

Brentano on Consciousness, Intentionality, Value, Will, and Emotion: Reply to Symposiasts

Uriah Kriegel

Forthcoming in *European Journal of Philosophy*

It is a regrettable feature of this book symposium that it appears only after the book itself. If I could solicit from three outstanding philosophers detailed analyses of substantial portions of the book *before* publishing it, the book would have been far better. Below, I indicate some of the ways the book would have been better.

(I will refer to my book, *Brentano's Philosophical System: Mind, Being, Value*, as BPS, and to Brentano's main book, *Psychology from an Empirical Standpoint*, as PES.)

1. Mendelovici on Brentano on Consciousness and Intentionality

Angela Mendelovici discusses my interpretation and defense of Brentano's accounts of consciousness and intentionality. In both cases, I make claims both about what Brentano is trying to account for and about how he tries to account for it – that is, both about his explanandum and about his explanans. In the case of consciousness, I claim that Brentano is targeting the phenomenon we call today 'phenomenal consciousness,' the mental phenomenon for which there is (a warranted appearance of) an explanatory gap with the physical world. His account of this phenomenon, I claim, creatively combines elements of first-order and higher-order intentionalism: according to Brentano, a conscious experience of *x* is a single mental states that has no separable parts but does have an internal structure, a kind of 'notional' structure in virtue of which it can be legitimately framed both as awareness-of-*x* and as awareness-of-awareness-of-*x* (for more on what this means, see BPS pp. 30-40). In the case of intentionality, I claim

that Brentano is targeting the phenomenon of so-called phenomenal intentionality, a felt quality of endogenous directedness at the world. His account of this phenomenon, I claim, is at odds with most contemporary work on intentionality and construes it as an intrinsic modification of the subject. Thus, for you to start thinking of a table (rather than a flower, say) is not for you to enter an internal state that bears some specific relation to a table (and not a flower); it is for you, the person as a whole (not some *state* of you), to undergo an intrinsic experiential modification in virtue of which you could be correctly classified as a table-thinker (but would be incorrectly classified as a flower-thinker).

Mendelovici takes issue with my claims that Brentano's notion of consciousness is the notion of phenomenal consciousness and his notion of intentionality is the notion of phenomenal intentionality. She argues that his notion of consciousness is rather the notion of *transitive* consciousness ('consciousness of') and his notion of intentionality is the notion of *generic* intentionality (directedness as such). She is willing to concede that Brentano's notion of consciousness *picks out* phenomenal consciousness, and that his notion of intentionality picks out what might *amount to* phenomenal intentionality, but she stresses that these latter claims are weaker.

I agree with Mendelovici on both counts and regret having framed certain claims in terms of what Brentano's *notion* of such-and-such is. But what I wanted to communicate is nonetheless stronger than merely that phenomenal consciousness and phenomenal intentionality are the *referents* of Brentano's notions of consciousness and intentionality. For what matters to me is that phenomenal consciousness and phenomenal intentionality are what Brentano took himself to be giving theories of. And for this to be the case, it is insufficient that his notions of consciousness *happen* to pick out phenomenal consciousness and phenomenal intentionality; it must also be that he *meant* them to pick out phenomenal consciousness and phenomenal intentionality. And so if I were to write the book, per impossibile, *after* this symposium, I would rephrase my claims as weaker than 'Brentano's notion of consciousness/intentionality is...' but stronger than 'Brentano's notion of consciousness/intentionality picks out...' It would be something like: Brentano's notion of consciousness/intentionality picks out phenomenal consciousness/intentionality *and he takes it to pick out phenomenal consciousness/intentionality*. This midway claim is strong enough to undergird what really matters to me: that Brentano's theories of consciousness and intentionality were meant as theories of *phenomenal* consciousness and *phenomenal* intentionality.

Mendelovici's other major objection concerns the role of the unconscious in Brentano's philosophy of mind. In BPS, I argue that (i) Brentano's insistence that there is no unconscious mentality has been disastrous to his legacy and (ii) it can in truth be painlessly excised from his system. It was just an unnecessary wrong turn. Mendelovici casts doubt on both claims. As against (i), she argues that all Brentano meant by his rejection of unconscious mentality is that there are states of transitive consciousness of which the subject is unaware; and even if this is false, it is nowhere near disastrously false. As against (ii), Mendelovici argues that if Brentano welcomed unconscious mental states, then given that he takes intentionality to be the mark of the mental, he would have to allow for unconscious intentional states; but if Brentano took intentionality to be an intrinsic experiential modification of the subject, as I so ardently claim, then it is unclear how an unconscious mental state, which after all would just a neurophysiological state, could constitute an intentional state. Contemporary defenders of unconscious intentionality typically take it to consist in a 'naturalistically kosher' (read: broadly causal) relation between neural states and states of the environment. But for Brentano the obtaining of such relations would not constitute intentionality. What would constitute intentionality is an intrinsic experiential modification of the subject, and unconscious states could never constitute *that*. So, to welcome unconscious mental states into his system, Brentano would have to give up either (a) his conception of intentionality as an intrinsic experiential modification or (b) his thesis that intentionality is the mark of the mental. Both (a) and (b) are central tenets of the Brentanian system, and so welcoming unconscious mentality would hurt after all.

My responses are as follows. With regard to (i), I really think Brentano's rejection of unconscious mentality is *more* than just the claim that there are no states of transitive consciousness of which the subject is unaware. In BPS, I take Brentano's lengthy argument (in PES) against 'unconscious consciousness' to be at bottom an argument that there are no unconscious mental states. One could legitimately deny that this is what that discussion is trying to establish, perhaps along the lines suggested by Mendelovici. It remains that nowhere in his system does Brentano mention or make reference to unconscious mental states. So if his argument against 'unconscious consciousness' is not an argument against unconscious mental states, then apparently he thought no argument against unconscious mental states was even needed. Either way he is committed to there being no unconscious mental states. Now, my claim that this has proven disastrous to his legacy was intended in the first instance as a claim

about the history of science. Psychology textbooks typically bring up early broadly introspectionist empirical psychology only to set it aside quickly as naïve and misguided. But even as they go through the effort of doing this, they dwell primarily on such figures as Wundt and Titchener. Brentano's role in the inception of empirical psychology, having spawned generations of researchers who described themselves as psychologists, has more or less disappeared from the textbook history-of-psychology narrative. And as we will see momentarily, I think this can be traced back directly to his rejection of unconscious mentality.

Regarding Mendelovici's claim that reintroducing unconscious mental states into entrain significant revisions in core tenets of Brentano's system, notably either that intentionality is not an intrinsic experiential modification or that intentionality is not the mark of the mental: I wholeheartedly agree that what Brentano would have to do is deny that the kind of intentionality he has in mind demarcates the mental. This might seem like a substantial revision, but the underlying point is that Brentano's entire project in PES, which he presented as that of offering conceptual, methodological, and metaphysical foundations for a science of the mind, would have been better presented as the project of offering foundations for a science of *consciousness*. In my opinion, all the claims Brentano makes about the mind in PES are in truth just claims about consciousness, and he presents them as claims about the mind only because he thinks the mind and consciousness are one and the same. Even if he were right about this, I think he would have done much better to present his work as concerned with the foundations for the empirical study of consciousness and simply bracketed the (separate) question of whether anything outside consciousness merits being called mental. Now, once PES is framed as concerned with the science of consciousness, it is no longer surprising that the relevant mark claim should be that intentionality is the mark of the *conscious*, not the *mental*.

This relates to Mendelovici's last reservation about BPS. It concerns my claim that Brentano's philosophy should be seen as a live philosophical program rather than a somewhat esoteric theoretical edifice of largely historical interest. Mendelovici is unconvinced: she points out that insofar as the whole project of PES is to establish the legitimacy of psychology, it might after all 'seem quaint to us by now.' Now, it is true that Brentano himself presents PES as concerned with establishing the legitimacy of psychology. But this is only because he took all mental phenomena to be conscious

phenomena. Had he proceeded slightly differently, defining psychology as the science of mental phenomena and then immediately bracketing all parts of psychology potentially concerned with unconscious mental phenomena, thus presenting his book as an attempt to establish the legitimacy of an empirical science of consciousness, then far from seeming quaint to us now his project would strike us as of urgent relevance. For we still have no clear foundations for a science of consciousness itself – as opposed to a science of the neural, cognitive, and behavioral correlates of consciousness – and the very legitimacy of such a science is commonly greeted with suspicion in the cognitive-scientific community. Thus Brentano’s project in PES, which in reality addresses the foundations of what many think of as science’s last frontier – the empirical study of consciousness – comes across as a quaint project because he refused to cordon off the question of unconscious mentality. This is exactly what I mean by ‘disastrous to his legacy’!¹

2. Olson on Brentano on Good and Better

Jonas Olson discusses my interpretation and defense of Brentano’s account of value. In BPS, I formulate Brentano’s ‘fitting attitude’ account of value, to a first approximation, as follows:

(FA) For any good g , (i) it is fitting to take a pro attitude toward g , and (ii) g is good *because* (i).

But in truth Brentano’s complete account supplements this analysis of goodness with corresponding analyses of badness and betterness:

(FA⁻) For any bad b , (i) it is fitting to take a con attitude toward b , and (ii) b is bad *because* (i).

(FA[>]) For any x and y , such that x is better than y , (i) it is fitting to prefer x over y , and (ii) x is better than y *because* (i).

Brentano’s complete account of value can be seen as the conjunction of FA, FA⁻, and FA[>].² In his commentary, Olson proposes an improvement on FA (and by extension FA⁻

and FA[>]) and presents two objections to Brentano's account that I have not considered in BPS, both of which are at least more perspicuous when directed at FA[>].

Olson's proposed improvement of FA is to replace 'because' in Clause (ii) with the locution 'what it is for...', so that the account reads as follows:

(FA^o) For any good *g*, (i) it is fitting to take a pro attitude toward *g*, and (ii) what it is for *g* to be good is for (i) to obtain.

This is an improvement, claims Olson, because it allows us to accommodate certain intuitions that FA has difficulty accommodating, notably the intuition that 'the fact that it is fitting to take a pro attitude toward *g* is not what *makes g* good.' FA appears to entail that the fittingness of a pro attitude toward *g* is what makes *g* good. FA^o does not entail this, according to Olson: what *makes g* good is whatever descriptive properties *g* has that make it fitting to take a pro attitude toward *g*. Thus FA^o makes clear that the fittingness of a pro attitude toward *g* 'explains *that g* is good, but it does not explain *why g* is good, or what *makes g* good.'

I agree that this would improve Brentano's dialectical position and therefore that BPS would have benefited from precisifying 'because' in the way Olson proposes. But even in this improved form, claims Olson, Brentano's account of value faces two important objections.

The first concerns quantitative value comparisons. Compare the following two truths:

- (1) The experience of eating chocolate ice cream is better than the experience of eating turnip.
- (2) The experience of eating chocolate ice cream is better than the experience of being tortured.

Both (1) and (2) are true, but there is a difference, which we may put, rather uncouthly, as follows: the experience of eating chocolate ice cream is 'more better' than the experience of being tortured than it is better than the experience of eating turnip. Now, given FA[>], Brentano would paraphrase (1) and (2) into (3) and (4) respectively:

- (3) It is fitting to prefer the experience of eating chocolate ice cream over the experience of eating turnip.

- (4) It is fitting to prefer the experience of eating chocolate ice cream over the experience of being tortured.

But how can Brentano make sense of the 'more better' fact just noted? There seem to be only two options. Either (a) it is *more fitting* to prefer ice cream over being tortured than to prefer ice cream over turnip, or (b) it is fitting to prefer *more strongly* ice cream over being tortured than over turnip. However, as Olson notes, (a) is incompatible with Brentano's construal of fittingness as an on/off affair for, and (b) would require us to constantly and anxiously supervise our preferences so they are neither too exuberant nor too nonchalant.

This is a formidable objection. But I wonder: might Brentano try to accommodate the 'more better' fact in terms of a *second-order preference*? The idea follows the intuition, in option (a), that preferring ice cream over torture is more fitting than preferring it over turnip. But rather than unpacking 'more fitting' as meaning 'has a greater amount of fittingness in it' (which makes fittingness gradient), we unpack 'more fitting' as meaning 'is to be preferred,' that is, 'it would be fitting to prefer it.' The upshot is that the 'more better' fact is assayed as the fact that it is fitting to prefer the preference for ice cream over torture over the preference for ice cream over turnip. This response would need to be more fully developed, and would likely entrain its own web of complications, but I see no better option for Brentano here.

Olson's second objection to FA[>] is that it cannot accommodate the fact that it is sometimes permissible to prefer one of two equal goods. Suppose Tim and Tom are having an equal amount of fun, but only Tim is your child. Intuition instructs, on the one hand, that it is fitting to prefer Tim's fun over Tom's, but on the other hand, that Tim's fun is not inherently better than Tom's. Yet FA[>] claims that Tim's fun is better than Tom's iff it is fitting to prefer Tim's over Tom's. Something has to give. Olson considers on Brentano's behalf three fairly ingenuous responses, but demonstrates that each is unsatisfactory.

Without undue confidence, I want to float a fourth possible response, namely, that the term 'fittingness' is ambiguous, with the sense implicated in FA[>] being different from the sense responsible for the intuition that the parent's preference is fitting. To see the two senses, consider this analogy in the realm of belief. When is the belief that *p* fitting? There is one way of hearing 'fitting' such that a person's belief that *p* is fitting

if the evidence at the person's disposal recommends believing that p . But there is another way of hearing 'fitting' such that the belief that p is fitting just if p is true. Consider Russell's chicken: 'The man who has fed the chicken every day throughout its life at last wrings its neck instead, showing that more refined views as to the uniformity of nature would have been useful to the chicken' (Russell 1912: 63). Let us suppose that the chicken's last belief is that *food is coming*. Is this belief fitting? In one sense yes and in another no. On the one hand, it is fitting for the chicken to believe that food is coming, insofar as the evidence in her possession supports the belief. On the other hand, it is *unfitting* for the chicken to believe that food is coming, insofar as it is *not* coming. What we should say is that the chicken's belief is fitting₁ but not fitting₂. I propose that an analogous distinction attends the fittingness of preferences. Very informally: in one sense, preferring x over y is fitting just when the subject is in some sense justified in having this preference; in another sense, preferring x over y is fitting just when x is better than y . Call the first *fittingness*₃ and the second *fittingness*₄. (Importantly, although talk of justification and truth, and their analogs in the realm of preference, is useful to 'focus the mind' on the different senses of fittingness, we are not here taking a stand on the ultimate order of explanation between each notion of fittingness and the thing we advert to in trying to focus the mind on it.) What I want to suggest, now, is that what intuition instructs is that it is fitting₃ for a parent to prefer her own child's joy, whereas what FA[>] implies is that it is not fitting₄ to prefer one's own child's joy. There is no strict contradiction here, though there is surely a kind of tension that would need to be addressed in a full account of betterness.

Neither Olson nor I believe Brentano's fitting-attitude account of value. Olson (2014) prefers an error theory, according to which nothing is good or bad and any claim that something is good or bad or better than another is simply false. I prefer a response-dependent account, according to which what it is for g to be good is for g to be disposed to elicit a certain type of appreciation in an ideal subject under ideal conditions. Olson's objections constitute two good reasons for our preferences.

3. Textor on Brentano on Will and Emotion

Mark Textor discusses my interpretation and defense of Brentano's account of the difference between will and emotion. It is clear that Brentano took states of the will and emotional states (*Gefühle*) to belong to a single fundamental category of mental states, which he referred to variously as 'phenomena of love and hate,' 'phenomena of interest,' or 'emotions/affects' (*Gemütsbewegungen*). The question is what distinguishes will and emotion as two subclasses within this fundamental category. On my interpretation of Brentano, the difference is in the attitude that emotional states and volitional states take toward their contents, and therefore goes to their intrinsic nature. On Textor's interpretation, the difference is rather extrinsic and has to do with (i) these states' contents and (ii) the beliefs they are based on. I will argue that my interpretation is better.

Brentano's fundamental classification of mental states divides them into states that present their objects under the guise of truth or falsity, states that present their objects under the guise of good or bad, and states that present their objects under no guise – without 'commenting,' so to speak, on the objects' truth or value. This is a difference in the *attitude taken toward* the object: states of the first category present-as-true/false their objects, states of the second category present-as-good/bad their objects, and states of the third category present-neutrally, or 'merely-present,' their objects. Willings and emotings belong to the category of states that present-as-good/bad. What distinguishes them, on my interpretation of Brentano, is the specific *kind* of presenting-as-good/bad they involve:

(K1) For any mental state M, M is an emotional state iff M presents-as-*prima-facie*-good/bad its object; M is a volitional state iff M presents-as-*ultima-facie*-good/bad its object.

On this interpretation, the key difference between emotion and will is the difference between presenting an object under the guise of the *prima facie* valuable and presenting it under the guise of the *ultima facie* valuable.

In the book, my case for my interpretation involves centrally considerations of alternative accounts. Two of them, I note, do find some textual support in the Brentano corpus. The first is that the difference between emotion and will goes to content rather than attitude. Specifically, volitional states are *token-reflexive*: a desire that I eat ice cream does not have the content 'I eat ice cream' but the content 'I eat ice cream as a

result of having this very desire.’ In contrast, emotions lack such a token-reflexive content: when I am happy that I am eating ice cream, the content is simply ‘I am eating ice cream.’ The other alternative is that the difference between emotion and will goes to certain beliefs that they involve. Specifically, volitional states are ‘based on’ beliefs to the effect that certain outcomes would be brought about by certain actions; emotions are not based on such beliefs. In BPS, I argue against each of these options. But one thing I do not do is consider a view that *combines* both these elements. According to Textor’s interpretation, Brentano takes the combination of token-reflexive content and the relevant beliefs to distinguish volitional from emotional states. He puts his account in terms of the difference between ‘desire’ and ‘love,’ but the difference is supposed to generalize to volitional states other than desire and emotional states other than love:

- (TI) S desires [or otherwise wills] the outcome A iff
- (a) S knows [or just believes] that his loving [or otherwise positively emoting about] A will tend to bring about A;
 - (b) S loves [or otherwise positively emotes about] A in part because (a).

Textor argues both (i) that TI is more plausible to attribute to Brentano than KI and (ii) that TI is the better view regardless. The two things are related, of course: where there is textual support for two interpretations, it is more charitable to attribute the more plausible view.

Starting with (ii), I confess that I do not think TI is a very good view to have. Suppose I want, more than anything else, to be published in *Phil Review*. I submit a paper every other year. It gets rejected without comments every time. As years go by, I no longer think it very likely that I will get a paper into *Phil Review*. But my desire has only intensified and I start submitting once a year. I do not submit, though, because I believe my submissions tend to result in publications. And even if I believed this, ultimately I submit because I believe a publication would give me professional glory, not because I simply believe that submitting tends to result in acceptance.

Naturally, according to Textor it is rather KI which is not very good. His sole argument against it, though, is this:

Kriegel’s distinction [between presenting-as-*prima-facie*-good and presenting-as-*ultima-facie*-good] presupposes that emotions and acts of the will are directed on something that forms part of a ‘mutual detestation society’ [i.e., collections of incompatible things].

But there are things that are not members of such societies. I have, for example, no opposite: there is no Anti-Textor. Yet, you may be angry with me or, much better, like me. But your anger cannot present as *prima facie* bad. Hence, the distinction between emotion and the will cannot be based on the *ultima/prima facie* distinction.

There are two problems with this argument. First, even if it were true that *x* can be sensibly described as *ultima facie* good only if *x* has incompatibles it is better than, it would not follow that *y* can be sensibly described as *prima facie* good only if *y* has incompatibles it fails to better. And so Textor could be presented as *prima facie* good (in liking him) even if he could not be presented as *ultima facie* good. And indeed, emoting about Textor seems psychologically possible whereas willing Textor seems impossible – a category mistake. Second, from the fact that for *x* to *ultima facie* good it must better all its incompatibles, it does not follow that *x* can be sensibly described as *ultima facie* good only if *x* has incompatibles. For if *x* has no incompatibles it *trivially* satisfies the condition of bettering all its incompatibles. Compare: from the fact that Sasha is an only child it does not follow that it is a category mistake to say that Sasha is her parents' oldest child; on the contrary, it is *trivially true* that she is.

I conclude that Textor's argument against KI fails and at the same time TI faces serious difficulties. Still, the worse view could have been Brentano's. But I think things are more complicated. I did find Textor's reconstruction of Brentano's discussion of the emotion/will divide in PES, as suggesting something like TI, quite persuasive. Although in BPS I acknowledged that the token-reflexive and belief-based interpretations have textual support, I did not appreciate just how deep that support goes, nor considered that they could be combined into a single unified account. The book would thus have profited from a more sustained discussion of TI. My current inclination is to see TI as capturing Brentano's 1874 account of the emotion/will distinction, but to see it as later superseded by KI, which, as I note in BPS, is aired in passing in lecture notes from 1894 at the latest, a 1908 letter to Oskar Kraus, some remarks in the embellished 1911 edition of (the later chapters of) PES, and most fully, in Brentano's 1907 manuscript 'Loving and Hating.'

Textor argues that a passage I quote from the lecture notes is not really concerned with the emotion/will distinction. Be that as it may, the other passage I quote in BPS, this time from §22 of the 1907 essay, says the exact same thing and is clearly concerned with the emotion/will distinction. Moreover, in the same section

Brentano makes certain claims that explicitly contradict TI. I did not dwell on this in BPS because I did not take TI (qua interpretation of Brentano) as seriously as I do now, but if we do take it seriously, we are naturally led to believe that Brentano changed his mind about the emotion/will distinction sometime between 1874 and 1907. Consider, then, this passage, excerpted from a single paragraph, which both makes clear that the topic is the emotion/will distinction and states explicitly that condition (a) of TI – that volitional states involve beliefs in their own causal effects – need not be satisfied by volitional states:

How then are we to draw the distinction between love of the more simple sort and desiring, wanting, and willing?... [I]n order to become an object of will, want, or desire, the thing that is loved must be preferred, not only to some one thing that is incompatible with it, but also to every possible object that is thought to be incompatible with it... It should be noted that I can thus want or desire a thing without at all believing it to be something I can bring about myself. I can want or desire that the weather be good tomorrow, but I have no choice in the matter. (Brentano 1989/1969: 151)

Clearly, the author of this essay does not think that what distinguishes volitional from emotional states is that they involve beliefs about what these states tend to bring about. The author of this essay is also 23 years older than the author of the passages in PES that Textor cites in favor of attributing TI to Brentano. And TI is also much inferior, I have suggested, than KI. Taken together, all these considerations recommend attributing KI to Brentano as his *ultimate* account.³

References

- Brentano, F.C. 1889/1969. *The Origin of Our Knowledge of Right and Wrong*. Trans. R.M. Chisholm & E.H. Schneewind. London: Routledge and Kegan Paul.
- Mendelovici, A. 2018. *The Phenomenal Basis of Intentionality*. Oxford and New York: Oxford University Press.
- Olson, J. 2014. *Moral Error Theory: History, Critique, Defence*. Oxford and New York: Oxford University Press.
- Russell, B. 1912. *The Problems of Philosophy*. Oxford: Oxford University Press, 1997.

¹ I am mindful, of course, that in her own work Mendelovici has taken a 'largely eliminativist' stance on unconscious intentional states (Mendelovici 2018: 191). But I am not predicting this will prove disastrous to her legacy! Mendelovici's project is to provide a comprehensive theory of intentionality that grounds it fully in phenomenal consciousness; success here is to be judged by the strength of argument for individual components of the theory and the unity of among the components of the overall theory. Brentano's project was very different: to provide some research area with conceptual, methodological, and metaphysical foundations. Here success would have required that these proposed foundations would command a measure of consensus among researchers and guide their research – none of which has materialized.

² The reason FA[>] is needed is that Brentano rejects an analysis of 'x is better than y' in terms of x having more goodness in it than y. His reasons for this are discussed in BPS pp. 240-1 and are nicely explained in Olson's piece.

³ Let me also flag here two errors, potentially typographic, in Textor's commentary. First, Textor writes 'Kriegel works through different answers to this question [namely, what distinguishes emotions and acts of the will?] suggested by Brentano and finds them all wanting.' But I do not in fact find KI wanting; I record that it entails that desire is an emotional rather than volitional state, but I *defend* this implication (BPS pp. 205-6). Second, Textor writes that I end up 'classifying desires and decisions as emotions.' This is true of desires, but not of emotions. On the contrary, I take decisions to be, and to be Brentano's, the *paradigm* of a volitional state (BPS p. 205). This comes through very clearly in the aforementioned 1908 letter: 'All acts of will, in the strict sense, consist of decisions. It is not possible to will incompatible things.' (Brentano 1889/1969: 114)