available. Not only is it fitting (and therefore rational) for an emotional response to take time; it is *not* fitting (and therefore not rational) to experience certain emotions concurrently with encountering their object. So SSR does not only fail to allow for rational delay, it fails to preclude *irrational haste*. There is a fitting pace to the development of emotional reactions, and there is such a thing as experiencing an emotion *too soon*.

Now, let us leave amusement aside and consider fit-related reasons for emotions such as resentment, joy, or grief. Again, it is not merely that the time it takes us to 'take in' the reasons for such emotions seems rationally permissible, but that it is difficult even to imagine an instantaneous response. Hide the red scarf behind your back and I immediately cease to see it; pull it back out and it pops in my perceptual field. But resentment, joy, and grief do not pop in and out of existence like images in perceptual experience. The idea of leaping immediately to the depths of grief or to

the heights of joy, without ever *falling* or *ascending*, renders these emotions unrecognizable and makes it unlikely that they have the evaluative content that is supposed, on some accounts, to make them fitting to their objects. To call the time it takes us to respond emotionally to the world around us a period of 'rational delay' is already to concede too much to the state-oriented view of emotional normativity. In fact, the process of taking the facts in and responding emotionally to them is itself part of a fitting, ongoing response to the relevant evaluative facts.

Consider resentment. 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. It may not be fitting to forgive immediately after the wrong is done without fittingly resenting it. To be sure, whether resentment is ever necessary in this way is a substantive question, and those with stoic tendencies might reject this idea, but the fact there is a plausible substantive view of the ethics of forgiveness that appeals to fittingly evolving resentment suggests that we should make conceptual room for the phenomenon of irrational haste.

It is also worth noting how often we understand the fittingness of emotional states as essentially determined by their history. 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. 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 is often 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. If these fittingness norms are not synchronic, neither are their corresponding rational norms.

The first response to Podgorski's argument from rational delay is to bring in diachronic state norms, which govern the agent's doxastic attitudes at two or more individual times. Here, a similar move might seem available. The proposal might go like this. We need to determine the fitting diachronic relations between the agent's emotional states. Once we lay out the relevant

emotional states on the temporal axis, they can function as reference points that dictate the fitting movement of emotion. The trajectory of fitting amusement, for instance, can be captured by a diachronic state norm that breaks amusement down to many, consecutive emotional states that together make up the movement of amusement. Both rational delay and irrational haste can be addressed, it might seem, by proper specification of diachronic state norms.

The move to diachronic state norms does not work in the case of affective attitudes for roughly the same reasons it does not work in the case of doxastic attitudes. What is the appropriate time gap between having a reason to be amused and being amused, or between having a reason to regret and regretting? It might be possible to articulate baroque diachronic state norms that are extensionally equivalent to the relevant process norms for emotions. But such state norms are more plausibly viewed as corollaries of more fundamental process norms. Process norms more clearly capture the fitting movement of emotion even if the same movement can, in principle, be articulated by reference to emotional states.

Furthermore, a diachronic state view of emotional norms construes fitting emotional movement as merely derivative of fitting emotional states. It seems more plausible that fittingness norms for emotions directly govern the movement of the agent's emotion. Thus, if an agent is emotionally rational, then her emotional movement produces emotional states with certain regularity. Consider the fittingness of emotional transitions. 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 past movement. Thus, a state-oriented view of norms for emotions cannot allow for (or at the very least explain) the fittingness of emotional transitions. Consequently, a state-oriented view of *rational* norms for emotions cannot capture the *rationality* of emotional transitions.[13]

---

[13] Some of the points I make in this section are taken from Na'aman [forthcoming].

4. Can the Synchronist "Idealize" Emotions?

Podgorski concedes that his argument from rational delay would not carry much weight against a certain view of rational requirements as 'ideals' [Podgorski 2017: 17]. On this view, rational requirements can only be met by agents with cognitive capacities that are far beyond our reach as human beings [Christensen 2004, 2007; Broome 2013; Hedden 2016]. On the basis of such a view, Brian Hedden argues that process norms are merely derivative on more fundamental synchronic norms. 'If we were perfectly rational,' says Hedden, 'we wouldn't need to engage in reasoning in order to satisfy the requirements of rationality' [Hedden 2016: 8]. Reasoning is a tool we use to be better at satisfying the synchronic requirements of rationality. As such, reasoning has merely instrumental epistemic value.

Even if the claim that process norms are merely instrumental can get some traction where reasoning is considered, the claim is utterly implausible with regard to emotional processes. Though emotional transitions are often instrumental in moving us from one emotional state to another, and though sometimes our lingering emotions are remnants of reasons we no longer have, as we have already seen, the idea that we would have been *more* rational had we been able to leap immediately from sadness to joy, or from anger to contentment, renders these emotions unrecognizable.

Furthermore, it is arguable that the emotional stages we undergo in processes of grief or repentance, for instance, are not merely instrumental to these processes' emotional end-states but essential to the rationality of these end-states. Thus, the diminution of guilt in the aftermath of a serious wrong is not warranted unless it results from a process of moral repair that involves various emotional stages, as well as other attitudes and actions.

Similarly, the idea that we would have been more rational had we been able to move from the news of our beloved's death immediately to a state of acceptance—and that the tortuous path to acceptance is only instrumentally justified given our emotional inability to 'leap over' grief, as it were—is ethically dubious. Consider, for instance, the public outrage over the prospect of "medicalizing grief." Even if we could, with the help of medication, skip over the pain of grief without psychological repercussions, there would be a plausible (pro-tanto) objection to doing so. To be sure, in fittingly responding to the loss of a loved one we undergo a process that can lead to some psychological relief. But it seems plausible that the same psychological relief would be neither

fitting nor rational unless it emerged from the relevant process. Thus, the idea that ideally rational agents would move between rational emotional states without undergoing what we normally consider to be rational emotional processes is far less plausible than the idea that ideally rational agents would move between rational doxastic states without engaging in reasoning.

5. Conclusion

I have argued that the argument from rational delay to the existence of fundamental process-oriented norms is even more persuasive in the case of norms for affective attitudes than it is in the case of norms for doxastic attitudes. Whatever reason we have to believe there are fundamental state-oriented norms, our considered judgments about fitting emotions give us strong reason to accept the existence of fundamental process-oriented norms. This conclusion has important implications for contemporary discussions in the normativity and rationality of emotions, for it opens up a new way of understanding the instability of affect and the diminution of backward-looking emotions. But the conclusion also contributes to contemporary debates in epistemology by showing that weak process rationality is a highly plausible view.[14]

REFERENCES

Bittner, R. 1992. Is It Reasonable to Regret Things One Did?, *Journal of Philosophy* 89/5: 262–73.

Broome, J. 2013. *Rationality Through Reasoning*, Malden, MA: Wiley-Blackwell.

Carr, J. R. 2016. Don't Stop Believing, *Canadian Journal of Philosophy* 45/5: 744–66.

Chappell, R. Y. 2012. Fittingness: the Sole Normative Primitive, *The Philosophical Quarterly* 62/249: 684–704.

Christensen, D. 2004. *Putting Logic in Its Place: Formal Constraints on Rational Belief*, Oxford: Oxford University Press.

Christensen, D. 2007. Does Murphy's Law Apply in Epistemology? Self-Doubt and Rational Ideals, in *Oxford Studies in Epistemology* 2, ed. Tamar Szabo Gendler and John Hawthorne, Oxford: Oxford University Press: 3–31.

D'Arms, J., and Jacobson, D. 2000a. Sentiment and Value, *Ethics* 110/4: 722–48.

D'Arms, J., and Jacobson, D. 2000b. The Moralistic Fallacy: On the 'Appropriateness' of Emotions, *Philosophy and Phenomenological Research* 61/1: 65–90.

D'Arms, J., and Jacobson, D. 2009. Demystifying Sensibilities: Sentimental Values and the Instability of Affect, in *The Oxford Handbook of Philosophy of Emotion*, ed. Peter Goldie, Oxford: Oxford University Press: 585–616. Online publication 2010: DOI: 10.1093/oxfordhb/9780199235018.003.0027

D'Arms, J., and Jacobson, D. 2014. Wrong Kinds of Reason and the Opacity of Normative Force, in *Oxford Studies in Metaethics* 9, ed. Russ Shafer-Landau, Oxford: Oxford University Press: 215–44.

Harman, E. 2009. 'I'll Be Glad I Did It' Reasoning and the Significance of Future Desires, *Philosophical Perspectives* 23/1: 177–99.

Hedden, B. 2015. Time-Slice Rationality, *Mind* 124/494: 449–91.

Hedden, B. 2016. Mental Processes and Synchronicity, *Mind* 125/499: 873–88.

Hieronymi, P. 2001. Articulating Uncompromising Forgiveness, *Philosophy and Phenomenological Research* 62/3: 529–55.

Hlobil, U. 2015. There Are Diachronic Norms of Rationality, *Thought* 4/1: 38–45.

Howard, C. 2018. Fittingness, *Philosophy Compass* 13/11: 1–14.

Howard, C. 2019. The Fundamentality of Fit, in *Oxford Studies in Metaethics* 14, ed. Russ Shafer-Landau, Oxford: Oxford University Press: 216–36 .

Kenan, A.M. ms. Correct vs. Justified Emotions, a Disambiguation. In preparation.

Kiesewetter, B. 2017. *The Normativity of Rationality*, Oxford: Oxford University Press.

Lord, E. 2018. *The Importance of Being Rational*, Oxford: Oxford University Press.

Markovits, J. 2014. *Moral Reason*, Oxford: Oxford University Press.

Marušić, B. 2018. Do Reasons Expire? An Essay on Grief, *Philosophers' Imprint* 18/25: 1–21.

McHugh, C., and Way, J. 2016. Fittingness First, *Ethics* 126/3: 575–606.

McMahan, J. 2005. Preventing the Existence of People with Disabilities, in *Quality of Life and Human Difference*, ed. David Wasserman, Jerome Bickenbach, and Robert Wachbroit, Cambridge: Cambridge University Press: 142–71.

Millgram, E. 1999. Moral Values and Secondary Qualities, *American Philosophical Quarterly* 36/3: 253–55.

Moller, D. 2017. Love and the Rationality of Grief, in *The Oxford Handbook of Philosophy of Love*, ed. Christopher Grau and Aaron Smuts, Online publication date: April 2017. URL= https://doi.org/10.1093/oxfordhb/9780199395729.013.35

Moss, S. 2015. Time-Slice Epistemology and Action Under Indeterminacy, in *Oxford Studies in Epistemology*, ed. Tamar Szabó Gendler and John Hawthorne, Oxford: Oxford University Press: 172–94

Na'aman, O. 2020. The Fitting Resolution of Anger, *Philosophical Studies* 177: 2417–30.

Na'aman, O. forthcoming. The Rationality of Emotional Change: Toward a Process View, *Noûs*.

Parfit, D. 1984. *Reasons and Persons*, Oxford: Oxford University Press.

Podgorski, A. 2017. Rational Delay, *Philosophers' Imprint* 17/5: 1–19.

Rosen, G. 2015. The Alethic Conception of Moral Responsibility, in *The Nature of Moral Responsibility: New essays*, ed. Randolph Clarke, Michael McKenna, and Angela M. Smith, Oxford: Oxford University Press: 65–88.

Schroeder, M. 2004. The Scope of Instrumental Reason, *Philosophical Perspectives* 18/1: 337–64.

Schroeder, M. 2007. *Slaves of the Passions*, Oxford: Oxford University Press.

Schroeder, M. 2009. Means-End Coherence, Stringency, and Subjective Reasons, *Philosophical Studies* 143: 223–48.

Setiya, K. 2004. Against Internalism, *Noûs* 38/2: 266–98.

Setiya, K. 2014. The Ethics of Existence, *Philosophical Perspectives* 28/1: 291–301.

Sharadin, N. 2015. Reasons Wrong and Right, *Pacific Philosophical Quarterly* 97/3: 371–99.

Sylvan, K. 2015. What Apparent Reasons Appear to Be, *Philosophical Studies* 172: 587–606.

Titelbaum, M. 2015. Continuing On, *Canadian Journal of Philosophy* 45/5–6: 670–91.

Wallace, R. J. 2013. *The View from Here: On Affirmation, Attachment, and the Limits of Regret*, Oxford: Oxford University Press.

Wallace, R. J. 2019. Trust, Anger, Resentment, Forgiveness: On Blame and Its Reasons, *European Journal of Philosophy* 27/3: 537–51.

Way, J. 2009. Two Accounts of the Normativity of Rationality, *Journal of Ethics and Social Philosophy* 4/1: 1–8.

Way, J. 2012. Explaining the Instrumental Principle, *Australasian Journal of Philosophy* 90/3: 487–506.

Way, J. 2012. Transmission and the Wrong Kind of Reason, *Ethics* 122/3: 489–515.

Whiting, D. 2014. Keep Things In Perspective: Reasons, Rationality and the A Priori, *Journal of Ethics and Social Philosophy* 8/1: 1–22.

Oded Na'aman is a Postdoctoral Fellow at the Martin Buber Society of Fellows for the Humanities and Social Sciences at the Hebrew University of Jerusalem.