Do emotions represent values and how can we tell?

Alex Grzankowski
Birkbeck, University of London
Email: alex.grzankowski@gmail.com

Abstract

Do emotions represent values? The dominant "content view" has it that they do. But there is a newcomer on scene: the attitude view. According to it, rather than representing value properties, there is a value-relevant *way* you represent the targets of emotion. For example, in feeling angry with someone, you stand to them in the relation of representing-as-having-wronged-you. But the central considerations in favour of these competing views are left wanting and it is hard to see how to choose amongst these alternatives. I argue that there is an empirical path to a decision.

KEYWORDS

motion, value and values, representational theory of mind, cognitivism, attitudinalism about emotion

1. INTRODUCTION

Emotions represent values. Or so goes the dominant philosophical conception.[1] Emotions "tell" us about the world, and they can be evaluated as correct or incorrect and, crucially, what they tell us about the world concerns value. According to some philosophers, emotions are an essential first point of contact with value properties,[2] *revealing* values to us. According to the representational view, in feeling angry with the person who just stepped on your toe, you represent them as having wronged you.

But there is an emerging alternative view to this dominant position. Rather than representing them as having wronged you, there is a *way* you represent them in feeling angry with them: you represent-as-having-wronged-you.[3] The person in question is *what* is represented and there is a manner of representing—a stance—that you take towards them. Let us call this newcomer view the "attitude view" [A9] and the more familiar representational view the "content view". The present essay considers whether there is a deep difference between these views and whether there is anything that might allow to choose between them, for they might look to be mere notational variants. I will argue that the extant considerations do not yet give one view a clear upper hand over the other but that there is indeed an interesting and substantive difference between the two views that is, perhaps surprisingly, empirically tractable. I do not know which of these views we ought to prefer but we will be left with a method for deciding that takes us further than the armchair considerations one presently finds in the literature (and which do not settle the matter).

2. THE DISPUTE

Everyone in the present dispute agrees that the emotions have something to do with value properties (and I will not question that assumption here). What disputants disagree about is how emotions are *related* to values. The disagreement is one over the metaphysics of the emotions. On the attitude view, values are *modifications* of representation. On the content view, values are *in the content* of emotions.

To better understand the attitude view, consider the act of kicking a ball. There are many ways one might kick the ball: gently, slowly, quietly, softly. Although it is strained to say that kicking is "directed at" the ball, the ball is the object of one's kicking—it is the thing kicked—and there are various *ways* one might *do* or *perform* the kicking. Similarly, on the

attitude view, there is an object of one's emotion (let us say a growling dog) and then there are ways one can represent it. In fear, we might say that one represents-in-the-danger-way the dog. Crucially, this is meant to be distinct from representing the dog as being dangerous where, on that view, dangerousness is not a *modification of an attitude*. Rather, it is part of the content of an attitude.

But is this not just pushing puzzle pieces around? What exactly is the cash-value difference between saying that value properties are ways of representing rather than ways things are represented to be? We have two views before us. According to one view, value properties are "in the content" and according to the other view they are not. Is this a difference that makes a difference? Those engaged in the debate certainly think so. But why prefer saying one of these things over the other? In the next two sections I will look at what have emerged as the key points that disputants often highlight in favour of their own view. I have no doubt that there are (or could be) further arguments that might be brought to bear for and against these views, but it is striking that these often-appealed-to, key considerations actually carry very little weight when properly understood. So, my first goal is to show that the extant central arguments still leave us in a position of indecision. Seeing why these arguments do not work is worthwhile for its own sake but it will also serve as motivation to seek a new way ahead. My second goal will then be to outline that path.

3. IN FAVOUR OF THE ATTITUDE VIEW?

Think about the role of value properties such as dangerousness and offensiveness in relation to the emotions. It's common to take these properties to provide the 'formal objects' of the emotions.[4] But other attitudes (non-emotions) have formal objects as well. We might argue over exactly which properties go with which attitudes, but to get the idea,

consider *truth* and *the good*. Beliefs *aim* at the truth, we might say. Desires aim for goodness (or perhaps good-for-me-ness). Let's not get distracted over the finer details since the main points I wish to make do not turn on these choices. The point presently is that there are normative aims or standards that apply more widely than just to the emotions. But now focus on the connection between these properties and the *contents* of belief and desire. Desire is the clearer case, so let's start there and then I'll have a bit more to say about belief.

Suppose you desire that you have a big piece cake. If you get a big piece of cake, your desire will be *satisfied*. Desires have satisfaction conditions and when those conditions are met, your desires are satisfied. We can capture this in terms of truth. Let your desire be a relation to a proposition: the proposition that you have big piece of cake. When that proposition is true, your desire is satisfied and when it is false it is frustrated. That is a pretty typical way of thinking about desire and satisfaction.[5] But notice that we have not yet mentioned *goodness* at all. And yet it is exceedingly plausible that in wanting to have a big piece of cake you 'see the good in it'.[6] But I do not think we want to say that your desire is *satisfied* just in case it is good for you to have the cake. Rather, your desire is satisfied if you *get* the cake, whether that is a good thing or not. As a matter of fact, there may be little or no good in having the cake. A big piece is greedy, it is overly calorific, it is not especially nutritious. So, you want the cake and you will be pleased if you get it, but, as it turns out, this is not a situation that "participates in the good". So whatever goodness is doing in relation to desire, it seems to be different from setting its satisfaction conditions. But satisfaction conditions are captured in terms of truth and that is why it is common to hold that (at least some) desires are propositional attitudes. So, we had better find a different home for goodness—it does not belong *in the propositional content* of the desire qua

satisfaction conditions. So, the formal object of desire is not a content feature. Or so we might think.

In my view, the above point about goodness and satisfaction can be made all on its own, but some theorists have reached it via considerations about belief, perception, and truth and this has now become a common form of argument in the existing papers on the attitude view:

> According to what I'll call "the attitude version", "good" is a feature of the form of the attitude, not its content. This view takes the relation between, on the one hand, desiring, intending, or acting intentionally, and, on the other hand, good, to be analogous to the relation between believing and true. In having a belief that p one takes p to be true even though 'is true' is not (or at least does not need to be) part of the content of one's belief that p. In the same way, in intending X, desiring X, or doing X intentionally, one holds X to be good even though 'good' is not (necessarily) part of the content of these attitudes. (Tenenbaum, 2018, p. 14)

Tenenbaum is arguing for an attitude view of desire in the face of wide adoption by others of a content view. Guise of the good theories of desire hold that in desiring something, one represents it is as good—perhaps by judging it to be good or perceiving it to be good. Tenenbaum's point is that to the extent that truth and goodness play analogous roles in relation to belief and desire respectively, we would be making a mistake by taking goodness to be part of the content of desire. We do not say that believing that p is believing that p is true,[7] so we should not say that in desiring that p one desires that p is/be good. So, again, the formal object of desire (and belief) is not a content feature.[8]

Theorists arguing for the attitude view of emotions seem to have something similar in mind:

> According to the attitudinalist, by contrast, the evaluation in emotion should not be understood in representational terms. The relation between the emotion and the evaluative property is claimed to be similar to the relation between, say, belief and truth. Belief does not as such represent the truth of the proposition believed. (Deonna & Teroni, 2024, pp. 47)

So what might we take away from the above? We seem to have a class of properties—formal objects—that connect to when an attitude is appropriate or fitting. Beliefs are fitting when they are true, desires when they are for the good, and emotions when they are directed at things that possess the relevant "thick" values. But in the case of belief and desire, we are not (or at least should not be) tempted to say that this is explained in terms of the contents of beliefs and desires and so we should not say such things when theorising about the emotions either. In short, the argument tells us that we must keep our house in order. Doing so brings the emotions into alignment with the best theorising about other attitudes such as belief and desire.

Although at first compelling, this argument is too quick. I agree that in believing that *p* one need not believe that *p* is true and in desiring that *q* one need not desire that *q* is/be good. But one cannot so easily now move to the conclusion that value properties are not represented properties. An important difference (the difference for example between satisfaction and the formal object of desire) has been brought out, but so far we have not been given a reason to say that value properties are not represented *anywhere*. More precisely, the representationalist has space to develop a view according to which what it is to believe, desire, or emote is to

represent things as being a certain way without holding that this is a mere add-on to what we might call the "surface content". And, in fact, when one takes a closer look at the standard representational theory of emotions and the most sophisticated guise of the good theories of desire, one can see that the above considerations about content simply do not speak against them. [9]

Start with emotion. In fearing the dog, on the content view, one represents it as dangerous. On the "surface" this simply looks to be a mental state *about* the dog—its content is simply the object, that very dog. But where is *dangerousness*? Notice that the theory does *not* say that in fearing the dog one *fears that* the dog is dangerous. Rather, in being afraid, the view says you evaluate the dog as being dangerous. The two most well-known variants of this idea say that the representing is judging or perceiving the dog to be dangerous. We might represent the idea as follows (where '< >' designates a proposition and the relations designated by "F" and "R" are stipulated to be distinct, as are "D" and "R"):

1. S fears o: F (s, o)
2. S represents o as dangerous: R (s, <o, dangerousness>)

And 2 is different from:

3. S fears that o is dangerous: F (s, <o, dangerousness>)

The content view analyses fearing an object in terms of a distinct relation: 1 is analysed as 2. Specifying the nature of that relation is not important for the points I wish to make so long as it is not the fearing relation.[10]

And the same considerations apply in the case of desire:[11]

4. S desires that S have a big piece of cake: D (s, <Ex (Cx, Bx, H(s,x))>)

5. S represents having a big piece of cake as good: R (s, <<Ex (Cx, Bx, H(s,x))>, goodness>)
6. S desires that S having a big piece of cake is/be good: D (s, <<Ex (Cx, Bx, H(s,x)>, goodness>)

The guise of the good theorist about desire does not simply add goodness to the content of desire as in 6. To do so would be to say that one has a desire that is satisfied just in case having a big piece of cake is good and we saw reasons to resist that view above. Rather, desiring is analysed as representing as good (as in 5). To desire that one have a big piece of cake is to represent that one's having a big piece of cake is good. In effect, the desiring relation is unpacked into two components: a new relation, call it R, and a representation of the original content of the desire *as good*. As above, we can go onto have a further argument about the nature of R – perhaps it is *sui generis*, or perhaps it can be accounted for in more familiar terms such as perceiving or judging.[12] What's important for us is that there is a further content (as in 5) beyond the "surface content" (as in 4) disclosed upon analysis and a further relation as well. The satisfaction condition of the desire is given in 4 by the surface content, but "under the hood" there is more to the story, and this is precisely where the value properties get into the picture.[13]

Notice that on the guise of the good theory of desire as laid out here and the account of fearing in terms of representing as dangerous, the value properties *are indeed parts of contents*. They are not part of *what one wants* or *what one fears*, but they get into content one layer down, in the philosophical analysis. They are still *represented* properties. So, when one notes that the formal object of fear and desire do not look to be part of content, we must not be too hasty. It is important to locate them in the right place and in the right way, but on the views we are considering now, value properties are still represented. And for fans of the view, this can look to be

advantageous for we now have a clear (or at least familiar) story about how attitudes relate us to values—in being in various states we *represent* those values. All this while avoiding the charge of conflating formal objects with what is, say feared or desired. Formal objects are represented, but not in the surface content.

So, one important upshot is that the observations about satisfaction and goodness and the observations about truth and belief simply do not get one away from a content view. They only speak against something like a surface content view. Those observations *do* show that the way truth features in belief cannot simply be by being tacked onto what one believes and they *do* show that the normative standards of desire cannot be equated with their satisfaction conditions, but they do *not* yet show that truth and the good are not relevant to belief and desire *due to being represented*. And fundamentally that's what a content theorist wants—beliefs, desires, and emotions are answerable to value properties such as truth, goodness, dangerousness, offensiveness, and so on because they involve representations of those properties.

With the above in mind, things can now really start to look like mere notational variation. Compare the following:

A: Fearing the dog: Represents-as-dangerous (s, dog)
C: Fearing the dog: R (s, <dog, dangerousness>)

What's *really* the important difference here? Both theorists take the target to be fearing the dog. One theorist says that "under the hood" one finds an instance of representing and specifically a representing of the dog and an attribution of dangerousness. The other theorist also says there is a representing of the dog but rather than an attribution of dangerousness to the dog there is a manner or way of representing the dog—the as-dangerous way. C might look like quite an attractive way to unpack A, in

fact. In order for the putative difference between A and C to really matter, then, the attitude theorist needs to say that value properties simply *are not represented properties*, not on the surface, not under the hood, not anywhere. Attitudes such as belief, desire, and the emotions are answerable to patterns of instantiation of value properties out in the world, but not because they in *any way* represent those properties. The relation must be explained in other terms. Although this is an interesting thesis, the arguments about formal objects above do not yet establish it.

4. IN FAVOUR OF THE CONTENT VIEW?

It is clear that both the attitude theorist and the content theorist think that our attitudes answer to various value properties. For example, fear is fitting or appropriate when the feared thing is dangerous and a desire is fitting or appropriate when the desired thing participates in the good. And, as noted, the content theorist says that these facts are to be explained in representational terms. This gives way to one of the most powerful arguments in favour of classical representationalism – that is, the content view.

There is something incorrect or irrational about loving a dog that is mean, dangerous, that tears up the house, and threatens your children. A dog that is loyal, playful, and gentle, on the other hand, is a perfectly fitting object of love. So, we can ask, in loving something, is one's love fitting or correct? And is not this now enough to establish that one represents the dog *as loveable*? True, loving the dog is not the same thing as loving that the dog is lovable, but if in loving the dog one can 'get things right' or can 'get things wrong', must not one somewhere and somehow be taking things to be some way, namely as lovable? And is not this sufficient to establish that things are represented as being some way? This looks to be a quick route to the content view.

But it is a substantive claim that fittingness or appropriateness is to be understood in representational terms. We *could* unpack appropriateness in terms of representing that the dog is lovable, but this needs to be argued for. This is shown by considering other things that can be appropriate or inappropriate but are not representational.[14] It is appropriate: to bring an umbrella when it is raining, to genuflect as a Catholic in front of the alter, to eat with a salad fork during the salad serving, to tell the truth, to react with disgust when smelling rotten egg. There is plenty one might wish to tease apart here, but it's difficult to see why one would think that, in all of these cases of appropriate action, one must be representing things as instantiating value properties. Just to focus on one example, the action of bringing an umbrella on a rainy day does not *represent* prudence – that is, bringing an umbrella does not entail that one represents anything as being prudent. The activity exemplifies or displays prudence. So those who hold that emotions are evaluable for appropriateness in virtue of representing value properties must substantiate this commitment.[15] Moreover, perhaps the attitude theorist can appeal to whatever we end up saying about the appropriateness of actions when trying to make sense of the appropriateness of emotion and other attitudes. In short, it is too quick to move from fittingness directly to representation.

So, the "master argument" for the content view is too fast and so, as with the attitude theory above, we are left in a position of indecision. There is no quick argument from evaluation to representation.

I have no doubt that theorists engaged in this debate will develop further arguments to prefer their own view over their competitor's. My goal in the present paper is not to canvass all known considerations (though at the moment the literature really does seem to me to be in a stalemate even when looking beyond the arguments just considered). But I hope I have shown that some of the central arguments are not as powerful as one might have hoped. Moreover, I have gone some way towards

motivating the idea that we could use a new plan of action in this domain. Fortunately, I do think there is a promising way ahead.

5. SUBSTANTIVE COMMITTMENTS

The core difference between the attitude view and the content view is that the attitude view holds that value properties are not represented *anywhere* but are, instead, ways of representing. In contrast, the content view commits to the idea that value properties are represented, even if only under the hood. I hope the above sections have sufficiently motivated that way of seeing the light of day between the views. But to claim that there is a representation of something present in your mental economy is a substantive commitment and indeed one that, at least in principle, we can explore empirically. This is how we will make headway on the dispute between the content theorist and the attitudinalist.

One worry about representing value properties is that one might think there is no naturalisable way to make sense of how such properties could be represented.[16] One might worry, for example, that such properties simply do not exist (on moral irrealist grounds) or one might worry that they are not causal properties and so could not be made to fit with a naturalised theory of the attitudes. I do not wish to press these concerns here although they do seem serious if one is committed to naturalising content. In short, my focus is not about *whether* value properties could be represented or *how* we could come to represent them. Rather, I wish to simply grant for the sake of argument that these questions can be answered. My focus is on the presence or absence of a representation in a mental economy.

To posit a representation is a substantive matter. What exactly is required for some entity to be a representation or for some state or event to be representational is a well-worn and controversial issue, but for our

purposes a comparison between two cases is enough to make the desired points. Consider the gear in a watch that controls the minute hand. As the gear turns, the minute hand turns and, moreover, the gear was *designed* to co-vary in position with the minute hand. It's possible for the gear to come loose and for it to spin freely without the minute hand moving and so to be faulty. And yet, despite meeting many of the classical markers of being a representation (co-varying, being designed to do so, and possibly malfunctioning), it is intuitively implausible that the location of the gear *represents* the location of the minute hand. No doubt, it's easy to read one fact off of the other when things are going well in the watch, but being able to be easily transform something into a representation is not the same as being a representation. A bit of mercury sitting in a puddle on a rock does not represent temperature even though it easily could be recruited to do so. So being a co-varying thing is not enough to be a representation and even being a co-varying thing that was designed to co-vary is not yet a representation. But now compare the gear and minute hand example with another:

> Foraging desert ants, Cataglyphis fortis, return to their nest by keeping a running total of their distance and direction from the nest. This mode of navigation was called path integration by Mittelstaedt and Mittelstaedt (Mittelstaedt and Mittelstaedt, 1982), who provided the first systematic studies of this phenomenon, and vector navigation by Wehner (Wehner, 1982; Wehner, 1983). More recent reviews and considerations on path integration are given elsewhere (Wehner and Srinivasan, 2003; Merkle et al., 2006). By path integration the ants acquire a home vector that enables them to return at any time along the beeline, so to speak, to the nest. However, after having played out their home vector, they do not always arrive exactly at the entrance of their nest, a tiny hole in the desert floor …

> If an ant fails to find the nest after having "run off" its home vector, it terminates its almost straight inbound run and starts a systematic search for the nest (Wehner and Srinivasan, 1981). During this search C. fortis performs loops of increasing radius around the supposed nest position (Wehner and Wehner, 1986). (Merkle *et al,* 2006, p.3545)

When all goes well, the ant finds food and heads straight home. But if, by mother nature or by experimental design, the ant is moved off track, the ant arrives at what should be home, but home is not found. The ant then initiates a search by looping wider and wider.

Focus on the ant's relation to the nest. When the nest is not found, something must explain not only why the ant begins its systematic looping search (and it cannot be the nest itself since the nest is not where the ant is) but also why that looping ceases when the nest is found. A very attractive explanation is that the ant has a something like a placeholder, "home', and, using its running total of distance and direction, when it fails to arrive at home, that placeholder is utilised. There is some way of checking "am I home?", a time when that checking is to occur, and then some way of initiating a search when the answer is "no!". Finally, the ant ceases the search only when the answer is "yes, I'm home" (and not for example when finding a white pebble or a leaf).

As others have argued, in addition to being designed to co-vary with other things, representations serve as stand-ins in situations of absence[17] and they serve to guide further processing and behaviour. This is exactly what we find with the ant and what differentiates that case from the gears of the watch. Representations are designed to carry information that can be utilised by a system and that is available when the worldly correlates are missing.[18]

The central commitments of positing representations are brought out nicely in a recent paper by Newen and Vosgerau (2020):

> Mental representations are a means to explain behavior. This, at least, is the idea on which cognitive (behavioral) science is built: that there are certain kinds of behavior, namely minimally flexible behavior, which cannot be explained by appealing to stimulus-response patterns. Flexible behavior is understood as behavior that can differ even in response to one and the same type of stimulus or that can be elicited without the relevant stimulus being present. Since this implies that there is no simple one-to-one relation between stimulus and behavior, flexible behavior is not explainable by simple stimulus-response patterns. Thus, some inner processes of the behaving system (of a minimal complexity) are assumed to have an influence on what kind of behavior is selected given a specific stimulus. These inner processes (or states) are then taken to stand for something else (features, properties, objects, etc.) and are hence called "mental representations". (Newen & Vosgerau, 2020 , pp. 178)

What does all of this have to do with the emotions? Recall that the content theory of emotion holds that value properties are represented, but we can now see that this is a substantive claim. In cognitive behavioural science it comes to the view that there are stand-ins for value properties that can be utilised by a representational system.[19] That is, the system's basic cognitive architecture and operations are, if the content theory is right, best specified in terms of internal representational states that are manipulated via computational algorithms and some of those algorithms manipulate representations of value. Moreover, the representations must be available for use by different algorithms and in the absence of any specific sensory

inputs or behavioural outputs (otherwise we might simply assimilate them to non-representational stimulus/response patterns). Those are predictions made by the content view of emotions. We can contrast this with attitude states, which are understood as *computational dispositions* or *operations* over internal representations. Such operations are not stored in a format that makes them available for downstream algorithmic use. It is this distinction that we can use to pull apart the content view and the attitude view. And indeed this is exactly the sort of project that cognitive science is in the business of. Theorists aim to uncover the representations and the algorithms that use them by devising experiments that tease apart action, activity, and changes in mental states in fine grained ways. We can form hypotheses concerning representations and then seek data that are best explained (or not) by the existence of the hypothesised representations and algorithms. The foraging behaviour of ants above provides one such case: positing a representation for "home" that is utilised by the system even in the absence of the home provides an explanation of the observed behaviour and this hypothesis can be further refined and tested. The ultimate goal is an account of what is represented and how behaviour is guided by computations and, moreover, an account that is well supported by ruling out competing hypotheses and showing how the account predicts the observed behaviour.

I am under no illusion – there are no *proofs* in the offing here. But, as with any scientific inquiry, we are in a position to offer sensible hypotheses and to stress test them. Devising experiments to test whether emotions represent value properties will not be a simple task. But my point from a philosophical perspective is not hindered. To hold that value properties are represented gives way to a known method of inquiry. We should, in principle, be able to develop tasks and tests that isolate the presence of information carrying stand-ins that are utilised (or utilisable) elsewhere in a representational system. In other words, representationalists

have posited the presence of a certain kind of entity that plays an explanatory role. I do not think we should shy away from arm-chair arguments for or against the content view or the attitude view – those are important too – but in this instance, we seem to have an empirical method to help break what I fear might be an emerging stalemate. If the content theorists could show that there is powerful empirical evidence in favour their view, this would be a major advance. If this cannot be done, we will be given powerful reason for looking for an alternative story, and the attitude theorist will then be in good position to come to the rescue.[20]

Before closing this section, it is worth warding off a worry. One might worry that a cognitive scientific notion of representation is not obviously what emotion theorists are concerned with. They might be interested in "representation as" based on rationalising explanation or a description of phenomenology.[21] But I do not intend the foregoing to be in conflict with these aims. It seems to me that all parties in the present debate can agree about the phenomenal and rational data points. What they disagree about is whether that data can and should be captured in representational terms. The emergence of the attitude view – a view that does *not* aim to *avoid* discussion of rationality or phenomenology – raises to salience that one cannot simply assume that the data is to be captured in terms of content. What the methodology from cognitive science aims to do is provide a way of teasing apart how best to capture the agreed upon data and to decide which path to follow.[22]

6. CONCLUSION

For forty years the dominant view of emotions amongst philosophers has been a content-oriented representational theory. It is exciting that there is a new-comer view that aims capture many of the advantages of this well known approach but in fresh terms. From a certain vantage, the differences between the attitude view of emotions and the content view are subtle (and

perhaps small) and the arguments in favour of one view over the other have not yet shown a clear winner. But on closer inspection, the content view makes a testable prediction that stands in contrast to the attitude view and following this path is a promising way to decide between these competitors. Perhaps with the help of specialists in experimental design this path can be taken.

ACKNOWLEDGEMENTS

Thanks are due to audiences at the University of Bristol, the Université Cote d'Azur, Nice, Cardiff University, and to the members of the Nature of Objects of Thought and Assertion Network. Thank you also to the editors and anonymous referees of Mind and Language for comments and helpful feedback.

REFERENCES

Ballard, B. (2021). Content and the fittingness of emotion. *Philosophical Quarterly* 71 (4),[A20] pqaa074, https://doi.org/10.1093/pq/pqaa074

Block, N. (1986). Advertisement for a semantics for psychology. *Midwest Studies in Philosophy* (Studies in the Philosophy of Mind), 10: 615–678.

Block, N. (2023). [A21] *The Border Between Seeing and Thinking*. Oxford University Press, New York.

Braun, D. (2015). Desiring, desires, and desire ascriptions. *Philosophical Studies* 172 (1):141-162.

Chappell, R. Yetter (2012). Fittingness: The sole normative primitive. *Philosophical Quarterly* 62 (249):684-704.

Deonna, J. & Teroni, F. (2024). Emotions and their correctness conditions: A defense of attitudinalism. *Erkenntnis* **89**, 45–64. https://doi.org/10.1007/s10670-022-00522-0

Deonna, J. & Teroni, F. (2015). 'Emotions as attitudes'. Dialectica 69 (3):293-311.

Deonna, J. & Teroni, F. (2014). In what sense are emotions evaluations? *Emotion and Value*, 15–31.

Deonna, J. & Teroni, F. (2012). *The Emotions: A Philosophical Introduction*. Routledge.[A22]

Elgin, C. Z. (2008). Emotion and understanding. In G. Brun, U. Dogluoglu & D. Kuenzle (eds.), *Epistemology and Emotions*.

Fara, D. G. (2013). Specifying desires. *Noûs* 47 (2):250-272.

Federico L. (eds.), *The Nature of Desire*. Oxford University Press. pp. 201-217.

Gregory, A. (2021). *Desire as Belief: A Study of Desire, Motivation, and Rationality*. Oxford: Oxford University Press.

Greenspan, P. (1988). *Emotions and Reasons: an Inquiry into Emotional Justification*. New York: Routledge, Chapman and Hall.

Grzankowski, A. (2019). To believe is not to believe true: Reply to Sankey. *Principia: An International Journal of Epistemology* (1):137-138.

Grzankowski, A. (2021). A puzzle for evaluation theories of desire. *Thought: A Journal of Philosophy* 10 (2):90-98.

Grzankowski, A. & Montague, Michelle (eds.) (2018). *Non-Propositional Intentionality*. Oxford University Press.
Harman, G. (1993). Meaning holism defended. in Fodor and Lepore (eds.), *Holism: A Consumers Update*, Amsterdam: Rodopi. 1993, pp. 163–171.

Helm, B. (2001). *Emotional Reason: Deliberation, Motivation, and the Nature of Value*. Cambridge: Cambridge University Press.

Helm, B. (2015). Emotions and recalcitrance: Reevaluating the perceptual model'. *Dialectica* 69.3, pp. 417–33.

Helm, B. (2024). Cognitivist theories of emotions. In A. Scarantino (ed.). *Emotion Theory: The Routledge Comprehensive Guide: Volume I: History, Contemporary Theories, and Key Elelments (1st ed.)*. Routledge. https://doi.org/10.4324/9781315559940, Ch. 19.

Howard, C. (2018). Fittingness. *Philosophy Compass* 13 (11):e12542.

Johnston, M. (2001). The authority of affect. *Philosophy and Phenomenological Research.* 63, 181-214.

Kriegel, U. (2018). *Brentano's Philosophical System: Mind, Being, Value*. Oxford: Oxford University Press.

Kriegel, U. (2018). Belief-that and belief-in: Which reductive analysis? In Alex Gzrankowski & Michelle Montague (eds.), Non-Propositional Intentionality. Oxford: Oxford University Press. pp. 192-213.

Kriegel, U. (2019). The intentional structure of moods. *Philosophers' Imprint*, 19:1-19.

Mendelovici, A. (2018). Propositionalism without propositions, objectualism without objects. In Alex Grzankowski & Michelle Montague (eds.), *Non-Propositional Intentionality*. Oxford: Oxford University press pp. 214-233.

Merkle, T., Markus K. and Rüdiger W. (2006). Uncertainty about nest position influences systematic search strategies in desert ants. *The Journal of Experimental Biology* 209, 3545-3549.

Merricks, T. (2009). Propositional attitudes? *Proceedings of the Aristotelian Society*. 109, Issue 1 pt 3, 1 October, pp. 207–232. https://doi.org/10.1111/j.1467-9264.2009.00265.x

Mitchell, J. (2017). The epistemology of emotional experience. Dialectica 71(1): 57–84.

Mitchell, J. (2021). *Emotion as Feeling Towards Value*. New York: Oxford University Press.

Mittelstaedt H. and Mittelstaedt M.L. (1982). Homing by path integration. In: Papi F. and Wallraff H.G. (Eds.). *Avian Navigation*. Heidelberg (Germany): Springer. p. 290–297.

Neu, J. (2000). *A Tear is an Intellectual Thing: The Meaning of Emotions*. Oxford University Press, New York.

Newen A. and Gottfried Vosgerau, G. (2020). Situated mental representations: Why we need mental representations and how we should understand them. in J. Smortchkova, T. Schlicht, & K. Dolega (eds). *What Are Mental Representations?* Oxford University Press. pp.178-212

Nussbaum, M. (2001). *Upheavals of Thought: The Intelligence of Emotions*. New York: Cambridge University Press.

Oddie, G. (2005). *Value, Reality, and Desire*. Oxford: Oxford University Press.

Orlandi, N. (2020) 'Representing as coordinating with absence', in J. Smortchkova, T. Schlicht, & K. Dolega (eds). *What Are Mental Representations?* Oxford University Press. Pp. 101-134.

Poellner, P. (2016). Phenomenology and the perceptual model of emotion. Proceedings of the Aristotelian Society 116(3): 261-288.

Quilty-Dunn, J. and Mandelbaum, E. (2018). Against dispositionalism: belief in cognitive science. *Philosophical Studies.* 175 (9):2353-2372.

Roberts, R. (1988) What an Emotion is: A sketch. *Philosophical Review* 97(2): 183-209.

Rossi, M. & Tappolet, C. (2019). What kind of evaluative states are emotions? The attitudinal theory vs. the perceptual theory of emotions. *Canadian Journal of Philosophy* 49 (4):544-563.

Sankey, H. (2019a). To believe is to believe true. *Principia: An International Journal of Epistemology* 23 (1):131-136.

Sankey, H. (2019b). Neither a truism nor a triviality: Reply to Grzankowski. *Principia: An International Journal of Epistemology* 23 (2):361-365.

Schafer, K. (2013). Perception and the rational force of desire. *Journal of Philosophy* 110. 258-281.

Schneegans, S. & Bays, P. M. (2017). Neural architecture for feature binding in visual working memory. *Journal of Neuroscience*, 37(14), 3913–3925. https://doi.org/10.1523/ jneurosci.3493–16.2017

Schroeter, L., Schroeter, F. & Jones, K. (2015). Do emotions represent values? *Dialectica* 69 (3):357-380.

Smortchkova, J., T. Schlicht & K. Dolega (2020). *What Are Mental Representations?* Oxford University Press.

Solomon, R. (1976). *The Passions*. New York: Anchor Press/Doubleday.

Stampe, D. W. (1987): The authority of desire. *Philosophical Review* 96. 335-381.

Tappolet, C. (2012). Emotions, perceptions, and emotional illusions. In Calabi Clotilde (ed.), *Perceptual Illusions*. Philosophical and Psychological Essays. Palgrave-Macmillan. pp. 207-24.

Tappolet, C. (2016). Emotions, Value, and Agency. Oxford University Press.

Tenenbaum, S. (2007). *Appearances of the Good: An Essay on the Nature of Practical Reason*. Cambridge: Cambridge University Press

Tenenbaum, S. (2018). The Guise of the Guise of the Bad. Ethical Theory and Moral Practice 21 (1):5-20.

Teroni, F. (2007). Emotions and Formal Objects. *Dialectica*, LXI, 3: 395–415.

Velleman, D. (1992). The guise of the good. *Noûs* 26(1):3–26.

Wehner, R. (1982). Himmelsnavigation bei insekten. Neurophysiologie und verhalten. Neujahrsbl. Naturforsch. Ges. Zürich 1984, 1-132.

Wehner, R. (1983). Celestial and terrestrial navigation: human strategies – insect strategies. In Neuroethology and Behavioral Physiology (ed. F. Huber and H. Markl), pp. 366-381. Berlin, Heidelberg: Springer.

Wehner, R. and Srinivasan, M. V. (1981). Searching behavior of desert ants, genus Cataglyphis (Formicidae, Hymenoptera). J. Comp. Physiol. 142, 315- 338. https://doi.org/10.1007/BF00605445

Wehner, R. and Srinivasan M. V. (2003). Path integration in insects, in K.J. Jeffery (ed.), *The Neurobiology of Spatial Behaviour*. https://doi.org/10.1093/acprof:oso/9780198515241.003.0001

Wehner, R. and Wehner, S. (1986). Path integration in desert ants. Approaching a long-standing puzzle in insect navigation. Monitore Zoologico Italiano - Italian Journal of Zoology, 20(3), 309–331. https://doi.org/10.1080/00269786.1986.10736505

Wilson, J. R. S. (1972). *Emotion and Object.* Cambridge, UK: Cambridge University Press.

[1] Content views are widespread and take many different forms. See Brady (2007, 2009), Damasio (1994), Döring (2009), Goldie (2000), Greenspan (1981, 1988), Helm (2001), Mitchell (2017), Neu (2000), Nussbaum (2001), Poellner (2016), Prinz (2004), Roberts (1988, 2003), Solomon (1976), Tappolet (2012, 2016), and Tye (2008). See Helm (2024) for a useful overview.

[2] See Elgin (1996) and Johnston (2001).

[3] In the literature on emotions the view has been developed and defended in most detail by Deonna and Teroni (2012, 2015, 2024). A structurally similar view that will be discussed below has been offered for desire. See Schafer (2013) and Tenenbaum (2007). Kriegel (2019) has argued that moods fit this profile and in his work on Brentano (2018 a, 2018b) he has offered a way one might develop such a view for mental states more generally.

[4] The seminal discussion can be found in Kenny (1963). For further discussion, see Wilson (1972) and Teroni (2007). I agree with Teroni that the formal objects do not individuate emotion types. In my view, functional roles are a better candidate for that job. The formal object of an attitude tells us which property must be instantiated by the intentional object of the attitude in order for the attitude to be appropriate or fitting, and it is possible that distinct attitude types might have the same formal object. For a recent overview of fittingness see Howard (2018). The details surrounding this issue are beside the point given the main aims in the present essay.

[5] See, for example, Fara (2013) and Braun (2015).

[6] See Gregory (2020), Oddie (2005), Stampe (1987), and Velleman (1992) for further discussion of the ‘guise of the good’ theory of desire.

[7] Why not? There are at least two reasons. First, one need not possess the concept *truth* in order to believe, say, that grass is green. A child might not yet be in a position to predicate *truth* of things but might be in a position to think about colours and objects. But belief is a conceptual state, so to believe that grass is green is true places a cognitive demand on a thinker that outstrips the child who can think that grass is green. Second, the view that to believe that p is to believe that p is true seems to introduce a regress. If to believe that p is to believe that p is true, then to believe that p is true is to believe that p is true is true. For more, see the exchange between Sanky (2019 a, 2018b) and Grzankowski (2019). See also Merricks (2009). If one adopts a redundancy theory of truth matters are more complicated, but I’ll put that issue aside presently.

[8] See Schafer (2013) who also argues for an attitude view of desire and leans on an analogy with the relationship between perception and truth.

[9] See Gregory (2021, Ch 1) for detailed discussion of the role of the evaluative in the content of desire. According to Gregory's brand of the guise of the good, to desire that *p* is the believe that you have reason to bring it about that *p*. Important presently is his careful discussion of the way in which "reason to bring about" does *not* appear in the content-clause of desire ascriptions—on his view it features in the content of the *analysis* of desire rather than in desire itself. Also see Oddie (2009, Ch 2-3) for a defence of the view that desires should be understood in terms of seemings of the good. Importantly, Oddie's view is not that one who desires that *p*, desires that p is/be good. His view is that we should understand desiring that *p* in terms of *p*'s seeming good to one. More in the main text concerning the application of this structure to the case of emotion.

[10] See Helm (2015, 2024) for recent overviews of the finer-grained dispute.

[11] Gregory (2020) is especially clear on this in the context of desire. As noted in footnote 9, his theory is a modification of the guise of the good in terms of reasons.

[12] See Grzankowski (2021) for further discussion and see footnote 9 above.

[13] It is important to point out that the layered conception of content offered here is not to be *equated* with something like a sub-personal layer of representation. Though it could perhaps be advanced in such a way, this is not the aim of typical content theorists about the emotions nor about guise of the good theorists about desire. First, the content ascribed in the analysis layer is sometimes appealed to in accounting for the phenomenology of the states in questions. For example, things might *seem* good when one desires them – see Oddie (2019) – and snakes might *seem* dangerous when one fears them – see Mitchell (2017). Second, the layer I have in mind can be appealed to when aiming to make sense of the rational profile of the states in question in an internalist vein. For instance, one might hold that in fearing the snake one commits, in a person-level sense, to the snake's dangerousness (see Helm, 2015). The question presently at issue is where in one's metaphysics one who appeals to content to help explain these facts locates the relevant representational features.

[14] Ballard (2021) has also recently made this point.

[15] See Rossi and Tappolet (2019) who pressure the attitudinalist to make good on their notion of correctness not in content terms. They argue that "either the notion of correctness at work in the [attitudinal] theory is understood in terms of correspondence, and the attitudinal account is likely to collapse into a perceptual theory; or the notion of correctness is understood in normative terms such as 'reasons' or 'ought', and the attitudinal theory cannot motivate another of its main tenets, namely, that the formal

object determines a mental state's correctness conditions" (p. 544). For present purposes, I am happy to give away that the attitudinalist has work to do on this front. That said, the examples in the main text concerning fitting action suggest that accounting for correctness always in terms of content also needs substantiating. One option for the attitudinalist is to opt for a "fittingness first" account such as that advocated by McHugh and Way (2016) or Yetter Chappell (2012). For an overview of fittingness see Howard (2018). What is needed given the present dialectic is only that the :quick route" to the content view discussed in the main text is not a step that can be made without substantive support. If anything, there appears to be yet another stalemate emerging.

[16] See Schroeter *et al* (2015) for a development of this concern. An anonymous referee helpfully notes that the Attitude view will face a challenge here as well. Presumably the attitudinalist wants to hold that emotions can be evaluated for fittingness or correctness as noted in the previous section. If these evaluations are to be substantiated, they must be committed to evaluative properties against which the attitudes are compared for fit. Thank you to the referee for drawing this important point to my attention.

[17] See Orlandi (2020) for further discussion.

[18] An anonymous referee worries that such a view, when applied to emotion, will beg the question against a perceptual theorist. For example, one might worry that the perception of a banana as yellow is not meant to be utilised and available when there is no banana. But the view is not meant to not exclude the perceptual view and, in fact, is in keeping with the usual motivation for a representationalist view. As is familiar from cases of illusion and hallucination, things can visually seem to one other than they in fact are. In such cases, perceptual experience still, prima facie, "speaks" in favour of taking things to be as they seem. On the view presently under consideration, perception is a representational state on the grounds that when one perceives that p (perhaps in error), one is in a state which "says" that p is the case and which can be utilised downstream when forming beliefs and judgements.

[19] To clarify, here I am not using "representation" to designate both an attitude (such as perceiving or fearing) along with its content. By "representation", I aim to isolate the content of the mental state we are investigating in order to provide explanations of behaviour and performance in terms of it. Representations are here conceived of as processes or states that are individuated by their contents. Those processes or states can then be further classified as, say, beliefs, desires, emotions and so on in light of the role they play in the mental economy. Hence, the sense of "utilising" I am relying on is of the sort one finds in much of cognitive science and neuroscience as well as computer science:

facts about one process or state explains facts about another by appeal to what the states represent. In some cases, it is only because of the presence of information in the system that some other process or state makes sense. A very simple example from computing would be storing information in RAM and then retrieving it later. The content represented in RAM can be "utilised" later and that content will explain why later processes proceed as they do.

[20] It might be helpful to work through a couple other examples to bring out the sort of testing that I have in mind. First, consider Quilty-Dunn and Mendelbaum's recent discussion of the role of representations in a theory of belief and specifically in the explanation of semantic priming:

> Semantic priming is one of the most robust and well-validated effects in cognitive science. When a subject reads the word 'doctor', and then has to discriminate words from non-words (e.g., hit the YES key in response to 'bread' and the NO key in response to "drabe"), her reaction time will be faster in identifying semantically related words like "nurse" than in identifying unrelated words like "tree" (e.g., Meyer and Schvaneveldt 1971). One plausible explanation of priming is that mental representations are stored in associative networks such that activating one representation (by, e.g., reading the word that expresses it) activates representations connected to it in the network. (Quilty-Dunn and Mandelbaum, 2018, p. 2368)

As another example, consider perceptual binding – the operation that groups different features together into objects. Schneegans and Bays (2017) investigated the relationships between colour, orientation, and location. Performance on carefully designed tasks provides compelling reason to hold that orientation binds to spatial location and colour binds to spatial location, but colour and orientation are not directly bound to each other. By interfering with information about location, one finds that information about which colours go with which orientations is lost. Experiments like this one allow one to build a case for how representations are structured and what exactly it is that they represent – see Block (2023) for further discussion. I'm advocating for something even more basic: do emotions represent value properties *at all*, but I hope that this example gives the reader a fuller sense of the sort of methodology that is recommended. As noted above, the key idea is to build a case for whether or not the system's basic cognitive architecture and operations should be specified in terms of how internal representations of value are manipulated via computational algorithms.

[21] Thank you to an anonymous referee for pressing this point and to the audience at Cardiff University for raising a similar question.

[22] Perhaps a way to press even further here would be to argue against the notion of content and representation provided by the cognitive scientific approach at its roots. One might, for example, seek a Fregean rather than referential view of content or might think of content as part of an interpretive strategy rather than in as robust of a realist vein as I've presented it. The finer details here would require a paper-length treatment of their own, but I can say two things in brief. First, the cognitive scientific approach is not itself in conflict with an interpretive strategy provided that one is prepared to be anti-realist or interpretivist about what one learns from successful explanations in cognitive science. I prefer a more realist bent and think it allows for a clearer articulation of the view, but one could loosen one's realist grip and say much of what I've said above. Second, a functional approach to the Fregean aspects of content can be pursued. The inferential role semantics advanced by theorists such as Block (1986) and Harman (1993) fit into a broadly cognitive scientific approach but look to go beyond reference in individuating representations.