EMOTIONAL TRUTH

by Ronald de Sousa and Adam Morton

I—Ronald de Sousa

ABSTRACT  Taking literally the concept of emotional truth requires breaking the monopoly on truth of belief-like states. To this end, I look to perceptions for a model of non-propositional states that might be true or false, and to desires for a model of propositional attitudes the norm of which is other than the semantic satisfaction of their propositional object. Those models inspire a conception of generic truth, which can admit of degrees for analogue representations such as emotions; belief-like states, by contrast, are digital representations. I argue that the gravest problem—objectivity—is not insurmountable.

I

Generic Truth. A ‘true likeness’ is not one that is not false. When we say that Tolstoy’s novels are true to life, we don’t mean to claim that they are, after all, non-fiction. In these and some other domains we speak of truth, but assume we are not speaking strictly. Must this be the case for emotional truth? The phrase sometimes refers to kindred properties such as authenticity, a difficult notion worth elucidating, but about which I have little to say. I propose instead to take literally the idea of truth-valued emotions.

The concept of emotion is Janus-faced. In one direction emotions face inward, either as ‘perceptions referred to the soul’ (in Descartes), or as perceptions of bodily states aroused by some exciting cause (in William James). In the other direction emotions face outward, suggesting that (at least some) emotions provide us with correct or incorrect representations of something in the world outside us. It is in this facing-out stance that emotions might claim to be literally true or false. In pursuit of this hypothesis, I shall offer some reasons for assigning a broader scope to the concept of truth as correspondence, and survey some of the difficulties that such an extension to emotions of the idea of literal truth may bring.

A mental state M can be said to be true or false, only if

(1) it is subject to a norm N;
(2) N is determined by M itself, yet
(3) N looks for its satisfaction to some reality existing independently of M.

These are necessary but not sufficient conditions for standard ascriptions of truth-value. I shall postulate that they are sufficient to capture the core of a generic notion of truth, which might be summed up in this slogan: *A story defines its truth, but whether it is true can never be part of the story.* My question, then, is how this might apply to the relation between an emotion and that which, if anything, the emotion represents.

In philosophy the entities to which truth value is attributed directly are commonly held to be not mental states, but propositions. Mental states that incorporate propositions in a suitable way (I shall speak generally of ‘attitudes’) inherit truth-value from their propositional objects. But it is unclear what propositions are. It seems safe to regard them as posits tailored to play just two roles: as objects of propositional attitudes, and as bearers of truth-value. Among the attitudes, beliefs then remain typical of those that admit of truth or falsity. They obviously satisfy the conditions just stated: a belief (1) specifies or ‘expresses’ a proposition; (2) it thereby determines a norm, according to which it is true or false; and (3) the satisfaction of that norm is independent of the belief’s existence.

A tight connection therefore seems to hold between truth and belief. For as is often noted—and as witnessed by Moore’s paradox—the *aim* of belief is truth. Nevertheless, it would be misleading to say that truth-value belongs essentially to belief. For, as Frege made clear, we need to allow that propositions may remain unasserted. Otherwise the antecedent and consequent of any conditional would be asserted merely in virtue of figuring in the conditional, trivially short-circuiting Modus Ponens. More generally, belief’s monopoly on truth might be infringed in two ways. First, some attitudes may lack propositional objects and yet also be true or false. *Perceptions* may provide examples. Second, one might attribute a truth-like property to other propositional attitudes, differing from beliefs in their aim. An example would be *desires*. In this second class of cases, one might ascribe truth-value derivatively to the attitude on the basis of the truth-value of its object. Thus one might say that my desire for
oyster ice-cream is true iff I get some. But a more interesting analogy with the truth of beliefs would focus not on the semantic satisfaction of propositional objects, but on the attitude’s attainment of its aim, its success. For belief, success is truth; but it lies elsewhere for other attitudes.

Emotions may stake a claim under both headings. Like perceptions, they sometimes lack a propositional object. And as in the case of desires, the truth of their propositional object does not define their success even when they can be said to have one. To see this clearly, recall Robert Gordon’s observation that some emotion ascriptions are ‘factive’, while others are ‘epistemic’.1 The former, such as $S$ is embarrassed that $p$, presuppose that the subject knows that $p$, while the latter, such as $S$ fears that $p$, presuppose that the subject does not know whether $p$.2 Obviously in the latter case, the truth-value of $p$ does not determine the appropriateness of the emotion. Even in the factive case, however, the truth of $p$ is not sufficient to vindicate the emotion. By contrast, the truth of $p$ always vindicates the belief or the assertion that $p$.

But if semantic satisfaction does not determine the aim of an emotion, what does? To answer this we need to proceed on two fronts. We must explore the way in which a state without a propositional object might be true; and we must ask what might define the success of an emotion, in a way precisely analogous to the sense in which truth defines the success of a belief. The former quest will look for inspiration to the model of perception. The second will explore what it is for an emotion to be, in the relevant sense, successful.

Proceeding somewhat indirectly, I begin by acknowledging some of the difficulties that might threaten to scuttle the project before it gets off the ground. The most insistent difficulty, concerning the prospects for emotional objectivity, will bring me back both to the analogy of perception and to the reconstruction of a relevant notion of emotional truth as success. I shall conclude by sketching some reasons to adopt the suggestion that the

2. What Gordon actually says is that emotions are factive or epistemic. But that cannot be quite right, for several reasons that don’t bear on the present point. (See de Sousa 1991).
specific domain of propositional truth, for which we are accustomed to reserve the literal meaning of truth, is distinguished by its digital as opposed to analogue mode of representation.

II

Some Logical Problems. A basic feature of the paradigm truth-valued states—propositions, assertions, or beliefs—is that they can be negated. Furthermore, where the embedded proposition exhibits a subject-predicate structure, it can be negated in two ways, yielding contraries distinct from contradictories. Thus $p$ is false if and only if $\neg p$ is true, while $Fa$ is false if $\neg\text{not-} F(a)$ is true. Is any such pair of standards for negation applicable to emotions?

Some named emotions are commonly felt to be polar opposites (love and hate, hope and despair, admiration and contempt, gratitude and resentment). Such pairs may plausibly be regarded as contraries, while equanimity—rather than indifference—might relate to both as their contradictory. But how are the norms of contrariety to be grounded?

Compare the case of desire. Beliefs demand consistency: if $p$ and $q$ are inconsistent, that inconsistency is automatically an indication that belief that $p$ and belief that $q$ cannot both be right. By contrast, someone might suggest that no such demand exists for consistency of desire. For two desires may aim at inconsistent states of affairs without entailing that at least one must be mistaken.

This is partly right, but harbours an important confusion. It presupposes that a single criterion of consistency is appropriate to both beliefs and desires. But that presupposition begs the question against the distinction alluded to above, by confusing the satisfaction conditions of desire with conditions of success. For any two beliefs, compatibility coincides with consistency. But for two desires to be consistent, it is not necessary that their contents be jointly satisfiable, but only that their contents be jointly desirable. So while a desire for $p$ and a desire for $q$ (where $q$ implies $\neg p$) are clearly incompatible, it does not follow that they should be regarded as inconsistent. And while this raises difficult questions about how to cash in the claim that two desires are

inconsistent, it makes room for inconsistent desires without requiring that consistent desires also be for compatible objects.

Emotions are similar, but messier. The reason is that there is no single proper object of all emotions. Each emotion is linked to its own specific evaluative continuum, and so defines its own proper object, and thereby the dimension along which contrariety might be defined for that emotion.

To illustrate how the distinction between truth and satisfaction might work out for a standard emotion, consider the example of fear. This can readily be construed as having been honed by natural selection to favour the avoidance of danger. The formal object of fear—the norm defined by fear for its own appropriateness—is the Dangerous. Fear that \( p \) is satisfied iff \( p \) is true, but it is successful iff \( p \) is actually dangerous. In general, for any emotion sufficiently complex to afford the identification of a propositional object:

\[
E(p) \text{ is satisfied iff } p \text{ is true} \\
E(p) \text{ is successful iff } p \text{ actually fits } E\text{'s formal object.}
\]

Where the emotion admits of a target \( t \) but lacks a propositional object (as in certain kinds of fear), semantic satisfaction consists in successful reference, while success still depends on whether the target fits the formal object:

\[
E(t) \text{ is satisfied iff } t \text{ exists} \\
E(t) \text{ is successful iff } t \text{ actually fits } E\text{'s formal object.}
\]

In all cases, the emotion's success is independent of semantic satisfaction. Fear of monsters is not semantically satisfied, but may be successful. The converse may be the case in fear of spiders.

Emotional truth, then, refers not to semantic satisfaction, but to success. I follow widespread practice in saying that fear's assessment of \( p \) or \( t \) as dangerous consist in some sort of evaluation of \( p \) or \( t \). Success is tied to the correctness of that evaluation, and I will need to say more below about how the

---

4. This skirts around a current debate about whether there can be non-conceptual contents of perception. Some hold, while others deny, that perceptual content may be non-conceptual. Even in Pittsburgh, however, where all content is conceptual and every concept owes its identity to the inferences it licenses, it doesn't follow that every perception must boast a propositional object.
evaluation relates to the rest of the emotional experience. But this suffices to suggest how the notion of opposition, if any, appropriate to a given emotion is internal to that emotion. And while this provides no handy criterion of emotional contrariety, it at least suggests a way in which such a concept might have application, as well as explaining why it is difficult to cash out in practice.

Another disanalogy is sometimes adduced between belief and desire and might apply *a fortiori* to emotions. When deliberating about what to do, there comes a moment when it is appropriate to say: *now is the time to decide*. And then one does so, definitively and rationally. But deliberating about what to *believe* is different. For that amounts to making up one's mind about what is *true*, and there is always a gap between the rationality of making up one's mind about *p* and the truth of *p*. At best there can come a moment when I am justified in making up my mind. But that cannot give me a rationally sufficient ground for the *truth* of the proposition. At best, it can be the right moment to decide only on the rationality of *behaving as if it were true*.5

Yet this contrast too is misleading. Admittedly, the pressure of time can furnish a reason to decide (that it is rational) to believe that *p*, but can never be *evidence* for *p*. What is rational to believe is only my best bet under current constraints. Nor is deciding to believe *p* equivalent to deciding to act as if *p* were true. But the following parallel still holds: while the pressure of time and other constraints can be a perfectly good reason for *deciding (that it's rational) to do p*, it can't be grounds for the proposition that *p* is *objectively best*, or even that *p* is what will seem best in the light of infinite consideration.

The crucial disanalogy between beliefs and other attitudes lies elsewhere. Only one of two incompatible beliefs can be true, and therefore only one can be successful. Among incompatible desires or emotions, on the other hand, no single desire or emotion need be uniquely successful.

The very idea of an *objective best*, however, may seem to beg the question against a prevalent view that neither emotions nor

---

5. This point was made by David Owens in comments on a paper by Gary Watson, at a Montreal conference on Akrasia, May 2001. Since I draw on Owens' oral presentation and conversation, I can't be sure that he would endorse my formulation.
desires can be assessed in terms of anything objective at all. To this I now turn.

III

The Claim of Objectivity. On the problem of emotional objectivity, Plato made an early start on two fronts. The *Philebus* argued that pleasure can be false, not merely in the derivative sense of being associated with or caused by a false belief, but in itself. That claim, extended to (other) emotions, presupposes that there can be an objective correlative to pleasure or emotion that is not a mere projection. That demand is made explicit in the *Euthyphro*, where Plato posed the problem of whether the gods love piety because it is pious, or whether calling it pious is merely to claim the gods love it.

The meaning of ‘objectivity’ is subordinate to the contrast with ‘subjectivity’, and that term has at least a dozen different senses. But a clear paradigm of objectivity can arguably be found in mathematical intuition. Imagine someone saying: *I understand your statement that all triangles have three sides, but I disagree.* One would be confident in objecting: *Your ‘disagreement’ suffices to show you have not understood.* But unexpectedly, at the other, most subjective end of the spectrum, it seems plausible to admonish someone who doesn’t share my individual tastes with a curiously similar demand for taste universalization:

*(TU) If oyster ice-cream tasted to you as it does to me, you could not fail to find it delicious.*

The cases seem very different, for in the case of taste we lack any independent way of supporting the counterfactual. Nevertheless, it is instructive to explore this further. Since taste exemplifies extreme subjectivity, it would not be surprising if it failed to meet the conditions for generic truth. If despite that it succeeds, on the other hand, we may assume that more complex emotions will also pass the test. If not, then to diagnose exactly how it fails might help us to discern how emotions must differ from taste if their claim to be truth-valued is to be vindicated.

If my taste for oyster ice cream (TOI) could be said to be truth-valued, the following must hold: (1) TOI must be subject to a norm of appropriate liking or aversion (the ‘valence’), (2) that norm of appropriate liking or aversion must somehow be defined by the character of the taste (the ‘quale’), but (3) the quale cannot suffice to determine the satisfaction of the norm in question.

One problem is to make sense of the quale’s defining its own norm of success—the appropriateness of liking or aversion. A second problem is how to make the relation between the quale and the valence contingent: it must be possible for an inappropriate valence to be present or an appropriate one to be absent. In other words, if taste is really objective then taste universalization (TU) must fail.

Suppose we think of all actual experience, on the model of a mathematical domain, as located in a multi-dimensional space encompassing all possible experiences, had by any possible conscious beings. Only some very limited ranges of experience are available to any specific kind of conscious being. (Human experience of colour, for example, can provide only partial insight into the experiences available to tetrachromatic animals.) At any point in that space, the valence of a specific experience is one of the qualitative dimensions of experience. It would then follow trivially that two experiences could not be qualitatively identical while being opposed in valence. On this picture, is there any prospect of prying apart the valence of an experience and its quale?

There are two possibilities. Either valence is a component of complex qualitative experience, or it is supervenient on other qualia. Understood in the first way, valence could always in principle be dissociated from the concomitant qualia. If my experience of oyster ice-cream consists in qualia \([A, B, C, \text{ liking}]\), and yours in \([A, B, C, \text{ aversion}]\), then we are not having the same experience. This would not preclude the required contingency. But what could be the measure of appropriateness between \([A, B, C]\) and one or another valence?

It is tempting to appeal to Human Nature to set a standard of correctness. Given any quale, an evaluative response that falls foul of the norm will lack appropriateness, and on that basis we can call it perverted, abnormal, or \textit{false}. The problem with human nature, however, is that if it refers to a set of interesting properties true of all and only humans, and robust enough to
support normative standards, then there probably is no such thing.\textsuperscript{7} Still, the suggestion is worth setting on ice for partial recuperation and reconstruction, as I shall suggest in a moment.

Now consider the second possibility, that valence supervenes on other qualia. If supervenience is understood deterministically, it will preclude the required contingency of the relation between quale and valence. But the laws governing that relation of supervenience might be stochastic, allowing two or more alternative outcomes. That would restore contingency, and providing one valence can be made out to be more appropriate than the other it would then satisfy the conditions for generic truth after all.

This last requirement remains very far-fetched in the case of taste. But it is much more likely to be met in the case of those emotions that are plausibly characterized as perceptions of value. Take, for example, the classic thought experiment in Mencius: you see a child about to fall into a well, and your apprehension of the situation immediately moves you, and you want to save the child. In this instance, what is apprehended is the \textit{need to intervene}. Or better it is the nature of the total situation, in which \textit{the need to intervene} roughly sums up the supervenient valence. Yet it is not impossible to witness the scene without being moved thus. Anyone whose experience lacks the appropriate valence, however, may be said to have an objectively false emotion.

This way of describing the situation avoids simple projectionism, insofar as what I perceive is not merely the shadow of my own response, but something about the character of a situation as a whole in the context not only of my own singular responses but of the feelings and interests of others. The choices to which I am led are products of a multi-dimensional landscape of values constituting a larger axiological whole. I call this view ‘axiological holism.’ It stipulates that we do not apprehend value in discrete units, but only in the light of a complex of factors that transcend individual experience. No single range of facts suffices for the success of an emotional response. Biological facts will speak to its origins and may thereby assign it a proper function in the sense of Millikan (1989), but they will not determine its relation to currently relevant norms. Social norms, in turn, are every bit as likely to be irredeemably nasty as biological ones.

\textsuperscript{7} de Sousa 2000.
(To endorse social norms as the touchstone of normativity would be to condemn all social reformers.) Individual biography sets up paradigm scenarios in terms of which each individual understands the world, but this defines only a narrow sense of fit between a current response and a present situation. That fit cannot be identified with value in any comprehensive sense, still less determine what is morally right.8

All of these factors—biological, social, or personal, and more—may properly be confronted with one another in the hope of arriving at something like reflective equilibrium. That holistic equilibrium is as close as we can come to reconstructing a notion of normative human nature. And perhaps it is close enough. In this way, we may find some emotional responses mistaken, just as Macbeth found (‘Art thou not, fatal vision, sensible to feeling as to sight?’) that a perception can fail the test of corroboration by different sensory channels. Vision provides distal information about our surroundings; yet visual illusions occur. Similarly emotions in general constitute apprehensions of axiological reality; yet not every emotion is equally to be trusted. We tell which is right and which is wrong much as we test the veracity of perceptual information: by appealing to corroborating evidence. Something like the method of reflective equilibrium is commonplace in science as well as in ethics; what is not often noticed is that the items that need to come to equilibrium are typically emotional responses. The search for reflective equilibrium plays an important role not just in moral deliberation but also where the issue is purely epistemic, where, as Christopher Hookway (1998) has argued, emotions such as the feeling of plausibility or doubt play a crucial role. Without such emotions, even the most comprehensively rational argument may remain powerless to move us.

Equilibrium, it may be objected, establishes only coherence. And it is an excessively weak theory of truth that is satisfied with coherence. Compare perception again. Each sensory channel provides a specific mode of information. But primary qualities are apprehended through different sensory channels. Multi-modal access is the warrant of objective reality. What then is the analogue of multi-modal access for emotions?

IV

*The Scope of our Emotional Access to Value: A Musical Analogy.* A helpful analogy is suggested by a fascinating paper by Dmitri Tymoczko on Milton Babbitt and John Cage. Tymoczko describes both composers as philosophers who, in their different ways, questioned the relevance of beauty or pleasure to aesthetic appreciation. Babbitt, in particular, aims to break the link between value and the ordinary listener’s emotional response to music:

Advanced music, to the extent that it reflects the knowledge and originality of the informed composer, scarcely can be expected to appear more intelligible than [higher mathematics] to the person whose musical education usually has been even less extensive than his background in other fields.

But are the compositions based on these principles *musically* intelligible, even to experts? ‘Do they lead to perceptible features of the music that can be understood through listening?’ asks Tymoczko. Babbitt’s defence of his esoteric compositions seems to presuppose that the relationships he elaborates are indeed perceptible as acoustic patterns, albeit only after special training. But Tymoczko gives some reasons for doubt, and points out that if, in fact, no amount of training can make the patterns perceptible even to the most sophisticated specialist, then the analogy with mathematics fails. For Babbitt’s compositions do not have their being in the acoustic domain to which music usually belongs.

Does that amount to an expansion of the scope of music, or to a *reductio ad absurdum* of Babbitt’s methods? There is, perhaps, a faintly discernible third possibility: namely that the domain of music is not actually circumscribed by our capacity to hear patterns, nor by the emotional responses typically evoked by the acoustic domain. Whatever the merits of this view of music, it provides a model for an alternative perspective on the objective correlates—the potential truth-makers—of emotions in general. Recall Tymoczko’s objection to regarding as music what cannot

be appreciated by the ear, however well trained: If 'the relationships are out there, in the objective world, but we cannot apprehend them', they must then cease to count as music. But precisely, Babbitt might respond, what matters is that the relationships are objectively there. And if they are continuous with patterns that can be heard as well as apprehended in other ways, that boosts their claim to be regarded as objective. So much follows from the principle of multi-modality as a touchstone of objectivity.

The retort would have some force; and the thought generalizes to non-musical emotions. If the values apprehended by emotions are objective, we might expect that they are not exhausted by actual emotional responses. This needn't commit one to the existence of a Platonic world of values which our emotions apprehend only dimly; but it does evoke the possibility that, just as Babbitt's music might be appreciated on paper, in the spirit in which a mathematician apprehends a proof, even by those whose auditory capacities are not up to hearing them, so one might, by a non-emotional process of ratiocination, apprehend values inaccessible to the emotional capacities of people at some given stage of personal, social or biological development.

On this view, something like emotional experimentation may, by analogy to musical experimentation, enlarge the domain of values to which we have access. But while the domain of values is not independent of the facts about conscious beings, it is neither simply projected from, nor ever exhausted by, the actual repertoire of human emotions—any more than all possible thoughts can be exhausted by the repertoire of actual humans thoughts past, present, or to come.

V

A Test Case: Huckleberry Finn. Several philosophers have discussed Huck Finn's decision to give up on morality and take up wickedness by stealing Jim out of slavery. Everyone agrees that it is Huck's emotions, as opposed to his explicit moral principles, which produce the true answer. In Huck, the two faces of emotion merge: his authentic emotion is also the true one. It corresponds to objective values which he apprehends, despite his conviction that he is doing wrong.
But there is some dispute about how to best describe the case. McIntyre (1990) argues that the main lesson of the story is that if akrasia is defined as a conflict between what one believes to be one's best reasons, all things considered and the real reason on the basis of which one acts, an akratic action may be entirely rational, because one may be mistaken about one's own best reasons. She contends that Jonathan Bennett (1974) wrongly sees Huck as irrational because he characterizes 'Huck's dilemma as one in which general moral principles and reasons conflict with 'unreasoned emotional pulls'.'11 Rather, she suggests, Huck need not be viewed as irrational even while he is akratic. For what he does is inconsistent not with his actual values but only with what he falsely believes to be his values.12 Her point is not merely that some objective reasons might exist to justify the akratic action, but that the so-called akratic might, after all, be doing the right thing from her own point of view. Rationality in action is 'evaluative consistency,' and that could be attained even if the action were akratic in the sense just defined:

> Evaluative consistency may exist, for example, in view of the fact that if the agent had had more time to reflect, she would have changed her mind about what the best thing to do would be. Thus she would have been saved from akrasia not by changing her behaviour but by changing her evaluation of it.13

These considerations bring us back to the problem of determining what constitutes emotional 'reflective equilibrium'. Three tentative morals may be drawn:

First, it appears to be neither necessary nor sufficient that the various emotions participating in the weighing in search of equilibrium be conscious.

Second, despite the fact that standards of contrariety for emotions are, as we have seen, obscure, it is principally emotions themselves, and not propositions, which are weighed against one another in the quest for reflective equilibrium.

Third, in the case of Huck Finn, the veracity of an emotion is hard to disentangle from its authenticity. We touch here on what

---

12. McIntyre, p. 386.
13. Ibid.
I called at the outset the inward-looking face of emotion. The sense of emotional truth I have sought to articulate is one which posits a correspondence between the emotion, characterized by a specific formal object, and some property of the human-inhabited world. But the values apprehended by emotions depend in part on who we are. They are no less objective for that; but what reflects my own individual nature—that makes for my emotional authenticity—therefore comes to seem, after all, potentially relevant to the objective world of value.

VI

Species of Truth: Emotional and Propositional. I began by advocating an extension of our notion of truth as correspondence, based on a core intuition that can be summed up in the slogan: a story defines its truth, but whether it is true can never be part of the story. (A corollary notoriously dooms the ontological argument: whether a thing exists cannot be part of its nature.) This required, in effect, the satisfaction of three conditions, which I conclude that at least some emotions are able to meet.

(1) Emotions are subject to a norm defined by their formal objects: what I fear must be dangerous; she of whom I am jealous must figure in a certain sort of triangle; what angers me must be a wrong.

(2) The norm in question is determined by the emotion itself. This is often manifest in the fact that there is an air of tautology about the characterization of the formal object: he whom I love must be lovable; what I regret must be regrettable.

(3) But the appearance of tautology is misleading, because the attainment of success for emotions—the actual fit between the object or target of the emotion and its formal object—depends on a vast holistic network of factors which transcends my actual response.

If emotions are properly said to be truth-valued in a generic sense, then the narrower class of truth-bearers traditionally targeted by philosophy—propositions or belief-like attitudes—no longer need to be regarded as the paradigm truth-valued attitudes. They form only a subclass of truth-valued states, a special case. What then is the difference that sets them apart?
EMOTIONAL TRUTH

My hunch about the answer is this: belief is digital; the representations involved in Truth's broader domain are analogue. A digital representation is necessarily part of a system of representation, and can function only once all possible signals are assigned to a finite set of discrete symbols. An analogue representation, by contrast, admits of varying degrees of precision and an indefinitely large set of possible symbols.

This hypothesis suggests another way in which emotion resembles perception. For while we can sometimes perceive that some proposition holds, in other cases the content of direct perception seems typically to be analogue. Furthermore it allays three worries raised by the notion of literal truth for emotions.

1. Generic truth legitimizes talk of more or less, by incorporating analogue correspondence, which can be more or less exact. But as traditionally conceived, truth admits of no gradations. A proposition is either true or false: tertium non datur. While this is rejected by advocates of ontological vagueness or fuzzy logic, it can be seen to apply at most only to a species of truth-bearer. Within a narrower domain of digitized representation, there are no degrees. So we get the kind of on/off truth we associate with well-defined propositions.

2. Digitality is not necessarily conventional, as shown by the example of the digital system embodied in DNA. But insofar as digital representation exists only in the context of a system of discrete values suited to indefinite copying, most digital systems are likely to be conventional. This should lead us to expect that the typical examples of truth in the sense precluding degrees will be bound to the conventional medium of language.

3. Digital representation is essential to secure fidelity in multiple reproductions of a stable 'message'. Insofar as they constitute a medium of social interaction, emotions tend to cluster into a limited repertoire of distinct entities, functioning as justifications and motivation for behaviour, regimented and 'digitized' as a system of limited significant types. But in the rich variety of their experienced reality, the significance of emotions is not

14. Robert Nozick made a very similar suggestion in Nozick 1989, p. 93: 'Emotions provide a kind of picture of value ... an analog representation of it.'
limited to their role in influencing behaviour and social interaction. Regarded as experiences representing something outside themselves, their variety instantiates a limitless continuum, and they have no need to take on the digital character of propositions.

4. That standard truth-bearers are digital representations helps to explain the grain of truth in the often expressed anxiety about the distortion of reality introduced by abstractions. Abstraction is, by definition, a process of pruning details, of ignoring certain distinctions and aspects of reality. Since all thought requires abstraction, all thought is risky. None escapes the danger that the most important aspect of reality for present purposes is precisely that which our abstractions have left out. A vague aspiration to the 'whole truth', which no utterance can contain, lies behind Nietzsche's charge that (propositional) 'truths are illusions which we have forgotten are illusions'. The whole truth is an impossible ideal: but it usefully evokes an analogue conception of representation. It has exactly the absurdity of a map on a scale of one inch to the inch, in which every nanometer is faithfully represented to scale. If emotions are conceived as analogue representations of an axiological landscape, it may come to seem natural that they should admit of variable degrees of definition, instantiating a concept of accuracy that merges with generic truth.

REFERENCES


EMOTIONAL TRUTH

by Ronald de Sousa and Adam Morton

II—Adam Morton

EMOTIONAL ACCURACY

ABSTRACT  It is accuracy rather than truth itself that is valuable. Emotional truth is a dubious though attractive notion, but emotional accuracy is much easier to make sense of. My approach to accuracy goes via an account of what makes a story accurate. Stories can be accurate but not true, and emotions can be accurate whether or not they are true. The capacity for emotional accuracy, for emotions that fit a person’s situation, is an aspect of emotional intelligence, which is as important an aspect of rational human agency as the intelligent formation of beliefs and desires.

I

Cheap Truth. Truth comes in many forms, some cheap and some valuable. Distinguish two dimensions of cheapness. One dimension extends in the direction of vagueness, indefiniteness and generality. If I claim that some flowers are coloured, or that music is sometimes nice, what I say is true, but cheaply so. Another dimension extends in the direction of the range of attitudes that can be counted as true. Truth can be extended from assertive sentences to beliefs to questions and requests at very little price. When a person attitudes that p, and p, we can count her attitude as true. So a Yes-No question is true if the answer is yes; a desire is true when it is satisfied. And we can say that all Jane’s desires on Tuesday were satisfied, which would be equivalent to ‘if on Tuesday Jane wanted cats to fly, then cats flew, and if she wanted 34 + 76 to be 994, then 34 + 76 = 994, and ...’. Similarly, we can say that ‘Hamlet killed Polonius is true’ iff Hamlet killed Polonius, and that ‘e\(^{\pi i} = -1\)’ is true iff e\(^{\pi i} = -1\), without worrying about where in the world to find Hamlet, Polonius, and imaginary numbers. None of this is very demanding; the conceptual price is low, as the minimalist literature shows.¹

¹  See Williams 1976, Horwich 1998.
Emotional truth is easily achieved if one wanders far enough out along these dimensions. My fear that the dog will bite me is true if and only if the dog will bite me. My elation that life has many joys and my depression that life is a grim business are both true since life is a grim business with many joys. But there's no philosophical pride to be had from bringing home these trophies; any child with a butterfly net could have gone out and got them.

Now to the more valuable kinds. The opposite of vagueness is precision, and precision combined with truth gives accuracy. Accuracy certainly adds value to truth. For one thing it allows non-perverse speculation: the difference between scientific cosmology and metaphysical rambling is that cosmology distinguishes between finely differentiated hypotheses—whether fundamental constants have this value or this slightly different one—and tries to distinguish the different consequences they would have. And on the other dimension, the opposite of minimalist content-matching is to insist on a world-to-mind direction of fit in which determinate aspects of the state have to match determinate aspects of the world. (A substantive theory of truth—correspondence, as I'm slanting it—thus aims not at telling us what propositions are to count as true, but what kinds of truth to count as valuable, a point ignored by Lewis 2001.)

Emotional truth that had these value-adding features would be something to aim for. There would be a point to directing the evolution of our emotional states towards it, just as there is a point to directing the evolution of our beliefs towards the more valuable, but only the more valuable, forms of truth. Analogous to the way precision in theory allows responsible speculation, precision in emotion allows responsible intensity. If you have the exact emotion for the situation, then you can feel it whole-heartedly, without the danger of inappropriate blundering. A bull that dances through the china shop. And analogous to the world-to-mind fit of beliefs would be some notion of an emotion that is demanded by the situation. Elation where elation is right, depression or anger where that is right, whether or not the person has grounds to motivate their feeling this right thing.

These remarks are meant to elicit sympathy for de Sousa's project. To the extent that we have a grasp of the right emotion for a situation, the objectively right emotion, we can see analogues in emotion of the valuable features of true belief. But they
are also meant to insinuate a doubt. The intuitions are linked not to the core idea of truth itself but to the value-adding aspects that make it worth having. In this paper I shall argue that some of these aspects are independent of the core. We can make sense of emotional accuracy without having to make sense of emotional truth, at least not in more than the cheap and easy way just described. Some of the consequences of accuracy-without-truth, though, are in many ways like those that de Sousa wants from emotional truth.

II

Accuracy Without Truth. Consider two stories:

Story 1: A carriage rolled north down Baker Street through a thick London fog on a cold December day in 1887. As it came to Marylebone Road the passenger rapped on the driver’s window and asked to be let out. Only the most acute of observers would have recognized the crippled Crimean war veteran who emerged as the famous detective Sherlock Holmes.

Story 2: A boat drew slowly along the Baker Street canal in the balmy weather of London in the winter of 1887. As it joined the Thames a passenger leapt to the bank. That person continued his journey on foot.

Neither story is true. Possibly neither is false. But the first is in two respects more accurate than the second. Baker Street does not have a canal, and even if it did it would not reach the Thames. The winter of 1887 was not balmy. That is the first accuracy, fit: the first story fits the world as it is, even though it does not say anything true about it. The first story is also detailed in a way that the second is not: it gives a specific name to its protagonist, and describes his appearance. Though both stories can be matched with many non-actual worlds, the first applies to fewer than the second does: it is more restrictive. (We are probably speaking of infinitely many worlds in both cases, so ‘fewer’ is problematic. It would be best to consider cases where one story’s worlds are a subset of those of another. But that would require four stories rather than the two I used.)

The two aspects interact. Detail allows fit. If a story has enough details that can be taken as true of an actual situation then it will fit it. Fit selects detail. If a story is taken as fitting
a particular situation then we can assess the detailedness of its description of that situation. This suggests a tentative definition of accuracy. One story is a more accurate depiction than another of an actual situation when there are more elements of the one that are true descriptions of aspects of the actual situation than there are of the other. (One story might be taken to be intrinsically more accurate than another when there is an actual situation such that there are more elements of the one that are true descriptions of aspects of that situation than there are elements of the other that are true descriptions of any actual situation.) That will do for now; the definition is not meant to be taken very carefully. (It surely will not survive rough handling: taking stories as closed under logical consequence and then literally counting true sentences, etc.)

Some think that stories are true of worlds, and thus simply true when they are true of the actual world. I do not want to get into this question. The important point is that one not-true story can be more accurate than another. Science fiction is not very accurate, at any rate not accurate about the technological possibilities (or even usually the laws of nature) of the present actual world. Cowboy fiction is said to give a very inaccurate impression of life in the Wild West. Zola or Hardy probably do give relatively accurate reports of life in the times and places they discuss. But none of these stories are true. In fact, a story can have a good measure of accuracy while lacking not only truth but also possibility. Kurt Vonnegut's Cat's Cradle is full of historical, sociological, and emotional accuracy while describing something that just can’t happen.

Accuracy as just described seems to presuppose truth. An accurate story has many elements that are true descriptions of an actual situation. But a more careful formulation takes care of this. A story can be taken as describing a situation no elements of which does it actually name. For example a story might begin 'The general had accumulated many powers, so many that concerned citizens plotted to assassinate him.' It might be taken as describing events in Rome in the first century BCE, or in many other times and places. But no element of it is simply true. Conversely a Jonathan Miller type production of 'Julius Caesar' might add enough detail that—incorporating all elements of the production into the story—it was an accurate portrayal of Tony
Blair and his entourage. The assassination itself would then be a non-descriptive detail that gained significance from its links to the descriptively accurate elements. Neither accuracy nor truth simply presupposes the other.

III

*Emotional Accuracy.* What does this have to do with emotions? The essential link is that a person’s emotions involve representations, and these representations can be more or less accurate depictions of her situation. Contrast two classes of cases.

(1a) An engineer is laid off by her company. She realizes that the economic climate is not good for getting another job of the same kind, feels relieved that she does not have to face more boring programming disguised as design, and goes back to university to do a MBA.

(1b) An engineer is laid off by her company. She takes this as showing that she has neither the technical nor the personal skills for success in a demanding profession, becomes very unhappy, and does not look for another job.

(1c) An engineer is laid off by her company. She reflects on the less competent and less hard working colleagues who have kept their jobs and of the lack of respect her boss has always shown to her. She gets very angry, goes into his office and pours a cup of coffee over his head.

(2a) An engineer is laid off by her company. She thinks of all the desired things that will now never happen and is overcome with sorrow. She becomes very unhappy at the fate of abandoned animals, and cries whenever she sees a dog walking without a leash, or a non-fat cat.

(2b) An engineer is laid off by her company. The next day she finds her mind is full of confusing thoughts moving in all directions. There is something exciting about the confusion and she develops an enthusiasm for the company. She starts a web site on which satisfied customers and grateful employees can register their good feelings.
(2c) An engineer is laid off by her company. Feelings of anger rise up in her and she directs them at American policy in the middle east. She becomes a fervent campaigner for the internationalization of Jerusalem.

The cases under 1 resemble story 1 in a way that the cases under 2 resemble story 2. That will only be true of some ways of imagining the cases, filling in the details. But they are the natural ones, the ones that would first occur to one. Suppose that we have a detailed filling in of one of these cases, including on the one hand the engineer’s beliefs, intentions, fears, and desires, and on the other hand her whole physical state and the state of her environment, the sensations she experiences, and the basic acts she performs. Call the first of these ‘the emotion-story’, and the second ‘the situation’. Then the emotion-stories of the (1) cases are more accurate depictions of the situation than the emotion-stories of the (2) cases. More of the facts are accurately represented in these stories.

Consider (1a). The engineer’s emotion is one of relief and redirected interest. These emotions are directed at specific aspects of her situation and do not make sense without them. They involve (or require, or even consist in) beliefs about the character of her work before she was laid off, beliefs about the character it would have assumed had she been one of those not laid off, desires to do one kind of work rather than another, intentions to act in one way rather than another, and so on. Contrast this with (2a). The engineer’s emotion is one of sorrow directed at the plight of animals. But, at least on one natural way of filling in the details, there are no specific episodes of animal suffering that give detail and specificity to the emotion: many associated beliefs are not true of the engineer’s life, and many associated desires do not lead to successful acts. (They’re not true desires, in the cheap way of speaking I suggested above.) This is generally true of natural ways of imagining the (2) cases: they do not latch onto actual features of the situation as it is. In fact, in order to imagine oneself into the situation of the engineer in the (2) cases one has to imagine her misconstruing and misrepresenting what is going on and what the connections between events are. This is much less so in the case of the (1) cases. The emotions there not only are sustainable in the face of an accurate grasp of the facts
and possibilities, they build on a network of representations of the details of the person’s situation.

I am trying not to put this in an overly cognitive way. On a 1970s-type account the emotion just is a complex of states essential members of which are propositional attitudes, which in accurate cases have true propositions as their objects. I take it that a number of writers, notably de Sousa and Greenspan, have shown us more plausible ways of recognizing that thinking is essential to emotion without turning emotions into thoughts. Without taking on the details of any of these accounts I shall assume that when one is in an emotional state there are patterns of belief and belief change, desire and desire change, and characteristic intentions, that are essential to ones being in that state rather than another. If a person is afraid then there is a pressure towards thinking that some things are dangerous, and a tendency towards wanting to avoid or escape some things, whether or not she succumbs to the pressure or goes along with the tendency. This is enough to make what I have called the emotion-story essential to the emotion, and thus to give the emotion an intrinsic degree of accuracy as a depiction of a person’s situation.

I said that accurate emotions are sustainable in the face of an accurate grasp of the facts and possibilities. Why possibilities? Consider someone who takes as fearful something that cannot hurt him, or who greets with joy something that cannot do him any good. The emotions don’t fit the situation not because the object will not harm or will not help; after all, it is appropriate to be afraid of a rattlesnake that in fact does not bite one. The lack of fit comes because something is thought to be capable of what it is not. More generally, an emotion can be inaccurate because it misrepresents the possibilities of the whole situation. Most emotions are action-guiding, taking action in a very general way to include strategies of thought. (This is a central idea in most of the papers in Goldie 2001.) They will not serve this role if they are unhinged from the actual situation of the agent; and they will not serve it if they do not respect what is actually possible and impossible, in fact what possibilities are more or less remote. So an accurate emotion must not only contain detailed representations that fit the actual situation; it must represent that

actual situation as rightly situated in the galaxy of could-have-beens and would-have-ifs around it.

This might seem to distinguish accuracy of emotion from accuracy of belief. I think it does not, though. A belief can be inaccurate though true in a detailed way of the actual world. Consider for example a rich and complete system of unnatural Goodmanian concepts, cutting across natural kind boundaries in weird and peculiar ways, and consider beliefs expressed in terms of them. The belief that all emerats are granimals is true (emerats are emeralds that come to human notice before 1 Jan 3000 or otherwise rats, and granimals are green things noticed before that date or otherwise animals). But it misrepresents what emeralds and rats are like and taken together with other similar beliefs would misrepresent what is possible for them. So respect for how a situation is situated among its possible variants is something we should write into a better definition of the accuracy of belief, too, taking accuracy even further from truth.

Accurate emotions are not well described as true. After all, the analogy is with an accurate story, and many very accurate stories are not true. The difference shows up in the non-uniqueness of accuracy. All of (1a)-(1c) are accurate, accurate to the same facts about the engineer’s life. I see no reason to think that any one has to be more accurate than the others; each could invoke as rich a body of beliefs and desires, fitting the person’s situation and its possibilities as well as each other one. (That is why (1b) is included: emotions that we think of as less wise or less admirable may still be accurate. But see Section IV below.)

Another way of putting it. An accurate emotion is like a rich myth, deeply engaged with the details of some aspect of the world. A less accurate emotion is like a shallow or artificial myth, a Walt Disney substitute, which tries to depict mythical events that bear no detailed relation to what actually happens in people’s lives. Or, the accuracy of an emotion is like the observational accuracy of a scientific theory, which can capture actual and potential observations more or less well. Theoretical and observational assertions and concepts can be intimately connected; neither may be intelligible without the other, and yet it is clearly true that observational accuracy does not guarantee truth. There can be rival equally observationally accurate theories, relative to any way of drawing the somewhat arbitrary line between observation and theory. And among non-true theories some will be
more accurate observationally than others. For some purposes, e.g. navigation or bridge construction, observational accuracy will be more important than truth. We want a rich and reliable body of connections with the ways the world impinges on us. So too with emotions: among the variety of attitudes we could take to the situations we find ourselves in, we want those that give a rich and reliable set of connections to guide our further acting and feeling.

IV

Emotional Intelligence. My main point has been that among the emotions a person can direct at a situation some fit it better than others. The point can be extended: among the varieties of anger, or of sadness or exhilaration, that a person can direct at a situation some fit it better than others. So accuracy cuts across our usual classifications for the emotions. You can be miserable, elated, or curious, and be so in a way that does or does not accurately represent your situation. No emotion is intrinsically accurate. But some distinctions between emotions are necessary for a creature that is to have accurate emotions. Sadness must be distinct from depression; remorse, guilt, shame, and embarrassment must be kept apart. Falling into one of these when another fits the situation is a sure route to emotional mess. And finding one’s way around a rich range of emotions is as demanding as finding one’s way around a complex set of beliefs. It requires a special and admirable quality that it makes sense to call emotional intelligence.

Emotional intelligence will not always result in emotional accuracy, any more than theoretical intelligence will always result in true belief. And just as truth bears a complex relation to the coherence of belief, emotional accuracy bears an equally complex relation to the coherence of emotions, with one another and with a person’s complex of beliefs and desires. Sometimes the more accurate emotions a person can have will not cohere well with one another or with the person’s other states. This will typically be when the other states are defective, or when the situation is so complex that the person is not capable of coherent attitudes that represent it well. (But then, the universe is like that, compared to our little brains.) And, to pile on the warnings, there is
no more guarantee that emotional accuracy will give us better
lives than there is that we will be happier if we have true beliefs.
To the perspicacious tyrant who kills you if you don’t believe he
is charming there corresponds the situation that is so unbearable
that your sanity will not permit you to react to the way it really
is. But, we all trust, these are aberrant outlying cases. In general,
the route to truth leads through evidence and results in satisfied
desire, and the route to emotional accuracy leads through the
acquisition of a range of possible feelings and attitudes and the
capacity to discriminate between them, and results in the har-
mony of thinking and feeling. More specifically, it tends to link
the evolution of our beliefs, our desires, and our feelings, and
allows the present state of each of these to put pressure on the
others. It allows us to be whole people, by having patterns of
thought that make two-way connections between what we believe
and what we feel. (Some of the connections in one way are clear:
when you discover the insect is harmless your fear should change.
The connections the other way must consist in part of your
emotions helping select relevance of evidence and direction of
thought. If you feel instinctively afraid of the insect you look for
reasons, both in what you can see around you and in what you
know, which might settle the question of its dangerousness.)

Imagine then a progression. It starts with our hard-wired
emotional responses, with their fixed affects and their simple
paradigm scenarios. Emotional intelligence then intervenes, and
we acquire the capacity to modulate our emotions to what we
learn and what we come to want. (At the beginning we feel dis-
may at a situation; at a later stage we anticipate regret for the
action we are choosing; at a yet later stage we anticipate regret
if we take one choice and remorse if we take the other.) Suppose
that the capacity were perfectly acquired. Then our emotions
would match our situations to the extent that our information
about then was accurate. Could they then be counted as
emotional truths? The main factor to consider is the way they
exclude one another. At the original primitive stage fear, say,
and delight are mutually exclusive. And the exclusion is not just
the effect of quirks and limitations: it is intrinsic to a simple fear
that it leads one to intend avoidance and to a simple delight that
it leads one to intend contact. They are emotions that cannot
both be held, though we can oscillate between them. But each
might be equally accurate. As de Sousa, following Tappolet (2000), would put it, the values of danger and of attractiveness are both present. So we shouldn’t count them as truths. (It would be a strange kind of truth, such that having it committed one also to falsity. To fear is to take as not attractive.) But at later stages the exclusion lessens. We acquire more subtle emotions, such as a delighted horror. (You see the notorious association between sophistication and perversity.) Then it is possible to acknowledge that the situation is both dangerous and attractive. So as our emotions become more and more refined they come to be capable of representing more and more of the values present in our situations, in such a way that to acknowledge one is not to reject another.

Might there be an ideal end to this progression, where in any situation an agent could have emotions which accurately represent it, and which do not exclude any others that accurately represent it? I have no idea. I do fear that these kinds of heroically accurate emotions would have become so much like beliefs that they could not easily serve the functions of emotions. After all, as Greenspan and earlier work by de Sousa taught us, emotions are essential for defining patterns of salience that create pressures on the evolution of our beliefs and desires. These patterns are essentially selective; they make things possible for us by limiting the possibilities. But perhaps creatures with sufficient emotional intelligence would be able to assume these deliberately limiting perspectives while also remaining open to alternatives. Perhaps. We don’t have to take a position on this, in order to conclude that there is such a thing as emotional accuracy, that it is valuable, and that intelligent thinking and feeling aims at it.

REFERENCES