DEBATES /*in* AESTHETICS

BSA   BRITISH
SOCIETY OF
AESTHETICS

ART & THE EMOTIONS:
Debating the work of Jenefer Robinson

Ryan P. Doran
University of Cambridge

Shelby Moser
Azusa Pacific University

Jenefer Robinson's work has been motivated by two central concerns: Firstly, to characterise the nature of the emotions, and second, to provide an account of how the emotions are involved in various aspects of art appreciation. In this issue, some of her critics attempt to put pressure on, or develop, some aspect of her influential contributions to these projects in their respective articles, and Robinson replies to each in turn (pp. 95-121).

Vanessa Brassey (pp. 15-29) turns to examine some of Robinson's more recent views on the relationship between art and the emotions. One of Brassey's primary aims in her article is to reveal one of the distinctive advantages of Robinson's personalist account of pictorial expression over its impersonalist rivals. Brassey suggests that we should favour Robinson's personalism on the grounds that what is expressed by some paintings is not exhausted by what is expressed by the figures, scene or design of a painting. This additional expressive meaning is not, according to Brassey, "derived from the formal (visible) qualities of the picture surface"[1], but rather comes from our impression of the person executing the picture.

Brassey offers the example of Hopper's 'NightHawks', observing that the juxtaposition of the mask-like faces with the

---

[1] Brassey 2019, p. 23.

sophisticated air of the picture overall results in the viewer experiencing a "richness and elusiveness of style" (following a discussion of Klee by Kendall Walton).[2] Brassey suggests that this experience of richness is dependent on having the impression that it is executed by a skilled painter adopting a naïve style rather than a "naïve illustrator".[3]

This idea is heartily embraced by Robinson in her reply (pp. 95-101), but one might have a couple of worries here. First, one might wonder whether experiencing "a richness and elusiveness of style" is indeed a case of expression. Even if this is granted, one might also worry that a personalist account of pictorial expression isn't needed to accommodate such cases. The impression of mystery in 'NightHawks' seems connected to features that are internal to the work, just as impersonalists such as Lopes allow: the sense of mystery seems to merely be generated by noting the tension between the mask-like rendering of the faces and e.g. the clean depiction of the streetscape. Why do we need an implied persona to account for interactions such as this?

Of course, on Lopes' particular variety of impersonalism, there may not be room for such cases. In her reply to Brassey, Robinson locates the problem Brassey identifies in her description of what a painting conveys overall as the 'design expression' (using Lopes' taxonomy) in the narrowness of Lopes' characterisation of this concept. The same problem seems to crop up here. Since the design expression is for Lopes "expression that is wholly attributable to a picture's design or surface and not to any figure or scene it depicts"[4], it cannot accommodate interactions such as that between the mask-like appearances of the figures and the depiction of the streetscape. As a result, perhaps a fourth kind of expression of the kind that Robinson outlines in her reply and which we might term (to use a

---

[2] Walton 1976, p. 52.

[3] Brassey 2019, p. 23.

[4] Lopes 2005, p. 57.

cue from Robinson) 'compositional expression' is needed by impersonalists.

Brassey also expresses some doubt about Robinson's claim that "the creation of expressive content should be understood as a transitive act of expression" on the grounds that the implied painter is a psychological extension of the painter.[5] Brassey suggests that this implies that painters cannot create implied personae that are distinct from themselves, and deploys the counterexample of Tolstoy: a writer who was able to express compassion in his work while lacking compassion in life.

While acknowledging the force of this point, Robinson responds by implying that the implied persona has *enough* in common with the painter to express the painter's state of mind. But this seems to be precisely what Brassey is denying: a novelist might *merely entertain* e.g. what it is like to be cruel and indifferent in the process of creating passages or works that are expressive of these, even if the novelist is herself only capable of being kind and concerned.

An alternative response which is not suggested by Brassey or pursued by Robinson here but is arguably latent in Robinson's work on pictorial expression would be to argue that such cases are of a piece with Calvin Klein adverts. These cases might best be regarded as *technê*, and not of transitive, or indeed, pictorial expression. According to this thought, such artworks would be mimicking expression, rather than truly expressing. The question then, however, would be whether this response would be convincing, given that such works are *prima facie* expressive?

Mary Beth Willard (pp. 30-44) engages with Robinson's answer to the paradox of fiction. Following Gendler and Kovakovich (2006), Willard formulates the paradox as the mutual incompatibility of three claims: (i) we have genuine and rational emotional

---

[5] Brassey 2019, p. 17.

responses to fictions; (ii) we believe that fictional entities don't exist; and (iii) we must believe that characters exist in order to have genuine and rational emotions towards them. Willard implies that Robinson, as an advocate of an embodied account of the emotions, rejects premise (iii): since emotions involve non-cognitive appraisals—which are not sensitive to the ontological status of their objects—we can have emotions in response to fictions.

One might worry that this does not quite capture the contours of Robinson's position. As Willard notes, and Robinson affirms in her response, Robinson regards our emotional responses to fictions as *irrational*. That is, Robinson seems to reject (i) and (iii). One problem here lies in the way the paradox has been formulated. Premises (i) and (ii) feature a conjunction of two distinct properties—genuineness *and* rationality. A gold ring can be genuine in the sense that it is actually gold and not, say, coated steel, but it is not thereby rational. It doesn't seem as though we can accurately capture Robinson's view on the paradox of fiction without being sensitive to this distinction: Robinson believes that we have *genuine* emotions in response to fictions in the sense that they're *actually* emotions (i.e. sadness in response to a fiction is the same kind of thing as sadness in life[6]), but she doesn't believe those emotions are *rational*. Why aren't they rational for Robinson? Because both the content of the affective appraisal and the cognitive monitoring system which together constitute emotions for Robinson represent states of affairs that can or cannot truly exist. The content of the affective appraisals and cognitive monitoring involved in episodes of

---

[6] Gendler and Kovakovich (2006) intend genuineness in Walton's (1990) sense of lacking motivational force and existential commitment. Setting the latter aside for reasons that Gendler and Kovakovich specify (2006, p. 249), the general idea here is that the emotions had in response to fictions are not the same—functionally or otherwise—to emotions had in response to their real-life counterparts. Robinson denies this.

fear is something like DANGER[7] (albeit non-conceptually in the case of the affective appraisal). And states of affairs can actually be dangerous or not. Sitting in a horror film is not dangerous but being pursued by someone with murderous intent is.

Setting this aside, Willard claims that Robinson *should not* say that our emotional responses to fiction are irrational on the grounds that the characters and situations presented are fictional. Indeed, Willard suggests that our emotional responses, when appropriate according to the relevant norms, are rational. It seems important to clearly locate where the disagreement between Willard and Robinson lies, if indeed any exists. Robinson (pp. 102-107) for her part claims that our responses to fiction are irrational in just the sense spelled out above. If truth is the standard by which our activities are measured, then emotional engagement with fictions is not rational: we represent objects which are not e.g. dangerous as if they were. But, setting this kind of irrationality aside, Robinson notes that there are other ways in which our emotional responses to fictional objects are rational. They may, for example, be rational in the sense that they are adaptive.

This is not the grounds for rationality that Willard has in mind though. Willard's sense of rationality is prudential *and* internal to the activity of art appreciation: it is rational to emote given our aim to appreciate artworks on their own terms.

Robinson welcomes this suggestion but disagrees with Willard's specific proposal for understanding rationality here. There is surely further work to be done on this, as Robinson and Willard both note. But to gesture towards one of the directions this might take in the future: Robinson's worry that Willard's second criterion is tautological may be able to be assuaged by understanding the criterion as governing when a response is *rational* in the relevant

---

[7] It is in virtue of the fact that they are about the same content that the cognitive monitoring system is *relevant to* the operation of the quick-and-dirty components.

sense (rather than appropriate), a perspective which is more in line with the general spirit of Willard's article in any case.

Quixote Vassilakis (pp. 60-73) puts pressure on Robinson's claim that the emotions are necessary to understanding certain works of art, and indeed suggests that the *absence* of emotional responses may actually enhance understanding of those works. In arguing for these claims, Vassilakis points out that psychopaths have emotional deficits but may have an enhanced ability to read the emotions of others, and that there is evidence that the emotions can lead us astray in forming literary interpretations.

In responding to the former argument, Robinson (pp. 111-114) makes a number of points: that psychopaths are not a relevant population from which to draw principles about how to read a novel; psychopaths can infer emotions but not how they feel; and psychopaths don't care. These responses suggest that Robinson and Vassilakis may be operating with different concepts. On the one hand, we might intend UNDERSTANDING in the strict sense of merely comprehending. We deploy this concept when we e.g. ask whether somebody understands the solution to a maths problem. It is this sense of understanding that Vassilakis seems to be targeting. On the other hand, we sometimes refer to understanding in another sense, namely to comprehend *and care* in the right way. We deploy this kind of understanding when we e.g. ask whether someone was understanding about a mishap caused by adverse circumstances. It is this sense of understanding that Robinson seems to intend. This distinction seems crucial, as emotional responses might be necessary for understanding in one sense but not the other.[8]

---

[8] In addition to operating with different notions of understanding, Vassilakis and Robinson seem to interpret the nature of the claim that the emotions are necessary for understanding some works in different ways. Vassilakis seems to interpret Robinson as defending a strong metaphysical claim. But in her reply, Robinson seems to suggest that she intends a weaker, prudential claim concerning "principles about how to read a novel in a normal way" (p. 113).

As Vassilakis presses, it may be the case that someone with emotional deficits could understand perfectly well what is going on in a novel, without having any concern for the characters that feature therein. But it is far more difficult to see how understanding in the second sense can be done without emoting. Indeed, to the extent that being sympathetic is an emotional state, it seems that the emotions may be logically built into this sense of understanding.

One might not be persuaded by Robinson's worry that the conclusions that can be drawn from psychopaths about the necessity of emotions for understanding are limited because psychopaths come from a small pathological population. If psychopaths demonstrate e.g. a selective inability to feel emotions generally and they can still understand others emotional reactions[9], then they are presumably doing this with non-affective psychological capacities which are shared by the readers that Robinson has in mind.

As a result, while it may be difficult for us non-psychopaths to imagine being able to understand emotionally-relevant situations without feeling, as many of the psychological mechanisms involved in emoting are automatic and mandatory (as Robinson notes), the case of psychopaths may provide us with a precious opportunity to see what the other non-emotional mechanisms we have in our heads may be doing when we're reading great realist fiction. In this regard, the method deployed by Vassilakis is structurally similar to Robinson's appeal to Joseph LeDoux's dissociation experiments in 'Deeper than Reason'.

---

[9] This characterisation is favoured by, e.g. Vassilakis 2019 and Prinz 2011. In her recent synthesis of the evidence on the deficits that psychopaths present, Maibom (forthcoming) suggests that shallow affect is characteristic of a subset of psychopaths—sometimes called *primary* psychopaths. It is important to note that psychopaths's recognition abilities are not entirely intact: they have been shown to have a deficit in recognising emotional facial expressions, e.g. Iria & Barbosa 2009. However, it is not clear that this damages Vassilakis' point given that he and Robinson are concerned with literary understanding.

A second way of carving up the domain of understanding arises from Robinson's discussion of Konrad et al's data argument. Eva-Maria Konrad, Thomas Petaschka and Christiana Werner (pp. 44-58) suggest that we can glean, for example, that Strether is mildly amusing, and that Anna is in a desperate situation without emoting. Robinson's response is intriguing: "there is an important difference between 'understanding' in the abstract, and understanding with your gut".[10] Here again, if the relevant sense of understanding is understanding with your gut, then it is difficult to see how the emotions couldn't be necessary, especially given Robinson's non-cognitive conception of the emotions. But it is interesting to gloss this distinction in a different way, namely as between superficial and deep understanding (as referenced in the title of Robinson's (2005) monograph). One question is whether *that* distinction carves along emotional-non-emotional lines?

One way into this issue is by thinking about whether the evaluative properties discussed by Konrad et al and Robinson—such as being amusing or desperate—should be understood in terms of emotional dispositions[11] or in terms of a second-order sentimentalism.[12] For comic moralists, for example, whether a character is amusing or not is a matter of whether they warrant laughter, not whether they do in fact tend to elicit laughter. As Elisabeth Anderson notes:

> A person may laugh at a racist joke, but may be embarrassed at her laughter. Her embarrassment reflects a judgement that her amusement was not an appropriate response to the joke. The joke was

---

[10] Robinson 2019, p. 110.

[11] As Konrad et al 2019 advocate, at pp. 55-56.

[12] See, e.g. D'Arms 2005.

> not genuinely good or funny; it did not merit
> laughter.[13]

If some such kind of neo-sentimentalism about evaluative properties is right, then it might naturally be thought that understanding that it is warranted to feel the relevant emotion in a given case (rather than anyone actually feeling the relevant emotion) is what determines whether one *deeply understands* that a character or situation is amusing or desperate. And such an understanding looks to be a cognitive rather than emotional activity.[14] Indeed, given that Robinson's account of the emotions as processes includes a hierarchy of processing—from affective appraisal to cognitive monitoring—such a neo-sentimentalist position, and the corollary of such a position for the necessity of the emotions for understanding, actually look apposite to Robinson's wider view of the emotions.[15]

As such, whether the emotions are even required for understanding whether a given aspect of a work really has evaluative properties that seem to intimately involve the emotions—such as having amusing characters—may seem far from trivial.

Standing back, one of the general points that may be thought to emerge from the exchanges between Willard, Vassilakis, Konrad et al and Robinson is that Robinson's idea that emotions are two-stage processes involving non-cognitive appraisals and cognitive monitoring may, when developed in the right way, unexpectedly put pressure on some of her positions on aesthetic questions—such as whether the emotions are necessary for understanding certain works, and whether our emotional responses to fictions are irrational.

---

[13] Anderson 1993, p. 2.

[14] A similar conclusion is suggested by Vassilakis when he argues that the emotions bias us, and suggests that it is the cognitive monitoring processes that are the truth-makers when it comes to literary interpretations.

[15] They also seem consistent with the view Robinson lays out in 2010, p. 77.

Whereas Vassilakis and Konrad et al both take Robinson's account of the emotions on its own terms and attempt to challenge the strength of Robinson's claims that the emotions are central to successfully engaging with some works, Irene Martínez Marín (pp. 74-94) takes the opposite tack. She argues that Robinson's account of the emotions is not able to deftly handle higher cognitive emotions (such as nostalgia) in addition to the basic emotions. Partly as a result of this, she suggests that Robinson underestimates the ways that the emotions are involved in engaging with certain artworks.

Focusing on the case of nostalgia, Martínez Marín argues that Robinson's non-cognitive account of the emotions cannot accommodate nostalgia on the grounds that nostalgia constitutively involves reflection on "the irrevocability of one's own past"[16]—an appraisal that is too cognitive to count among the affective appraisals that Robinson claims lie at the heart of the emotions. Martínez Marín continues by suggesting that these higher cognitive emotions are central to the appreciation of some of the 'intellectual' artworks that Robinson implies do not need to be engaged with emotionally. Nor indeed, according to Martínez Marín, is Robinson's work on empathy able to be purposed to accommodate such cases.

Robinson (pp. 114-119), for her part, defends against the thrust of Martínez Marín's suggestion by questioning whether we should consider nostalgia an emotion on the grounds it is partly constituted by a set of reflections. Robinson suggests that we should regard nostalgia as a 'long-term' emotion rather than an emotion proper.

Some might worry that the considerations Robinson brings to bear to justify this deflation are not entirely decisive. It is not clear, for example, why the fact that nostalgia is partly constituted by reflections disqualifies it from being an emotion. Certainly, nostalgia seems to *prima facie* be an emotion—we commonly talk about "feeling nostalgic", and nostalgia seems to have a complex valence and phenomenology. As Robinson notes, the phenomenology of

---

[16] Martínez Marín 2019, pp.75-76.

nostalgia is more than a blend of sadness and happiness. Nor is it clear why nostalgia should be thought to be a long-term emotion. The object of the emotion is necessarily some temporal distance from us, but it is not clear why the *emotion itself* needs to be attributed to individuals for a long period of time. It seems common to feel *flashes* of nostalgia when one tastes, smells, hears, or sees something that reminds one of one's past and is regarded as irrevocable.

Notwithstanding this, even if nostalgia is an emotion (long-term or otherwise), it is not clear that the basic and higher cognitive emotions do indeed form a single natural kind[17] such that a unified account is desirable. This may be significant for Martínez Marín given that Robinson's aim in 'Deeper than Reason' is, at least in part, to characterise a set of psychological states that function in the same way, rather than make prescriptions about what we should call emotions based on an analysis of the term.

Perhaps then, as Robinson herself notes in the close of her reply, further independent work (such as that provided by Martínez Marín) is warranted on the place of states such as nostalgia in the ecology of the mind, and how they relate to works of art.

The second half of this issue focuses on issues that are germane to both the spirit and content of the first half, and indeed the guiding principles of this journal more generally. In his piece, Hans Maes (pp. 122-130) reflects on his recent book of interviews with prominent philosophers of art—including Jenefer Robinson—to trace the way that formats such as conversations can contribute to philosophical progress in the current publishing environment. In addition to noting the dialectical similarities between philosophy as it is done in journal articles and in conversation, Maes traces six ways in which conversations are uniquely placed to contribute, including

---

[17] See Griffiths 1997.

revealing tensions and unifying commitments in a given philosopher's oeuvre.

Angelo Cioffi's (pp. 131-176) interview with Murray Smith exhibits the lattermost advantage of doing philosophy-in-conversation exceptionally well. The most prominent idea to emerge from the interview is the strength and depth of Smith's commitment to the naturalistic enterprise in philosophy: Smith not only believes that philosophers should present theories that are, at a bare minimum, consistent with the relevant scientific findings, but they should preferably try to synthesise relevant findings from different fields and at different levels of explanation or even collect data themselves. In their commitment to respect the findings of the cognitive sciences, Smith and Robinson, and indeed some of Robinson's critics here, share a common perspective.

rd592@cam.ac.uk

slmoser@apu.edu

References

Anderson, Elizabeth. (1993). *Value in Ethics and Economics*. Cambridge, MA: Harvard University Press.

Brassey, Vanessa. (2019). "The Implied Painter". *Debates in Aesthetics* 14(1), pp. 15-29.

Cioffi, Angelo, & Smith, Murray. (2019). "Aesthetics Naturalised: An Interview with Murray Smith". *Debates in Aesthetics* 14(1), pp. 131-176.

D'Arms, Justin. (2005). "Two Arguments for Sentimentalism". *Philosophical Issues* 15, pp. 1-21.

Gendler, Tamar, & Kovakovich, Karson. (2006). "Genuine Rational Fictional Emotions". In: *Contemporary Debates in Aesthetics and the Philosophy of Art*. Matthew Kieran (Ed.). Oxford: Blackwell Publishing, pp. 241-253.

Griffiths, Paul. (1997). *What Emotions Really Are: The Problem of Psychological Categories*. Chicago, IL: Chicago University Press.

Iria, Catarina. & Barbosa, Fernando. (2009). "Perception of Facial Expressions of Fear: Comparative Research with Criminal and Non-criminal Psychopaths". *Journal of Forensic Psychiatry & Psychology* 20, 66-73.

Konrad, Eva-Maria, Petraschka, Thomas, & Werner, Christiana. (2019). "Are Emotional Responses Necessary for An Adequate Understanding of Literary Texts?". *Debates in Aesthetics* 14(1), pp. 45-59.

Lopes, Dominic McIver. (2005). *Sight and Sensibility: Evaluating Pictures*. New York, US: Oxford University Press.

Maes, Hans. (2019). "You Talking to Me?". *Debates in Aesthetics* 14(1), pp. 121-128.

Maibom, Heidi. (Forthcoming). "Responsibility in the Age of Neuroscience: Moral Understanding and Empathy in Psychopathy". In: Oxford Handbook of Moral Psychology, John Dorris & Manuel Vargas (Eds.). Oxford: Oxford University Press.

Martínez Marín, Irene. (2019). "Robinson & Self-Conscious Emotions: Appreciation Beyond (Fellow) Feeling". *Debates in Aesthetics* 14(1), pp. 74-94.

Prinz, Jesse. (2011). "Is Empathy Necessary for Morality?". In: *Empathy: Philosophical and Psychological Perspectives*. Amy Coplan & Peter Goldie (Eds.). Oxford: Oxford University Press, pp. 212-29.

Robinson, Jenefer. (2005). *Deeper than Reason: Emotion and its Role in Literature, Music, and Art*. Oxford: Clarendon Press.

 —. (2010). "Emotion and the Understanding of Narrative". In: *A Companion to the Philosophy of Literature*, Garry L. Hagberg & Walter Jost (Eds.). Malden: Blackwell, pp. 71-92.

—. (2019). "Response to Critics". *Debates in Aesthetics* 14(1), pp. 95-121.

Vassilakis, Quixote. (2019). "Emotion in Narrative Understanding and Interpretation". *Debates in Aesthetics* 14(1), pp. 60-73.

Walton, Kendall. (1990). *Mimesis as Make-Believe*. Cambridge, MA: Harvard University Press.

—. (1976). "Points of View in Narrative and Depictive Representation". *Nous* 10(1), pp. 49-61.

Willard, Mary Beth. (2019). "Reclaiming the Paradox of Fiction". *Debates in Aesthetics* 14(1), pp. 30-44.

THE IMPLIED PAINTER

Vanessa Brassey
King's College, London

1   Introduction

In this paper, I discuss Jenefer Robinson's personalist account of pictorial expression.[1] According to personalism, a picture possesses the expressive properties we attribute to it because we take it that *someone* expresses E in the work. Robinson's particular strategy exploits the concept of an implied persona who 'unifies' and 'specifies' what is expressed.[2] Dominic Lopes challenges this view by attacking what he takes to be a flawed assumption motivating the personalist account: the priority of figure expression.[3] Once we acknowledge this flaw, he argues, there is no good reason to prefer personalism to an impersonalist theory. I will argue that Robinson qua personalist can pre-empt this strike by clarifying the relation between (a) what a picture expresses and (b) what is depicted as expressing in the picture. Rather than leading with the idea that personalism unifies or determines the meaning, I argue that impersonalism diminishes it and that, should it be accepted, we would be left with an impoverished view of what specific art pieces can mean.

---

[1] Also referred to as a 'persona-theory' or 'persona-view', see Robinson 2005, 2007 and 2017.

[2] Robinson 2017, p. 260, p. 263.

[3] Lopes 2005, pp. 50-65.

2   Personalism

''NightHawks' is melancholy' (J1)
'Hopper is melancholy' (J2)

If Hopper feels sad, angry or nervous, Josephine might pick up on how he feels simply by looking at the expression on his face. When Hopper's mental states are revealed in this way, his expression is said to be 'transitive'.[5] Some behaviours are intransitive, that is they have the outward appearance of a transitive expression (a smile) but there are no mental states expressed. An example would be a smiling zombie. Josephine sees the zombie with the look of a smile, thinks of something mentalese, but nothing is in fact conveyed.

Whether we can map this distinction onto instances of pictorial expression is controversial. Intuitions run in both directions. Some think that we see the "mind, sensibility and skills"[6] of the painter in the work. Others think that the look of (J1) is parasitic on the public criteria or look of (J2), and so expression can be successfully tokened intransitively.[7]

Personalists argue that pictorial expression is always transitive, although they do not claim that all pictures are expressive in this sense. To give an indication of the kind of pictures that are considered expressive, 'The Scream'*,* 'The Raft of the Medusa' and 'Guernica' are typically discussed as promising candidates. Meanwhile, scientific illustrations and maps seem to lack these expressive qualities. There are of course difficult cases, like children's drawings or the work of mental patients, but I will just put these to one side for now.

---

[5] The term receives sustained discussion in Part 11 of Wittgenstein 1958. In the literature the terms 'expressing' and 'being expressive of' are used to express the same idea.

[6] Gayford 2010.

[7] Kivy 1980, pp. 67-68; Davies 1994; Lopes 2005.

Robinson, qua personalist, constructs her strong version of the view by conjoining two claims,

> (Pa) viewers view or should view the expressive content as transitive, and
> (Pb) the creation of expressive content should be understood as a transitive act of expression.[8]

Even if a personalist did commit to (Pa) and (Pb), (Pa) does not imply (Pb), and vice versa. However, as we shall see, Robinson's account endorses both claims.

3   Robinson

According to Robinson, pictorial expression is transitive. A picture expresses *only if* the artist expresses, by articulating her mental states (M) through the work: (Pb). For the sufficient conditions of expression to be met, a competent viewer must be able to pick up the expression: (Pa).[9] The artist articulates (M) through an implied persona, whose (i) expression is picked up by viewers and (ii) who is a psychological extension of the actual artist. This two-step process makes her view weaker than so-called transmission theories, according to which the actual artist's emotions are said to be directly transmitted through the vehicle of the painting to a receiver.[10] The implied persona, functioning as an expresser, unifies and determines what is expressed.[11] In a discussion of Spiegelman's 'Maus'[12], Robinson says:

---

[8] Theorists who claim (Pa) include Levinson 2006 and Vermazen 1986. Theorists who claim (Pb) include Robinson 2005 and Wollheim 1987.

[9] Robinson 2007, p. 36; Robinson 2005, p. 270; Robinson 2017.

[10] Collingwood 1938; Tolstoy 1962.

[11] Robinson 2017, p. 260, p. 263.

[12] Ibid., pp. 255-257.

> Barring any reasons for scepticism about Spiegelman's sympathies, it seems clear that the horror and dread expressed by the implied author of the picture reflect traits of compassion and distress that can safely be attributed to the artist himself.[14]

In this way, it is suggested that a psychological link ties implied artist to actual artists. Because implied artists are a construct of the actual artist's mind, they are constrained by the actual artist's own psychological profile and so contain "traces" of them when extended into the pictorial world.[15] She says 'The Scream' is,

> full of repressed desires, melancholy, and angst, and this is of course an important side of Munch's own personality.[16]

However, there is more to pictorial expression than merely depicting a figure in an expressive posture or with an expressive facial expression. Some figures may be transitively expressive, such as Munch's howling figure in 'The Scream', but others may be intransitively expressive. Those depictions which merely *look* sad or happy, such as models in a Calvin Klein advert, or emoticons, are not really cases of pictorial expression. They are examples of *technê*: skilled depictions of emotion.[17] Instances of *technê* map to Zombie 'expressions' as they convey nothing, whereas transitive expressions map to (J2) since they convey the artist's (M). The difference between the two apparently rests on how the artist has used what he paints to "articulate or individuate an emotion" often without knowing "what he will express until he has expressed it".[18] For this reason, some

---

[14] Ibid., p. 257.
[15] Robinson 1985, p .227.
[16] Robinson 2017, p. 11.
[17] Robinson 2005, pp. 2005-228.
[18] Ibid., p. 267.

depictions are and some are not also pictorial expressions. Pictograms and sentimental *technê* elements are "typically subordinate to and *explained by* the overall expression of emotions or emotional attitudes by the artist or his persona in the picture", where the implied persona acts as a proxy for the mental states of the actual artist.[19]

4   Lopes' challenge

Lopes develops one of the few sustained philosophical accounts of *intransitive* pictorial expression, in which he denies both (Pa) and (Pb). His argument against (Pa) is most germane to the analysis I develop here.[20]

Lopes identifies three ways a picture expresses emotions.[21] In figure expression (hereafter FE), such as occurs in Daumier's 'Fatherly Discipline', the depiction of a toddler in a tantrum straightforwardly represents a toddler having a conniption.[22] In-scene expression (hereafter SE) elements of nature may be depicted expressively to correspond or clash with FE.

> The shipwrecked, starving figures aboard Delacroix's Raft of the Medusa express despair; the roiling sea in which they are set adrift expresses dumb, haughty malignance; and the tiny ship on the horizon that might signify safe harbour instead expresses blind indifference.[23]

Lopes insists he is not claiming that by FE and SE "expression is depiction".[24] His view is that "by depicting a figure or scene as

---

[19] Robinson 2017, p. 263.

[20] Lopes 2005.

[21] Ibid., p. 57.

[22] Ibid.

[23] Ibid., p. 53.

[24] Ibid., p.56.

expressing the emotion" pictures express mimetically.[25] Mimetic expression may be bolstered or undermined by a third mode: design expression (hereafter DE). In DE, some brute fact about the way we correlate colours, shapes and textures with emotions links the "picture's design or surface" with expressive qualities. For example, Mondrian's complaint that curves are "too emotional" reveals "the trouble is with the curves themselves, not with anything that they depict".[26]

Lopes argues that Robinson (qua personalist) has taken FE to be *the* paradigm case of pictorial expression and as a result, developed an erroneous genealogy of personalism in order to retrofit SE and DE to FE. He argues that, for personalists

> [s]cene expression raises a missing person problem. Unless there can be expression in the absence of a being, to whom the expressed emotion is attributable, then either there is no scene expression or the being in question is one not depicted.[27]

But, he continues, there is no good independent reason to "attribute the emotion that is putatively expressed by a scene to some person who is not depicted" and so the "expression of emotion does not require that there be anyone to whom the emotion expressed is attributable".[28]

Robinson's crucial error according to Lopes, is to assume that viewers sustain an FE figure expression line of thought in the absence of figures.[29] But "once this assumption is dropped" he notes we "may adopt an impersonal theory [...] a dog can smile when it is not happy

---

[25] Ibid., p. 56.

[26] Ibid., p. 57.

[27] Ibid., p. 58 (my italics for emphasis).

[28] Ibid., p. 59.

[29] Ibid., p. 70.

(and so can zombies)".[30]

5   An Error

Lopes' aim is to discredit the motivation for personalism. But there is an error in his analysis. By confronting it, Robinson would neutralise an objection to arguments in support of (Pa) and (Pb).

Lopes' error is to conflate his categories of expression with a picture's expression of an emotion E. He says FE is, for personalists, the central manifestation of a transitive expression. But if the personalist is not motivated by the missing person problem, as described by Lopes, then his objection is neutralised. And, indeed, the personalist claim is not motivated by the missing person problem. Personalists do not argue that simply seeing depicted figures expressing emotions FE is sufficient for seeing a picture as being an act of expression. The flickr algorithm might compile rows of illustrated smiling faces, each showing those faces expressing happiness, yet I do not see the flickr webpage as an act of expression.

Instead of rejecting Lopes' analysis, Robinson tacitly accepts the notion of the categories and develops her argument in response to them. In doing so, she accedes to his construal of the 'missing person'. She argues that the missing person is 'a persona' of the artist and also an 'internal spectator' in the picture-world.[31] By pointing out that 'form and content' are interdependent, Robinson argues that Lopes is wrong to claim there are brute facts about the expressive qualities of DE and that it comes apart from FE and SE.[32] She says,

> Kokoshka's Self Portrait [...] shows the painter
> looking anxious and insecure (as in 'figure
> expression'), but also conveys anxiety and
> insecurity in the agitated passages of paint, the

---

[30] Ibid.

[31] Robinson 2017, p. 261.

[32] Ibid., p. 262.

awkward perspective and the vague, unstable use
of space (design expression).[33]

There is an unresolved tension in Robinson's reply to Lopes. The above quotation suggests that she thinks that DE is solely responsible for giving us the implied persona expression (FE is responsible for the figure 'looking anxious', DE for how the picture 'conveys anxiety'). "Design expression" must refer to DE, since she denies that there "is a fourth species of pictorial expression, in addition to figure, scene and design expression".[34] However, this contradicts a different response she makes to Lopes, namely, that: "Figure, scene and design 'expression' in and of themselves are only 'expressive elements,' not genuine artistic expressions of emotion".[35] It may be that the categories are significant in the way we come to see the picture as expressive, but on the basis of her Kokoshka example alone, it is difficult to see how they relate to each other.

My main concern, however, is that Robinson leads with the idea that Kokoshka's expression specifies and unifies what we see in the picture. I think that there is an alternative way to phrase the motivation for the personalist case, now that Lopes' challenge to the personalist starting point has been neutralised. This alternative links to Robinson's observations concerning the formal qualities of the picture as well as the things we see depicted in it, and brings out the different levels at which we understand a picture. For example, at first glance 'NightHawks' is a straightforward scene executed in a "light touch" noir-ish style. But given due attention, we can see a more serious or 'deeper' meaning in the picture. 'NightHawks' exhibits some stylistic features that seem incompatible with each other: for instance, the faces are rendered in a quite crude and mask-like way, while the ambience of the picture overall is sophisticated and self-assured. This can encourage the view that a naïve painterly

---

[33] Ibid.

[34] Ibid., p. 265.

[35] Ibid.

style is being put to use by a mature, skilled artist. Walton suggests that this kind of multi-level impression of an implied painter operates at a "deeper level" than just taking it that the work is painted by a naïve illustrator and can lead theoretically to the viewer experiencing a 'naïve illustrator' in some paintings that "derives from an obscure partial awareness of a multi-level situation of this kind".[36]

Since Lopes' categories of expression can only account for meaning derived from the formal (visible) qualities of the picture surface, meaning is restricted to what goes on *inside* the represented pictorial world. Robinson can push the worry that this is insufficient to capture the full extent of expressive meaning. The appeal of (Pa), in contrast, is to be found in how it explains these additional levels of meaning, unavailable to interpreters working solely with Lopes' categories. If so, then there are good reasons to consider the argument for implied expression.

An obvious counter from the impersonalist could involve denying that implied personae have any place in the interpretation of paintings. However, Robinson would appear to be on firm ground as (1) we need to account for more than the meaning of the visible formal properties of the picture, and (2) implied (narrator) personae are well established in the (relatively mature) philosophy of literature, where they are broadly understood as conveying intentions, beliefs or norms that do not belong to any of the characters in the novel.[37] Without a principled reason *against* extending (2) to pictures, Robinson is free to push for (Pa).

There is, however, the following objection to (2). Although some pictures, like some novels, give us the impression of the kind of person who created them, we should be cautious about how we extend the concept from narrative literary works to pictorial expression.

---

[36] Walton 1976, p. 52.

[37] Booth 1983.

Walton notes that for those novels that have implied narrators, the narrator plays a crucial role because,

> He mediates the reader's access to the rest of the fictional world; we know what happens in the fictional world only from his reports about it.[38]

But since the way viewers represent the pictorial world can be much more direct, the mediating gaze of the implied painter does not seem as crucial. I judge that a (fictional) tragedy has occurred by being told by an (implied) narrator, "baby shoes for sale; never worn". In contrast, I can judge (in the pictorial world) that there is a listless dynamic between the people in the bar from by how they appear directly to me. Although this deserves much more explanation and unpacking, it seems, from this initial observation, that the expression of an implied painter is going to play an incidental, rather than crucial, role in how the viewer gleans the expressive meaning of the picture. These considerations provide further support for my view that implied expression, wherever it may be found, is an additional level of expressive meaning that enriches viewer understanding, rather than a type of meaning that specifies and coheres what is expressed in the picture.

For reasons of space, I have simply noted this qualification on the arguments for (Pa). I now put the issue to one side since even if one can make out the case for this multi-level experience of (Pa), there is a more pressing issue that needs addressing, concerning the move from (Pa) to (Pb).

6   Moving from (Pa) to (Pb)

Having dealt with Lopes' objection to the motivation for personalism, I would like to now turn to the way Robinson's argument moves from

---

[38] Walton 1976, p. 50.

(Pa) to (Pb). I note some difficulties with the move and also some reasons to be cautious about the way (2) extends to pictures.

Robinson proceeds in more or less the following way. In real life, we form reliable impressions of people by looking at how they dress, the way they talk and so on.[39] An extrovert may wear bright colours; a confident character may strut; a wit produces pithy humorous observations and it is these traits that necessarily and inevitably reveal character. Similarly, in reading a novel or looking at a picture we may form an impression of the kind of person who authored or painted it. Sensitive prose and elegant brushstrokes will reveal aspects of the artist's personality which readers and viewers will pick up through some kind of epistemic seeing.

Even putting to one side the worries about whether we can reliably infer facts about, say, Josephine's psychology from her preference for ditzy chintz skirts, and simply granting Robinson the inference in real life, it is still not at all clear how this transposes to the case of pictures. This is in part due to her two-step notion of transitive pictorial expression: the claim that viewers can infer facts about the *artist's* psychology on the basis of a *construct*, the implied persona. For example, if it is the case that Hopper expresses his melancholy in 'NightHawks' then when the viewer views the melancholy as issuing from his implied persona (a construct) in the work in virtue of the implied persona that Hopper has constructed (hereafter *Hopper'*), the viewer views Hopper's melancholy.

Robinson has told us that because *Hopper'* is the creation of Hopper, *Hopper'* is necessarily shaped by and linked to the personality of Hopper. So if one picks up on an expression of melancholy by *Hopper'* we also pick up on Hopper's melancholy. Robinson relies on this psychological link to move from (Pa) to (Pb).

This is tendentious as it implies painters cannot create implied personas that are psychologically distinct from themselves. A well-known counterexample to this, however, is Tolstoy's ability to

---

[39] Robinson 1985.

write with compassion while lacking compassion in his personal relationships. For this reason, it expects too much of impressions since it cannot be that viewers go from (Pa) (viewing the painting as an expression by *Hopper'*) to (Pb) (understanding the painting to be an expression by Hopper), since *Hopper'* is merely a construct and so has no psychological reality. I cannot, when engaging with the ventriloquist's doll, legitimately move from the impression of sentience to believing I am picking up psychological states of the doll that are somehow given reality by an extension of the ventriloquist's states. So, it is not clear how viewers pick up on the emotions or traits of Hopper by constructing *Hopper'*. Even if viewers form the impression of an implied persona, since there is no entailment from (Pa) to (Pb), the fact that a viewer has the impression of an implied painter who expresses melancholy is not an argument for (Pb). The concern is that the distinction between the two claims that a personalist can make, (Pa) and (Pb), is being glossed over.

Robinson could argue that viewers infer a sincere connection between artists and their implied personae, by relying on extra-pictorial biographical information. The problem with this attempted solution, however, is that even if the viewer identified *Hopper'* as Hopper, it would still be the case that the viewer was connecting with *Hopper'* and not Hopper. Put in the counterfactual mood, the viewer would understand the picture even if they did not make the identification. So, the link between *Hopper'* and Hopper is not conceptual and, as we have noted, we do not have grounds to allow that it is psychological.

For these reasons, one cannot accept that the expression of an implied persona is also the actual expression of the artist. This should, however, not discourage further reflection on the significance of our impressions of a painting's origin. Arcing drips, bold palette knife work, and delicate glazes of paint may suggest a sensuous, agitated or careful personality in the style of the work. We can recognise this and seek an explanation for why the viewer chooses to represent the implied persona as if it were a psychological

continuation of the actual artist. That is, we may question why viewers make this connection, without supposing that the connection reflects a constitutive tie between two distinct identities. A viewer who sees the painting as an expression by Hopper would be making a harmless transition from viewing the expressive content as transitive to seeing the creation of the expressive content as a transitive act of expression. On this story, Robinson's move from (Pa) to (Pb) could be understood as a benign further claim, entertained but not entailed by the central issue. Meanwhile, implied expression can still be understood as a significant source of expressive meaning for paintings and, moreover, one that the impersonalist struggles to explain.

7   Conclusion

I have argued that impersonalism impoverishes our comprehension of some expressive paintings because it fails to fully reveal all the meaning in the work that only implied expression will explain. While this falls short of providing an argument for Robinson's controversial claims about the actual painter, it provides a reason to accept (Pa). However, there remains a highly complex and unresolved issue: namely, whether, in picking up on implied expression, we pick up on the mental states of painters or their implied personas, or if we merely imagine that we do so.

vanessa.1.brassey@kcl.ac.uk

philosophical theory and the artist's practice through interviews, public seminars, live events and film.

References

Booth, Wayne C. (1983). *The Rhetoric of Fiction,* 2nd edition. Harmondsworth: Penguin.

Collingwood, Robin George. (1938). *The Principles of Art.* Oxford: Clarendon Press.

Davies, Stephen. (1994). *Musical Meaning and Expression.* Ithaca, N. Y.: Cornell University Press.

Gayford, Martin. (2010). *Man With a Blue Scarf: on Sitting for a Portrait by Lucian Freud.* London: Thames & Hudson.

Kivy, Peter. (1980). *The Corded Shell: Reflections on Musical Expression.* Princeton, NJ: Princeton University Press.

Levinson, Jerrold. (2006). "Musical Expressiveness as Hearability-as-Expression". In: *Contemporary Debates in Aesthetics and the Philosophy of Art,* Kieran, Matthew (Ed.). Malden, MA; Oxford: Blackwell Pub, pp. 99-108.

Lopes, Dominic McIver. (2005). *Sight and Sensibility: Evaluating Pictures.* Oxford: Clarendon.

Robinson, Jenefer. (1985). "Style and Personality in the Literary Work". *The Philosophical Review* 94*,* pp. 227-247.

—. (2005). *Deeper Than Reason: Emotion and its Role in Literature, Music, and Art.* Oxford: Oxford University Press.

—. (2007). "Expression and Expressiveness in Art". *Postgraduate Journal of Aesthetics* 4(2), pp. 19-41.

—. (2017). "The Missing Person Found. Part I: Expressing Emotions in Pictures". *British Journal of Aesthetics* 57, pp. 249–267.

Tolstoy, Leo. & Maude, Aylmer. (1962). *What Is Art?: Essays on Art.* London: Oxford University Press.

Vermazen, Bruce. (1986). "Expression as Expression". *Pacific Philosophical Quarterly* 67, pp. 196.

Walton, Kendall. L. (1976). "Points of View in Narrative and Depictive Representation". *Noûs* 10, pp. 49-61.

Wittgenstein, Ludwig. (1958). *The Blue and Brown Books.* Oxford: Basil Blackwell.

Wollheim, Richard. (1987). *Painting as an Art.* London: Thames & Hudson.

RECLAIMING THE PARADOX OF FICTION

Mary Beth Willard
Weber State University

1

It is a truism of modern parenting that entertainment aimed at children needs to be carefully screened to avoid violence, sexual situations, and coarse language. Fortunately, there is a wealth of programming that avoids all of this. What is desperately needed, however, when choosing entertainment for a four-year-old, is a ratings system aimed at minimising suspense. My son, for example, finds any kind of suspense, any kind of tension or fictional peril, to be nearly unbearable. He immediately wants to stop watching because he fears too much for the well-being of the characters.

In children's movies, dramatic tension usually resolves happily. Bad guys don't win. But my sensitive son resonates like a theremin to the tension of the narrative and wails: "I don't like this! Turn it off!"[1]

As a parent, my thoughts turn to how to calm him down so that he might successfully finish a film. As a philosopher, my thoughts turn to the paradox of fictional emotions, standardly formulated as:

1. We have genuine and rational emotional responses towards fictional characters and situations.

---

[1] Lest any reader be concerned that I am mocking my dear son for the sake of a lively philosophical introduction, I confess that he comes by his thereminic qualities honestly. At age five I ran out of a theatre because the villain was making Rainbow Brite scrub the floors.

2. We believe that fictional characters and situations do not exist in reality.

3. In order for us to have genuine and rational emotions towards a given character or situation, we must believe that they exist in reality.[2]

All three are plausible. Any two can be held consistently. Holding all three together results in a contradiction. Eureka! Philosophy shall be a consolation for my son! Thus, I whisper:

"Don't worry. It's not real."

He shoots me a look of withering scorn. "I know *that*, Mum."

So much, then, for philosophy.[3] He is young, and new to fiction; he does not know the conventions of story-telling. Maybe this story ends with a terrifying princess flambé. From his perspective, what's happening *now* is distressing. It's possible, of course, that as a small child, his response signifies nothing more than his inexperience at regulating his emotions. But I think it offers a clue about something that's long bothered me about the paradox of fiction.

Attempts to resolve the paradox of fictional emotions focus on the non-existence of fictional characters, and star as examples in the larger debate between judgment theories of the emotions and embodied cognition theories of the emotions.[4] Judgment theories typically resolve the paradox by rejecting 1.: if emotions require cognitive judgments of danger, then whatever our physiological

---

[2] See Levinson 1997 for an overview of philosophical solutions to the paradox of fictional emotions; Gendler and Kovakovich 2005 delineate a particularly clear formulation of the paradox.

[3] In one sense, the theories are right. Pointing out that the characters aren't real does *interrupt* his engagement with the fiction. My son has replaced fear with scorn. But arguably all this shows is that interrupting someone's engagement with fiction makes it hard to engage with fiction.

[4] For example, Walton 1978, 1990, and 1997; Currie 1990; Stecker 2011; Tullman 2012; Tullman and Buckwalter 2014; Matravers 2014. But see Cova and Teroni 2016 for some pushback on this.

responses to fiction are, they're not genuine emotions. Embodied cognition theories, by contrast, explain that emotions result from affective appraisals of our circumstances, bodily changes that register what is significant to our well-being, without any intervening cognitive judgment.[5]

Cognitive appraisals of a situation can modulate our responses to the circumstances, but they do so by creating secondary affective appraisals; in other words, there is no change in emotions without first registering the change in the body. On Robinson's view, becoming emotionally involved in a movie requires nothing more than instinctively responding to the visual stimuli; bodies just aren't all that sophisticated. Robinson quips,

> It does not matter to my emotion systems (fear, sadness etc.) whether I am responding to the real, the merely imagined, the possible or the impossible. […] This might be irrational in some sense, but it happens all the time.[6]

Robinson shouldn't concede irrationality, or at least not on the grounds that the characters do not exist. As I shall argue, we should reject premise 3 in the standard paradox; existence turns out not to matter all that much to whether our emotions are rational. Instead, whether our emotional responses toward fiction are rational depends on our compliance with what I'll call the 'norms of aesthetic engagement': norms that govern the range of licensed emotional responses to the artwork.

2

Arguably, the paradox of fictional emotions has always, at its core, been a question of aesthetic appraisal, despite the direction the

---

[5] Robinson 2010, pp. 72–73.
[6] Ibid., p. 86.

philosophical literature has taken. Recall one of Radford and Weston's original examples, in which a man weeps irrationally over imagining the counterfactual, in which his sister suffers from infertility. They conclude that the man is irrational because he is a "sort of Walter Mitty, a man whose imagination is so powerful and vivid that, for a moment anyway, what he imagines seems real, that his tears are made intelligible, though of course not excusable".[7] The clear implication is that if it is irrational to weep over purely imagined states of affairs, then it is also irrational to respond emotionally to fictional characters and situations, which after all are imagined states of affairs, the production of which has been outsourced to an artist. So, the historical discussion has proceeded by assuming that fictional characters and situations are on a par with merely imagined fictional characters and situations, and that the problem is one of reconciling rationality with non-existence.

But if we look at the norms that govern our emotional reactions to fiction, we see that the question of existence fades into the background in favour of the work's licensing attitudes, behaviours, and beliefs. The question of the appropriateness of emotions directed toward a fiction is not fully answered by whether the attitudes are toward existing objects. One might exhibit the *wrong* emotions, such as if someone were to laugh uproariously at 'Schindler's List'. One might suffer from a *deficit* of emotions, such as when someone is too cynical to find the casual violence of a gritty film noir shocking. One might suffer from an *excess* of emotions, as my son does when watching anything with dramatic tension.

In these cases, whether the characters exist settles nothing. We would not respond to the person who laughs at 'Schindler's List' by saying, "Don't laugh! It's only a fictional retelling of Oscar Schindler's heroism.". Nor would we point to the non-existence of the gritty noir world as a reason to have *more* shock at the violence. And even if my son did relax upon being reassured that the characters for

---

[7] Radford 1975, p. 74.

whom he fears do not exist, he wouldn't be engaging with the fiction properly if he felt nothing. Small children are *supposed* to be shocked and concerned when they see Lightning McQueen crash, but not to such an extent that they cannot bear to continue the film.

Failing to react emotionally to fiction signifies not a wise person who refuses to be taken in by an illusion, but someone who isn't engaging the fiction *properly*.[8] Our emotional responses to fiction are governed by the norms of aesthetic engagement. Aesthetic normativity, broadly construed, concerns what is good and bad in beauty and art. Norms of aesthetic engagement govern the responsibilities appreciators of art incur when they engage works of art. Roughly, they say, "to experience the work correctly, appreciators should respond in this way". These norms include appropriate emotional responses, and two norms are relevant for our purposes here:

1. Being open to engaging with fiction.
2. Exhibiting the appropriate emotional response.

First, consumers of works should be open to allowing the work to influence them emotionally. This norm is relatively easy to follow. As Robinson observes, our initial emotional response happens automatically. Robinson writes that "emotion is a *process* at the core of which is a set of bodily responses activated by an *affective appraisal* that is 'instinctive' and automatic".[9] The emotional systems bequeathed to humans by evolution simply aren't all that complicated, so they can be triggered by images, sounds, imagination, contemplation, and even mistakes in judgment. In the case of works of art, emotional responses may be occasioned, as in the case of literary fiction, by vividly imagining the characters, or listening to music with great focus, or by paying attention to the detail of a fine painting. On the embodied cognition model, we experience emotions

---

[8] See also Robinson 2004b.

[9] Robinson 2010, p. 73.

in these cases only if our imagination, focus, and attention are registered physiologically. [10]

Second, consumers should exhibit the appropriate emotional response authorised by the fiction. [11] Robinson writes that "[t]his automatic appraisal gives way to *cognitive monitoring* of the situation, which reflects back on the instinctive appraisal and modifies expressive, motor, and autonomic activity accordingly, as well as actions and action tendencies". [12] I suggest that cognitive monitoring of our emotional responses to fiction is informed by knowledge of aesthetic and artistic practices. Having the correct emotional and behavioural response requires at least tacit knowledge of the artistic and aesthetic conventions of the artwork.

For example, some artworks are more demanding than others, requiring greater knowledge of those consumers who would engage with them. A newcomer to opera might find some of the conventions risible rather than moving. Moreover, the genre of a work can characterise what emotional responses are appropriate when engaging a work. For example, familiarity with the conventions of horror films will lead the savvy appraiser to feel apprehensive when the beautiful young heroine leaves her quiet farmhouse on a still night to check out a noise she hears coming from the shed. The noise itself isn't frightening, but the knowledge of how horror movies usually go induces tension in viewers. A satire is characterised by how the genre inverts attitudes that would be appropriate to the situation in the real world, so that to engage with a satire properly requires at least tacit knowledge that one is engaging with a satire. [13]

Appropriate engagement with 'Die Hard' requires allowing the character John McClane a certain amount of leeway as the hero of an action film, as heroes of action films are customarily permitted to

---

[10] Robinson 2004, p. 43.

[11] See Walton 1990.

[12] Robinson 2010, p. 73.

[13] Liao 2013, p. 274.

perform impossible feats of endurance without disrupting narrative engagement.[14]

Not every artwork will be successful in evoking the desired emotional responses in its consumers; in fact, we can often see the aesthetic and artistic conventions most clearly in cases where the work is not successful in evoking the correct emotional response. Consider the famous sequence in 'Battleship Potemkin' of the massacre on the Odessa steps. It is still a powerful scene, but one of its most iconic images, the orphaned baby in the carriage rolling helplessly down the stairs, has been copied by so many films that what was once no doubt a startling scene full of urgent pathos feels somewhat pedestrian. In other cases, we may recognise that a maudlin work is trying to make us feel sorrow, or that a comedy suffers from stale and flat jokes. In such cases, we recognise upon reflection that there is a norm that explains how we should feel, but one that we find impossible to fulfil.

Nor should we think of aesthetic affective engagement as a dramatic martinet, dictating the precise response of the audience. Rather, it is more like norms of social engagement. A norm that says that I should greet the barista politely before ordering my coffee, which admits of a wide range of social interactions: a friendly smile, a warm hello, a "good morning", a joke about manners before caffeination, a conversation, or just a nod and a brief and efficient transaction. Similarly, an aesthetic norm that instructs us to respond emotionally to a fiction will allow for a broad but related range of responses.

I suggest that an appraiser's emotional response to fiction is rational only if she satisfies the first two norms of aesthetic engagement. The physiological responses that we have once engaging with a fiction properly are nearly automatic, but whether they are *appropriate* requires a secondary cognitive judgment, which in the case of fiction derives from a knowledge of aesthetic norms. Here I

---

[14] Hazlett & Mag Uidhir 2011, p. 41.

propose that aesthetic norms parallel folk theories of the emotions. On Robinson's analysis, physiological responses are simple and automatic, but folk theories of the emotions are complex, culturally specific, and determine how the emotion is received. Cultural practices distinguish emotions by the context in which they are appropriate (e.g. anger vs. indignation). Some initial emotional responses are disregarded as inappropriate or inscrutable, so cognitive monitoring maps the initial affective appraisals onto a topography of what is salient or valued.

For example, *liget*, an emotion peculiar to the culture of the Ilongot in the Philippines, described as "a powerful energy running through the body" or the feeling of "want[ing]to take a man's head and throw it", can be explained only with regard to the circumstances of that particular tribe.[15] As initial emotional responses are coarse-grained physiological responses, a Westerner can experience the initial physiological analogue of *liget* but won't be embedded in a culture where describing the feeling as 'liget' makes sense. Nor will others in her culture be inclined to blame or excuse her behaviour based on her experience of *liget*, nor teach their children to expect feelings of *liget* to emerge after a profound personal loss. Similarly, for those who are not members of the relevant culture, the Japanese emotion of *amae*, a propensity to 'depend or presume upon another's love' or the Western emotion of *schadenfreude*, the feeling of taking joy in someone else's misfortune, simply will not be conceptualised in the same way.[16]

I suggest that just as the concept of *schadenfreude* rationalises the mixture of joy and fascination that a Westerner feels in response to the misfortune of a rival, the norms of aesthetic engagement rationalise the emotions we have in response to an artwork. Crying for no reason at all may be a sign of mental instability; but being moved at the fate of Anna Karenina is understood culturally as an appropriate

---

[15] Spiegel 2017.
[16] Robinson 2004, p. 39.

response to a great work of art. The initial physiological response happens as we engage and imagine the world depicted by Tolstoy, but the norms of aesthetic engagement help us identify this emotion as rational. Thus, whether the ostensible objects of one's emotions exist turns out to be a misleading question. Radford's wool-gathering man is irrational when he weeps because we do not have a common cultural practice of working ourselves into tears over isolated, imagined situations. Were he to share his imaginings as a fiction about a woman who dearly wanted children but couldn't have any, he might find that the other would applaud his weeping, and weep with him.

Moreover, thinking of the emotions as modulated by the norms of aesthetic engagement helps explain the very subtle character of the emotions we might feel towards fictional characters, in ways that are not easily explained if the only question we have to answer is whether the character exists or not. We might welcome the death of a character, even as we feel sad, knowing that the death will lead the hero to a satisfying revenge. We might tremble with dread at a happy scene, recognising from foreshadowing that disaster will soon befall the characters. My son's emotional overreaction to fictional peril can be tempered by reassuring him, "Don't worry. The good guys will win in the end.".

So, when Robinson quips that feeling emotions toward fiction might be in some sense irrational, but that it happens all the time, she concedes too much even by her own lights. "It happens all the time" because our physiological systems are not terribly sensitive, initially, to the source of the emotional stimulus. If the emotion that results from an encounter with a stimulus is believed to be rational, the culture must have a way of conceptualising it.[17] It is plausible that some of those norms concern our interaction with works of art.

So, we may revise the third premise of the paradox of fiction: In order for us to have genuine and rational emotions towards a given character or situation, we must believe either that they exist, or our

---

[17] Note that the claim is not that emotions are relative to cultures, but that *folk theories* of the emotions are relative to cultures.

reflective emotional response must be sanctioned by the norms of aesthetic engagement.

Note that I'm proposing, in accordance with the traditional formulation of the paradox, a necessary condition, not a sufficient condition. It is possible that a reflective emotional response could be sanctioned by aesthetic practices and nevertheless be irrational for other reasons.[18] A subculture that treated 'Schindler's List' as a farce would be irrational even if they had a well-developed aesthetic practice, because their practice would be monstrous!

More worryingly, we might object that aesthetic engagement is not necessary, because novices to artwork can exhibit appropriate emotions without any knowledge of the norms of aesthetic engagement. Imagine a person recently escaped from a hyper-Platonic enclave, in which all fiction had been banned. Upon her escape, the said person encounters a copy of 'Anna Karenina', reads it, and is profoundly moved although she has no theory of fiction or knowledge of aesthetic norms. But surely, she isn't irrational to respond with sorrow and grief; on the contrary, her perfectly appropriate reaction might be thought to demonstrate the timelessness and universality of the artwork.[19]

First, we should distinguish between the initial emotional response and the emotional response that develops as the result of cognitive monitoring, which is where knowledge of the norms of aesthetic engagement refine and conceptualise our emotional response. Our Platonic escapee will experience the story of 'Anna Karenina' as sad, on the assumption that despite her Platonic training, she is otherwise a typically functioning human being. In this restricted sense, her emotional response is rational.

One might object further, however, that my proposal implies that *knowledge* of aesthetic norms is necessary for genuine rational emotions. If our initial affective appraisal is automatic, and we refine

---

[18] Thanks to an anonymous reviewer for encouraging me to clarify this point.
[19] Thanks again to the exceedingly helpful, anonymous reviewer for this objection.

it only with cognitive monitoring informed by the norms of aesthetic engagement, we must know what those norms are. Thus, we should conclude that whether the Platonic escapee is rational or not depends on her knowledge of aesthetic norms. Aesthetic norms vary, and by hypothesis, she is from a culture that has no norm that concerns fiction. Doesn't this imply that *rationality* is dangerously relativistic?

This conclusion can be resisted. In this specific case, I find that my intuitions regarding her rationality depend on how we flesh out what our escapee thinks she's doing when she reads 'Anna Karenina'. Does she believe herself (à la 'Galaxy Quest') to be reading a historical document of a doomed woman? Then it seems that she is rational, though mistaken, to the extent that it is rational to mourn the believed bad end of any stranger. Has she heard only that within this new culture she's encountered, they regularly read fictional tales to induce strong emotions, and that one may feel sad even if one knows it is no more than a story? Then it seems to me that what this example shows is that only a very minimal knowledge of aesthetic conventions is necessary in order to conceptualise and refine one's emotional response to fiction. At least with respect to the very basic emotions, this seems right; one does not need a lot of exposure to fiction to understand the basic norms of aesthetic engagement. Does she fail, because of her hyper-Platonic heritage, to cognitively monitor her initial affective appraisal because she has no assumptions at all about 'Anna Karenina'? Then it's very hard to say how she would be able to reflect on what she's read; if that's all irrationality means in her case, it does not seem to pose a problem for my proposal.

More broadly, it's plausible to suppose that the norms of aesthetic engagement do not develop in isolation from other moral, social, and ethical norms. If that's the case, then while we might find that the norms of aesthetic engagement vary from culture to culture, the variance is no greater than we might find in folk theories of the emotions or moral theories, and if we can avoid a pernicious commitment to relativism in these cases, it is reasonable to suppose

that we can similarly avoid a commitment to relativism with respect to aesthetic engagement.

3

I have sketched two features of aesthetic engagement that allow us to treat the paradox of fictional emotions as a puzzle of aesthetics. We experience emotions by being open to engaging with fiction, and once engaged, we experience physiological changes that arise automatically. Our knowledge of aesthetic conventions and norms can revise our emotional responses to a work of fiction. Consequently, only emotions that are sanctioned by the norms of aesthetic engagement are appropriate to have toward fictional characters and situations. On the assumption that the embodied cognition model is correct, we should vigorously deny that the existence or non-existence of the characters has anything to do with the rationality of our emotions when it comes to fiction.

Much more needs to be said about aesthetic normativity, of course, especially as it related to ethics and practical reasoning. What grounds an aesthetic norm? How do we learn aesthetic norms? What about cases where we invert aesthetic norms, as when we enjoy a work, but only ironically? But I hope here to have shifted the conversation about the paradox of fictional emotions. Philosophers have ably responded to the paradox by developing theories of the emotions, but at its heart, the paradox is one about how we respond to the arts. Aesthetics should take this paradox back.[20]

marybethwillard@weber.edu

ABOUT THE AUTHOR   Mary Beth Willard is associate professor of philosophy at Weber State University. She earned her Ph.D. from Yale University in 2009, and works primarily on issues at the intersection of metaphysics and aesthetics. The topics of her recent publications range widely, from a criticism of simplicity as a criterion of theory choice in metaphysics to the ethical dimensions of street art. Her current project concerns art, public space, and cyberspace.

References

Cova, Florian & Teroni, Fabrice. (2016). "Is the Paradox of Fiction Soluble in Psychology?". *Philosophical Psychology* 29(6), pp. 930–942.

Currie, Gregory. (1990). *The Nature of Fiction*. Cambridge, UK: Cambridge University Press.

Gendler, Tamar, & Kovakovich, Karson. (2006). "Genuine Rational Fictional Emotions". In: *Contemporary Debates in Aesthetics and the Philosophy of Art*. Matthew Kieran (Ed.). Oxford: Blackwell Publishing, pp. 241-253.

Hazlett, Allan, & Christy Mag Uidhir. (2011). "Unrealistic fictions". *American Philosophical Quarterly* 48(1), pp. 33–46.

Lamarque, Peter. (1981). "How can we fear and pity fictions?". *British Journal of Aesthetics* 21(4), pp. 291–304.

Liao, Shen-yi. (2013). "Moral Persuasion and the Diversity of Fictions". *Pacific Philosophical Quarterly* 94(3), pp. 269–289.

Matravers, Derek. (2014). "Fiction and Narrative". New York: Oxford University Press.

Prinz, Jesse. (2004). "Which Emotions are Basic?". In: *Emotion, Evolution, and Rationality*, Dylan Evans & Pierre Cruses (Eds.). New York: Oxford University Press, pp. 69-88.

Radford, Colin. (1975). "How can we be Moved by the Fate of Anna Karenina?". *Proceedings of the Aristotelian Society, Supplementary Volumes* 49, pp. 67–93.

Radford, Colin. (1977). "Tears and Fiction." *Philosophy* 52(200), pp. 208–213.

Robinson, Jenefer. (1995). "Startle". *Journal of Philosophy* 92(2), pp. 53–74.

—. (2004a). "Emotion: Biological Fact or Social Construction". In: *Thinking About Feeling: Contemporary Philosophers on Emotions*, Robert C. Solomon (Ed.). New York: Oxford University Press.

—. (2004b). "The Art of Distancing: How Formal Devices Manage our Emotional Responses to Literature". *Journal of Aesthetics and Art Criticism* 62(2), pp. 153–162.

—. (2010). "Emotion and the Understanding of Narrative". In: *A Companion to the Philosophy of Literature*. Garry L. Hagberg & Walter Jost (Eds.). Malden: Blackwell, pp. 71-92.

Spiegel, Alix. "Invisibilia: A Man Finds an Explosive Emotion Locked in a Word". NPR, June 1, 2017. https://www.npr.org/sections/health-shots/2017/06/01/529876861/an-anthropologist-discovers-the-terrible-emotion-locked-in-a-word. Retrieved 11/15/2017.

Stecker, Robert. (2011). "Should We Still Care about the Paradox of Fiction?". *British Journal of Aesthetics* 51(3), pp. 295–308.

Tullman, Katherine. (2012). "HOT Emotions: Dissolving the Paradox of Fiction". *Contemporary Aesthetics*. https://digitalcommons.risd.edu/liberalarts_contempaesthetics/vol10/iss1/9/. Retrieved 5/7/2018.

Tullmann, Katherine & Buckwalter, Wesley. (2014). "Does the Paradox of Fiction Exist?". *Erkenntnis* 79(4), pp. 779–796.

Van Leeuwen, Neil. (2016). "The Imaginative Agent". In: *Knowledge through Imagination*, Amy Kind & Peter Kung (Eds.). New York: Oxford University Press, pp. 85–109.

Walton, Kendall L. (1978). "Fearing fictions". *Journal of Philosophy* 75(1), pp. 5–27.

—. (1990). *Mimesis as Make-Believe: On the Foundations of the Representational Arts*. Cambridge, M.A.: Harvard University Press.

—. (1997). "Spelunking, Simulation, and Slime: On Being Moved by Fiction". In: *Emotion and the Arts*, Mette Hjort & Sue Laver (Eds.). New York: Oxford University Press, pp. 37–49.

ARE EMOTIONAL RESPONSES NECESSARY FOR AN ADEQUATE UNDERSTANDING OF LITERARY TEXTS?

Eva-Maria Konrad
Goethe-University of Frankfurt

Thomas Petraschka
University of Regensburg

Christiana Werner
University of Göttingen

1   Introduction

In this paper, we want to examine the claim that emotional responses are relevant for the understanding of literary texts. This claim can be encountered in different versions. We will distinguish between a weak and a strong variant, both of them prominently presented by Jenefer Robinson. When Robinson suggests "that our emotions help us in the construction of a satisfactory summary reading of the novel or play or movie, or what is often called an 'interpretation'"[1], she advocates a weak version of the claim, because emotions are considered to be merely a *non-necessary* means to facilitate understanding.

We have no objections against this claim. There are numerous rather uncontroversial ways in which emotions can 'help' us to understand literary texts. Let us suppose John read Tolstoy's 'War and Peace', and that it made a deep impression on him.

---

[1] Robinson 2005, p. 105.

Subsequently, John was curious about other books written by Tolstoy and read 'Anna Karenina' as a result. One could say that epistemic emotions like curiosity, interest, the urge to broaden one's horizon, etc. played a role in understanding the work, because without them, John probably never would have read—and therefore, if all went well, understood— 'Anna Karenina'. Therefore, it seems perfectly reasonable to us to say that emotional involvement indeed helped John understand 'Anna Karenina'.

But Robinson is not satisfied with the claim's profoundness and proceeds to advocate a stronger version:

> Even if you agree with me that an emotional experience [...] helps us to understand [a novel], you might still think that it's possible to come to the same understanding by a more cerebral engagement with the text. [...] I'll try to show that this is not true: nothing else can do the job that emotions do. Without appropriate emotional responses, some novels simply cannot be adequately understood.[2]

The last sentence of the passage contains what we consider to be the most ambitious and also the most interesting variant of the claim, stating that emotions are relevant for an understanding of literary texts. It is a strong variant, because it considers emotions to be a

---

[2] Ibid., p. 107. Robinson restricts her claim to "realistic novels, plays and films [...] as part of the 'Great Tradition' of the English novel" (ibid., p. 106). In order to avoid an undue generalisation of Robinson's thesis, we follow Robinson in her choice of examples and concentrate on realistic novels and plays as well. Additionally, it should be noted that it is not the aim of this paper to discuss general conditions of adequate understanding. However, we assume, like Robinson, that an understanding of literary works *can* be more or less adequate. Accordingly, once again we will follow Robinson's own suggestions of what an adequate understanding might be in the examples she considers.

necessary condition of understanding.[3] In the following, we will concentrate on this version of the claim and analyse three central arguments Robinson presents in its favour.

2   The 'Trigger' Argument

The first argument could be called the *trigger argument*.[4] Its basic idea is this: every reader of a literary text is confronted with an enormous quantity of information she has to process. Emotions, so the argument goes, provide orientation and guidance in negotiating this thicket: They are "sources of salience"[5] in the sense that they "alert us to important aspects of the story".[6] Therefore, if we react emotionally to a certain episode, event, or character, "we are in a good position to try to discover *why* we respond emotionally as we do, and this in turn can lead us to seek in the work the origins of this response. [...] [T]hereby [we] acquire a deeper, fuller understanding of the work".[7]

The trigger argument has some persuasive power indeed. Aspects that cause strong emotional reactions are very often important for an understanding of the text. Consider, for example, Anna's last encounter with her son Seryozha in Tolstoy's 'Anna Karenina'. Not only does this episode usually elicit pity, the observation of Anna and her awareness of what her son means to her also causes the reader to realise that, from this moment on, Anna's life will be nothing but miserable since she simply cannot live

---

[3] Such a strong claim is not only advocated by Robinson, but also supported for example in Novitz 1987, p. 74f., and Miall 1989, p. 65. We suppose that understanding and appreciation are to be distinguished and that a full understanding of a literary text is possible without appreciating (parts of) the work.

[4] Cf. Robinson 2005, p. 108ff. & 126ff. Cf. also Elgin 2008.

[5] Elgin 2008, p. 43.

[6] Robinson 2005, p. 107.

[7] Ibid., p. 111.

without her son. Additionally, the trigger argument explains why repeated reading of the very same text can potentially reveal new information again and again. Every time we read, we are in a different emotional state and our emotions bring different aspects of the text into focus, thus potentially providing new insights.[8]

Nevertheless, the potential trigger-function of emotions does not prove the strong claim that when it comes to an adequate understanding of literary texts "nothing else can do the job that emotions do".[9] To support this claim, the trigger argument has to provide evidence that being triggered by her emotional reactions is *the only way* for a reader to become aware of significant episodes of a text. We cannot see how the trigger argument can do this. Think, for example, of a reader who is perfectly trained in all areas of literary studies. Her awareness could be triggered by a number of textual aspects that do not solely rely on emotional reactions, but are remarkable with respect to form or content as well: significant episodes might, for example, be placed right in the middle of the book (for example the fatal encounter of the English and the Scottish queen in Schiller's 'Mary Stuart'); they might be first or last instances of their kind (like the last encounter between Anna and Seryozha); or there might be special narratological features such as a change of focalisation or a change of discourse time (as for example the alternation of present and simple past at the beginning of Thomas Mann's 'The Magic Mountain'). To take one's emotional responses into account is hence to draw conclusions from *one of many* possible sources of salience. Emotional reactions are not exclusive in this respect.

Additionally, it seems problematic to suppose that, as Robinson says, only "*appropriate* emotional responses"[10] can trigger

---

[8] This is emphasised in Elgin 2008, p. 45.

[9] Robinson 2005, p. 107.

[10] Ibid. (our italics).

our awareness of significant passages of a text.[11] Suppose a sadist reads 'Anna Karenina' and because of her sadistic constitution she does not feel pity for Anna but enjoys her sadness and pain instead. Why should the inappropriate emotion of joy trigger the reader's awareness of central episodes of the text (like Anna's last encounter with Seryozha) any less than the emotion of pity would? Apparently, *any* emotion can function as a trigger. The trigger argument therefore fails to prove the strong thesis that "[w]ithout appropriate emotional responses, some novels simply *cannot* be adequately understood".[12]

---

[11] What appropriate emotional reactions are exactly and what makes them appropriate is a difficult matter that we need not delve into deeply here. It seems quite clear that some notion of "appropriate emotional responses" is firmly embedded in our everyday behaviour (cf. Elgin 2008, p. 38). It is appropriate to feel joy after finally reaching a goal one has worked hard to attain and to feel pity when we see someone suffering—but surely not the other way round. Since the literary examples we take into consideration are mostly clear-cut cases like this, we do not intend to dwell on complicated borderline cases. Robinson herself also goes for this modest approach and explicitly restricts herself to "examples in which it is pretty clear that the emotional responses are appropriate to the text and were probably intended by the author" (Robinson 2005, p. 108).

[12] Robinson 2005, p. 107. As Robinson admits, the trigger-function of emotions alone cannot grant an adequate understanding of a literary text either: By triggering the reader's awareness, emotional reactions apparently only help the reader to know where she should *start* her interpretation, i.e. which aspects, scenes, or episodes of a text she could take into consideration; but they do *not* help her with regard to the *content* of this interpretation itself, i.e. to interpret the emotionally triggered data in a reasonable way. A reader whose awareness is perfectly triggered by emotions could still be completely misguided in her interpretation of a text. To result in an adequate understanding of a literary text, the trigger must be followed by a thorough process of interpretation (cf. ibid., p. 122: "It is important to distinguish among *experiencing* the work, *reflecting* on our experiences of it as they occur, and *interpreting* it by reflecting on and reporting our experiences of the work after finishing it, by summing it up as a whole.").

3   The 'Empathy' Argument

Robinson considers a second way in which emotional reactions are necessary for understanding works of literature, namely emotional engagement with fictional characters:

> The argument in a nutshell is this: (1) Understanding character is essential to understanding great realistic novels I have in mind; (2) understanding character is relevantly like understanding real people; and (3) understanding real people is impossible without emotional engagement with them and their predicaments.[13]

In the following we will take (1) for granted but discuss claims (2) and (3). In order to examine the assumed epistemic value of emotional engagement with a real person we will start with Robinson's claim (3).

The phrase 'emotional engagement' refers to a wide range of emotional reactions, like falling in love with someone or being angry at one another; some reactions include forms of perspective-taking or feeling for or with another person, such as sympathy, emotional contagion, and empathy. Given the wealth of emotional reactions and the brevity of this article it is necessary to narrow down the scope of this phrase. Since it is often claimed that empathy is a way of understanding another's mind or at least a process that helps us to understand others,[14] empathic emotions seem to be the most

---

[13] Robinson 2010, p. 78.
[14] Cf. Stueber 2006.

promising candidates for emotional reactions that might play a decisive role in understanding a character or person.[15]

Following Coplan, we understand empathy as an imaginative process in which an empathiser simulates a target's mental situation *exactly*, where the latter involves not only states like beliefs and desires but also affective states.[16] Picking up Robinson's proposal, one could hypothesise that empathy's affective components, i.e. empathic emotions, have an epistemic value. We can, according to this specification, reformulate (3): Understanding real people is impossible without empathic emotions (towards them).

How can empathic emotions be a source of understanding others? Understanding someone's mental situation involves information about her mental states, which in turn involves not only cognitive but also affective states. The phenomenology of an affective state is a central aspect of these mental states. If one does not know what it is like to be, e.g. in the state of grief, one lacks important information about grief and hence about the mental situation of a grieving person. Experiencing an affective state is a way of getting information about what it is like to be in that particular state. We do not want to argue here that it is metaphysically necessary to experience a certain affective state in order to know everything about the phenomenology of this state, but we claim that it is psychologically necessary: in order to acquire *all* the information about the phenomenology of an affective state, human beings have to experience this state.

Even if this is true, we still have not shown that experiencing a particular affective state *empathically* is necessary in order to get information about what it is like to be in that state and hence for a proper understanding of the other's mind. Think of the following situation: A tries to understand B while B is in the affective state c,

---

[15] Other emotional reactions will be discussed in the following section on the 'data' argument.

[16] Cf. Coplan 2011.

and A has experienced *c* before. For a proper understanding of B, is it necessary for A to experience *c again*, this time empathically? We do not think so: If A has had the experience of *c* before, she knows what it is like to be in a state of *c* (let's take it for granted that she can remember the experience). If this is true then an empathetic experience of an affective state does not provide additional information about the phenomenology of the affective state in question if the empathiser has prior experience of the respective affective state. As a result we claim that empathic emotions are only necessary for those empathisers who have not felt the affective state in question before in order to fully understand the other's mind.

Up to this point we have only focused on the epistemic value of empathic emotions in understanding real people. Now, we see that even if (2) is true, readers do not necessarily have to empathise with the fictional characters in order to know what it is like to be in their mental situation—a prior experience of the affective state in question is sufficient.

Let us have a look at premise (2) now. The question is whether understanding a character's mental situation is relevantly similar to understanding the psychology of a real person. Even though there is no consensus on the question of which mental process the term "empathy" refers to, it seems that all accounts presuppose that there is a target with mental states. Independent of whether we follow a realist or antirealist account regarding the ontology of fictional characters, the problems for the fictional case start with this basic assumption: according to antirealist positions, there are no such things as fictional characters. Hence, there are no such things as mental states of fictional characters, whereas according to realist-creationist accounts, fictional characters exist. However here fictional characters are seen, e.g., as abstract artefacts[17] or posits of literary criticism.[18] In any case they are supposed to be

---

[17] Cf. Thomasson 1999.
[18] Cf. van Inwagen 1977.

entities without mental states. If one of these accounts is correct, either fictional characters, including their mental states, do not exist, or they only exist as entities without mental states. Against this background, the target of empathy in the fictional case is an open question. If we allow non-existent targets of empathy we still need to clarify what the mental states in question are.

A solution to this problem might be that empathisers in the fictional scenario simulate or imagine the mental states the fictional characters are in *according to the text*, e.g. if the fictional target is in a state of fear the empathiser has to simulate this particular state. Hence, the reader gets information about the phenomenology of the character's particular state. But many literary works are not or at least not always explicit about the mental states of the characters. Therefore, figuring out the mental states of the fictional character turns out to be a matter of textual interpretation. Being directly confronted with a real person whose mental state we want to figure out, however, is not very different from this procedure. It is, of course, not a matter of textual interpretation, but still, we need to interpret her outer appearance and/or her behaviour. It is worth mentioning that even real targets of empathy are sometimes not perceived directly. An empathiser can empathise with a real person while or after reading a text about her. Thus, it seems that interpretation is involved in all these cases. So, we assume that in this respect there are no decisive differences between the real and the fictional case.

It is, however, important to point out that if there were two or more equally plausible interpretations of the outer appearance and/or behaviour of a real person, the interpretation would not have an influence on the fact that a real person is in a particular state (or that she is not). This seems to be rather different in the case of fictional characters. There might be fictional texts with more than one equally plausible interpretation with regard to the mental situation of a character at a certain point in the story. Moreover, there could be an interpretation *a*, according to which the character

is in affective state *x* but not in affective state *y*. According to interpretation *b*, however, the character is not in affective state *x* but in affective state *y*. Even though the interpretations are contradictory, they might both be equally adequate, or equally "optimal", as Gregory Currie puts it.[19] Readers will then, following these different interpretations, simulate different mental situations and will therefore get information about different mental states by means of empathy with a fictional character. One empathising reader, following interpretation *a*, will get information about what it is like to be in state *x*; another empathising reader, following interpretation *b*, will get information about what it is like to be in state *y*. As long as empathy is based on equally adequate interpretations in both cases, either piece of information the readers get as a result of the empathic process is equally epistemically relevant concerning the mental situation of the fictional character,[20] while in the case of a real target there is a fact of the matter that decides the epistemic relevance of the information.

As a result, we claim that empathy with fictional characters is decisively different to empathy with real people in relevant aspects. We have argued, nevertheless, that empathy with real people as well as with fictional characters can provide information that is necessary for a proper understanding of the fictional character or real person in question. However, empathy is not a unique way of gaining the sort of information discussed above. A prior experience of the affective state in question can be a source of this information as well. Hence, in order to understand (certain) literary texts properly, empathy is only necessary for those readers who lack prior experience of the affective state in question.

4   The 'Data' Argument

---

[19] Cf. Currie 2003, esp. p. 293.

[20] It is in fact an open question whether both readers empathise with one or rather two different fictional characters.

The results of our discussion of the argument from empathy are furthermore relevant for the third argument we shall now discuss. Its core idea is this: to adequately understand a literary text, it is not enough to grasp the content of the story and to process all the information given by the text in a purely cognitive way. According to the argument, there are additional features of literary texts that can only be grasped via emotional reactions:

> [I]n reading a complex novel such as Anna Karenina, we don't just emote about Anna. We use our emotional responses towards her as data in arriving at an interpretation of her character.[21]

As discussed above, empathic emotions toward fictional characters can be sources of information for those readers who have never experienced the affect in question before. However, Robinson seems to be thinking of a further kind of information that is relevant for an interpretation of a character and hence of the whole work, but is not necessarily related to a characters's mental situation. In the following, we will focus on the question of which additional kind of data our non-empathic emotional responses might provide.

In the course of interpreting a text we often ascribe properties to texts or to fictional characters. Robinson ascribes the properties of being amusing and being repulsive to Strether and Macbeth, the protagonists of Henry James' novel 'The Ambassadors' and Shakespeare's play 'Macbeth', respectively.[22] Properties like these can be referred to as response-dependent properties, because whether an entity has such a property depends on a corresponding response. Roughly speaking, because we are repulsed by Macbeth's ruthless striving for power and amused by Strether's awkward yet loveable behaviour, Macbeth is repulsive and Strether is amusing. If

---

[21] Robinson 2005, p. 116.
[22] Cf. ibid., p. 111ff.

the response which decides the issue is an *emotional* response—just as in the cases of being amusing or repulsive—it seems very plausible to say that our emotional reactions provide access to a special kind of information, viz. information on whether a fictional character has got a certain response-dependent property or not.

We are not, however, convinced that this amounts to proof for the strong claim. It is not clear to us why information about characters having certain response-dependent properties should be provided *only* by emotional reactions. We think it is possible to find out whether a character has a certain emotional-response-dependent property or not without reacting emotionally to the character oneself. For the sake of simplicity, let us suppose that a character has the property of being amusing or repulsive as soon as at least one person finds him amusing or repulsive. It is important to note, however, that although being amusing or repulsive depends on the emotional reaction of at least one person, it does not depend on *a specific person's* emotional reaction. For this reason, we think that a reader can also find out that Macbeth is a repulsive character or that Strether is an amusing character if she, for example, observes the behaviour or the facial expression of another reader. Suppose Mary smiles, chuckles, and maybe sometimes even laughs out loud when she reads this or that passage of 'The Ambassadors' (and also suppose that Mary's facial reaction and behaviour is connected to her emotional reactions in a reliable way), such that the reader can infer that Mary finds Strether amusing, and consequently, that Strether has the response-dependent property of being amusing. The reader's *own* emotional reactions to Strether do not play any role in this scenario. If this is correct, information about emotional-response-dependent properties can, at least in principle, not *only* be gained by means of emotional reactions but also by other means.

To sum up: Although Robinson forcefully argues for the claim that emotional responses are necessary for an adequate understanding of literary texts, we remain doubtful about the strong version of her claim. While emotional reactions do of course often

trigger our awareness of important episodes in literary texts, the reader's attention can in principle also be grasped by other peculiarities of the same episode, which are perceivable without reacting emotionally. Robinson is also right in arguing that empathic reactions to fictional characters can grant information about the phenomenology of the affective states they are in. Empathic emotions, however, are not the only source of this particular information; prior experiences of corresponding emotions are sources as well. Therefore, empathic reactions are only necessary to fully understand the fictional character in question for those readers who have never experienced the emotion in question before. Finally, we have tried to show how it is possible to gain information about characters who have certain response-dependent properties even if one does not exhibit an emotional reaction oneself. If what we say is correct, the claim that emotional responses are necessary for an adequate understanding of literary texts is in need of additional support.

konrad@lingua.uni-frankfurt.de

thomas.petraschka@ur.de

christiana.werner@mail.uni-goettingen.de

ABOUT THE AUTHORS   Eva-Maria Konrad is a Research Assistant at the Department of Literary Theory and Comparative Literature at the Goethe-University of Frankfurt, Germany. Her research interests concern fiction, fictionality and knowledge, and the philosophy of literature. She is currently working on counterfactual literature.

Thomas Petraschka is an assistant professor of German literature at the University of Regensburg, Germany. His research focuses on the philosophy of literature and German literature from the 19th and 20th century.

Currently he is working on the relationship between literature and emotions.

Christiana Werner is leader of the junior research group "Language, Cognition and Text" at the Graduate School of Humanities Göttingen at the University of Göttingen. Her main research focuses on the philosophy of mind and philosophy of the emotions in particular and the philosophy of fiction. Currently she is working on an epistemological project about the affective components of interpersonal understanding.

References

Coplan, Amy. (2003). "Understanding Empathy". In: *Empathy: Philosophical and Psychological Perspectives*, Amy Coplan, Peter Goldie (Eds.). Oxford: Oxford University Press, pp. 3–18.

Currie, Gregory. (2003). "Interpretation in Art". In: *The Oxford Handbook of Aesthetics*, Jerrold Levinson (Ed.). Oxford: Oxford University Press, pp. 291–306.

Elgin, Catherine Z. (2008). "Emotions and Understanding". In: *Epistemology and Emotions*, Georg Brun, Ulvi Doğuoğlu, Dominique Kuenzle (Eds.). Aldershot: Ashgate, pp. 33–50.

Miall, David S. (1989). "Beyond the Schema Given. Affective Comprehension of Literary Narratives". *Cognition and Emotion* 3(1), pp. 55–78.

Novitz, David. (1987). *Knowledge, Fiction & Imagination*. Philadelphia: Temple University Press.

Robinson, Jenefer. (2005). *Deeper than Reason: Emotion and its Role in Literature, Music, and Art*. Oxford: Oxford University Press.

Robinson, Jenefer. (2010). "Emotion and the Understanding of Narrative". In: *Companion to the Philosophy of Literature*, Garry L. Hagberg, Walter Jost (Eds.). West Sussex: Blackwell Publishing, pp. 71–93.

Stueber, Karsten R. (2006). *Rediscovering Empathy. Agency, Folk Psychology and the Human Sciences*. Cambridge, MA: MIT Press.

Thomasson, Amy. (1999). *Fiction and Metaphysics*. Cambridge, UK: Cambridge University Press.

Van Inwagen, Peter. (1977). "Creatures of Fiction". *American Philosophical Quarterly* 14(4), pp. 299–308.

EMOTION IN NARRATIVE UNDERSTANDING AND INTERPRETATION

Quixote Vassilakis
Graduate Center, CUNY

1   Introduction

In this paper, I appraise Jenefer Robinson's account of how and in what ways emotion and literature interact, as this is developed in her book 'Deeper than Reason'. [1] Robinson argues for a specific framework for understanding human emotion—that it is a three-part process of affective appraisal, physiological change, and cognitive monitoring—every aspect of which "feeds back on the other aspects". [2] She further claims that this entire process must be experienced by readers of realistic fiction if they are to plausibly understand fictional characters, to fill in narrative gaps, and to gather data for critical interpretations of realistic fiction. While I agree with Robinson that the emotional engagement of readers is fundamental to their understanding of many works of literature, I suggest that the degree to which emotional responses, as she defines them, aids in understanding literary works remains an open question. In this paper I point out some reasons why I think that this is the case. Factors such as variability in emotional intelligence and the power that emotional bias can exert on readers' responses form part of my considerations.

---

[1] Robinson 2009, pp. 105-135.

[2] Ibid., p. 85.

I believe that the questions I raise are worth pursuing, not only because they can lead to a deeper understanding of Robinson's book, but also because important values are at stake. As a whole, 'Deeper than Reason' opens doors for readers, arguing for expansive access to great literature through their emotional responses. But if readers must respond to certain great literary works in accordance with particular emotional patterns in order to properly value them, as Robinson argues, then readers who do not respond in these ways are excluded from the possibility of valuing these works. But we may have good reasons to resist this conclusion. As Robinson admits, "little [work] has been done on the role of emotion in understanding narrative".[3] But, as I note below, some of the work which *has* been conducted, including intriguing findings on the ways in which observers with severe emotional deficits are able to accurately interpret the emotions and actions of others, provides a credible basis for asserting that much of this story is still to be told. Thus, though Robinson's theory purports to be inclusive, it may in fact exclude groups without good reason.

2   Robinson's theory of emotional response

Robinson's central argument for the role of emotion in readers's experiences of literature is that, to understand realistic novels like those written by the likes of Leo Tolstoy and Henry James or other great literary works like the sonnets and plays of William Shakespeare, we must experience them emotionally. Emotional involvement can focus our attention on important details of a story's plot, characters, setting, and points of view in ways that no other means of perception can, revealing nuances in character and plot that are otherwise inaccessible to us. Robinson notes that her theory does not hold for all fiction, excepting, for example, many post-

---

[3] Ibid., p. 101.

modern and genre works. Her detailed studies of emotion are restricted to great works of literature.[4]

For Robinson, emotional involvement with the events and characters in a novel has several consequences. We feel compassion for the imaginary characters of literature such that their interests and fate affect us personally. We respond to the them as if they are real people in our lives—or even as if we were those characters ourselves. These attachments trigger involuntary affective appraisals—gut-level reactions—regarding the characters well-being. And these appraisals, in turn, trigger physiological and autonomic responses in our bodies. All of this reinforces and focuses our attention on the characters and events in a novel and may even elicit action tendencies—states of physical readiness to take actions in the real world, even though our responses are to the characters and events of a fictional world.[5]

It is important to note that Robinson's theory requires there to be a physiological change in the body of the reader for the reader to be considered to be responding emotionally to a text. She cites William James, who claims that it is physiological change "that puts the 'emotionality' into emotion".[6] Robinson holds that only through an emotional response, as defined in this way, can a reader properly understand a literary work. The understanding she refers to is not solely emotional, such as the happiness or sadness that a reader might feel at appropriate moments. Rather, the understanding she has in mind is also cognitive. An emotionally unengaged reader cannot properly fill in narrative gaps in a novel by Henry James or Leo Tolstoy and so cannot build a cognitively accurate account of its events:

> When we respond emotionally to a text, our
> attention is alerted to important information

---

[4] Ibid., pp. 99-107.

[5] Ibid., p. 114.

[6] Ibid., p. 36.

about character and plot that is not explicitly asserted in the text.[7]

The parts of the story that are left unsaid are, in other words, narrative gaps that the reader must bridge. The psychological subtleties of characters can also be seen as narrative gaps that only emotional engagement can uncover. For example, in 'Anna Karenina', Vronsky's horse racing accident is horrifying, but appears differently when the reader realises that his affair with Anna might be responsible for his fall. While the text does not explicitly make this connection, the emotionally responsive reader might feel nervous as Vronsky competes in such a trying and dangerous sport while preoccupied by news of Anna's pregnancy. It is Robinson's point that by being emotionally engaged readers we shall be able to connect the dots between Vronsky and Anna's relationship and Vronsky's fall. We should, as emotionally responsive readers, understand that Vronsky's fall was not by chance, but rather, caused precisely by his distress over his relationship with Anna.

Cognitive monitoring, the third element of emotional response to literature that Robinson posits, aids the reader in curtailing the misunderstandings and action tendencies evoked by his or her initial affective appraisals and physiological responses. It reminds the reader that she is reading fiction and not responding to real-life events. It also provides "crucial data for an interpretation of the book as a whole" as the reader reflects on whether a novel did a good job of eliciting her emotional responses.[8]

Robinson holds that authors of great realistic novels rely upon readers's tendencies to draw upon their emotional experiences of characters to fill in narrative gaps. We understand that Tolstoy's Anna finds it difficult to openly follow through on her relationship with Vronsky because we experience, along with Anna, the emotional effort required to overcome an unfeeling spouse who has

---

[7] Ibid., p. 120.

[8] Ibid., p. 116.

control over her finances. Tolstoy relies on readers to simply *understand* that this is so by triggering their own emotions and connecting the dots.

3   Objections

Much of Robinson's work on narrative is based on theories of psychologists and philosophers who posit cognitive faculties as responsible for narrative understanding. An important way in which Robinson's work extends, and in some cases challenges, their theories is by arguing that direct emotional responses to literature, in addition to cognitive responses, are essential to achieving a plausible understanding of those fictional narratives.[9] This is what I specifically wish to expand upon and even challenge in Robinson's theory: her contention that a reader must experience a physiologically emotional response to a literary work in order to properly understand its narrative.

Low Emotional Quotient (EQ) and Literary Understanding

Robinson raises the scenario of people with low emotional intelligence attempting to understand a novel and concludes that, just as these people miss the significance of real-life relationships and events, so they will in fiction.[10] The ways in which low or damaged emotional intelligence has been shown in research to interfere with understanding other people and making good judgments suggests to Robinson that,

> cognition is not enough for understanding other
> people and that emotional understanding is
> crucial. If this is right, then it seems reasonable to

---

[9] Ibid., pp. 122-123.
[10] Ibid., p. 129.

> think that it is also crucial to understanding the
> 'people' who populate well-crafted realistic
> novels.[11]

Robinson's argument as a whole, however, would benefit from further discussion of this issue. How might she respond, for example, to research on psychopathy that investigates the ways that people with this severe emotional deficit may have a potentially enhanced—rather than impoverished—ability to understand the emotional cues of others, despite the fact that they do not experience many emotional responses themselves?[12]

In a 2014 article, researchers discuss this finding and explore how this indicates that recognising emotions cognitively—and experiencing them emotionally—"may be distinct processes, where one can be intact without the other".[13] In a separate study of lexical decision tasks, Amanda Lorenz and Joseph Newman found good evidence to support the hypothesis that psychopathic individuals are able to appraise affective language as accurately as non-psychopaths. The authors termed this phenomenon the 'emotional paradox' to indicate the difficulty of understanding how psychopathic individuals are able to understand emotional cues without feeling the emotion themselves.[14] Lorenz and Newman's research has been supported in a more recent article from 2018, in which researchers found that individuals with psychopathy are able to accurately classify the valences of affective language as well as non-psychopathic individuals *so long as they had ample time*.[15]

Consider the role of emotional responses in understanding literature in light of this finding. The so-called 'emotional paradox', or the duality between recognising and experiencing emotion that

---

[11] Ibid., p. 130.

[12] This phenomenon has been noted by Book et al (2007), who describe it as a "special intelligence", p. 532.

[13] Sandvik et al 2014, p. 585.

[14] Lorenz & Newman 2002, p. 91.

[15] Vitale et al 2018.

some attribute to psychopathy, indicates that it may be possible to understand the emotional dimension of a work of literature without actually experiencing the physiological response. It may be possible to 'get' what is funny about Henry James's Strether and what is sad about Tolstoy's Anna without the affective appraisal and autonomic response that characterise emotions for Robinson.

Psychopathy is an extreme case, but it still resides within the wide spectrum of human emotional response. More common along this spectrum is the emotionally normal reader who is simply not emotionally involved with a novel to the degree that he or she responds to it physiologically. It is plausible to argue that all readers—whatever their emotional capacities—may have access to non-emotional cognitive tools for understanding narrative.

Some recent research has even suggested that one function of sleep, the REM dream phase in particular, is to consolidate memories and process their emotional tone—as Van der Helm and Walker put it, "we sleep *to forget* the emotional tone, yet *sleep to remember* the tagged memory of that experience".[16] If there is a parallel between the content of dreams and narrative in literature, then this finding could suggest that our memory and thereby understanding narratives in literature succeeds by suppressing emotional responses to it.

The problem of emotional bias

Richard Gerrig, a psychologist known for his investigations into reader responses to fiction, is one of Robinson's sources of empirical evidence for demonstrating the role of emotion in narrative understanding. Gerrig theorises that readers use a "core of automatic responses" to understand narratives.[17] They automatically engage their working memories as they read in order to construct causal networks which allocate attention and situate events within the

---

[16] Van der Helm & Walker 2012, p. 781.
[17] Robinson 2009, p. 121, quoting Gerrig 1993, p. 48.

context of the whole work.[18] Through this process, they commit important events to memory while discarding those of less importance. Some aspects of understanding a work, Gerrig holds, are "under the strategic control of the reader", but the "core of automatic processes" is not.[19] As Robinson notes, the automatic processes Gerrig describes are largely logical or cognitive, though he also identifies 'participatory responses' that arise in readers. These are non-inferential responses, some of which involve emotion.[20]

In his experimental studies, Gerrig found that readers' "expressions of hope and preferences have measurable consequences for the memory representations constructed in the course of experiencing a narrative world".[21] If readers did not prefer an outcome, they were less likely to remember that it *did* happen and what the outcome of it was; their participatory responses as readers resulted in a measurable effect on memory.[22] In other words, when narrative events run counter to readers's preferences, those events are less likely to be accurately captured and stored in their memories in such a way that they could be used to understand later developments. They may escape the notice of the reader altogether, or they may be recast in a more favorable light and thus misconstrued.

Bias poses a significant problem for any theory of narrative understanding that emphasises the role of emotion because many literary works present characters and events that challenge our expectations and our preferences for real world outcomes.

In my own experience, an author whose works present situations contrary to commonly held real-world preferences is Fyodor Dostoyevsky. His portrait of the Underground Man in 'Notes from the Underground' exemplifies a narrative development

---

[18] Ibid., p. 121, quoting Gerrig 1993, p. 76.
[19] Ibid., quoting Gerrig 1993, p. 69.
[20] Ibid., p. 120.
[21] Ibid., p. 122.
[22] Ibid., p. 123.

apparently so painful to some readers that it is frequently misunderstood. [23] Readers have come away from 'Notes' thinking that the Underground Man should have simply relaxed and gotten along better with his old school mates. A reader may prefer that the protagonist's feelings toward Liza, the young prostitute, are genuine to such a degree that they forget the protagonist was merely playing a deceptive game all along. Considering this type of response through the lens of Gerrig's findings, it is reasonable to suggest that these readers are perhaps actively suppressing (or never forming) their memory of the actions of a protagonist who is so thoroughly disaffected. These readers, then, may miss the work's depth and psychological significance, or even thoroughly misunderstand its narrative, confabulating, for example, alternative endings that align better with their own preferences and understandings.

Bias is built into emotion, in part, because emotion necessarily elicits action tendencies, which can be difficult to reverse—or even to be aware of—to a degree that I believe Robinson never fully explores. Although, as Robinson notes, readers employ cognitive monitoring to moderate action tendencies provoked by emotional responses to fictional characters and worlds, most of us are unable to quiet our physiological responses at will. When we wake up from a nightmare the emotions we felt during the dream remain with us. Difficulty shaking off a bad emotion is a common experience even when the emotion seems irrational.

This same phenomenon can be found among readers who become so immersed in stories that they feel as if fictional characters are real and become unable to control their emotional responses to them through cognitive monitoring. This problem is the inverse of the paradox of fiction (the apparent contradiction of fictional events inspiring real-life emotions), which Robinson explicitly discusses as a potential objection to her theory. She rebuts this objection, in part, with reference to research that shows there is no clear line between

---

[23] Dostoevsky 1993/1864.

the real and the fictional in our emotional responses. Indeed, I believe this underscores my point. Readers who are deeply emotionally involved with a novel can construe fictional worlds as real; they can retain their emotional attachment to characters over time, and some never lose the sense of the characters and the events of a work. Such a deep sense of the reality of a world can block a full recognition that it is a fictional world. This in itself can interfere with forming a plausible understanding of a work, if for no other reason than that it is, in itself, a key misapprehension of the context of that work. This leads to another potential objection to Robinson's theory, which she raises as an objection to her theory only to rebut, that readers can have inappropriate emotional responses to fictional worlds.[24]

Inappropriate emotional responses

Robinson acknowledges that works such as 'Anna Karenina' admit a range of valid interpretations. And she agrees that some interpretations that are based upon readers's emotional responses are simply not credible. Readers can easily ignore too large a portion of the text or may suffer from a 'lack of awareness' while reading. Or it may be that "idiosyncratic personal interests forbid or prevent" readers from enlarging their "emotional horizons by sympathetically engaging with people who are different from them in significant ways".[25] In other words, sometimes readers can just get it wrong, even when they are emotionally engaged.

Robinson's point is well taken. The fact that there can be exceptions to a rule—in this case, the rule that readers need to respond with appropriate emotions to a work in order to understand it—does not mean that there is no rule. However, it remains fruitful to press the point further. Robinson casts reader error as incidental and largely correctable. I am not so sure this is the case. It can

---

[24] Robinson 2009, p. 141.
[25] Ibid., p. 142.

credibly be argued that the ubiquity and persistence of idiosyncratic emotional response is a central feature, not a rare anomaly, of emotions.

Interpretation

Robinson distinguishes between "*experiencing* the work, *reflecting* on our experiences of it as they occur, and *interpreting* it by reflecting on and reporting our experiences of the work after finishing it, by summing it up as a whole".[26] Only after a reader has finished a book can he judge his reflections and experiences of "the work by offering an *interpretation* of the work as a whole".[27] Interpretation is part of—and an extended reflection on—the readers's emotional responses to a novel as a whole, including its characters and narrative, in order to determine what provoked their emotion and whether this particular provocation was justified. It is, Robinson holds, an extension of the emotional process because it includes as important data the readers's emotional response to the work.

But, as noted above, emotional responses, while indeed providing readers with important data about a work, can also interfere with their ability to accurately notice and remember narrative detail important to the understanding of the work. For the sake of argument, let us assume that a reader follows Robinson's prescriptions and sets about interpreting a novel by drawing upon his or her basic working understanding of its historical context, genre, the plain meaning of its language, and a wealth of personal emotional responses to it, including extended reflections. Thus equipped, the reader must figure out what it was in the story that provoked his or her emotional responses and whether they were justified. Especially considering that readers's physiological 'participatory responses' can become an obstacle to their correct

---

[26] Ibid., p. 123.
[27] Ibid., p. 125.

understanding, I am sceptical that most readers's interpretations will prove sound. In my view, the most plausible part of this hypothetical reader's interpretation is their search for the provocations and justifications for their emotional responses. In this case, their interpretation is reliable insofar as it is, in fact, cognitive. Robinson situates cognitive appraisal as part of readers's emotional responses. Clearly, however, readers also have access to cognitive appraisal that are separate from emotion.

4   Conclusion

In this paper I have questioned Jenefer Robinson's description of the degree to which emotional response—as she has defined it—contributes to readers's understanding of works of great literature. I have suggested that her theory strikes me as suggesting a too-narrow gate-keeping that might unnecessarily exclude some readers from benefiting from these works. I have argued that this exclusion may be premature, as research salient to this point is still evolving.

I have objected to Robinson's extension of—and challenge to—earlier theories that narrative understanding is largely based on cognitive assessment, pointing out that cognition that does not include a physiologically emotional response may be adequate for narrative understanding, while emotional responses may introduce bias into readers's narrative understanding. I largely base this objection on ideas illuminated by Richard Gerrig's research.

Robinson covers important ground in this work, pulling together a large body of research and thinking by philosophers, psychologists, and others to present readers with an enriched means for approaching great literary works. My objections to some narrow— but I believe important—aspects of her work are meant not as a rejection of her ideas—but as an expansion.

qvassilakis@gradcenter.cuny.edu

ABOUT THE AUTHOR   Quixote Vassilakis is an adjunct lecturer in the philosophy department at Brooklyn College, CUNY. He is writing his master's thesis on Hannah Arendt's political philosophy in the Liberal Studies Program at the CUNY Graduate Center.

References

Book, Angela S., Quinsey, Vernon L., & Langford, Dale. (2007). "Psychopathy and the Perception of Affect and Vulnerability". *Criminal Justice and Behavior* 34(4), pp. 531-544.

Dostoevsky, Fyodor. (1993) [1864]. *Notes From the Underground*. (Richard Pevear & Larissa Volokhonsky, Trans.). New York, NY: Random House.

Gerrig, Richard J. (1993). *Experiencing Narrative Worlds: On the Psychological Activities of Reading*. New Haven: Yale University Press.

Hu, Peter, Stylos-Allan, Melinda, & Walker, Matthew P. (2006). "Sleep Facilitates Consolidation of Emotional Declarative Memory". *Psychological Science* 17(10), pp. 891-898.

Robinson, Jenefer. (2005). *Deeper than Reason: Emotion and Its Role in Literature, Music, and Art*. Oxford: Clarendon Press.

Sandvik, Asle M., Hansen, Anita L., Johnsen, Bjørn Helge, & Laberg, Jon Christian. (2014). "Psychopathy and the Ability to Read the 'language of the Eyes': Divergence in the Psychopathy Construct". *Scandinavian Journal of Psychology* 55(6), pp. 585–592.

Van Der Helm, Els, & Walker, Matthew P. (2012). "Sleep and Affective Brain Regulation". *Social and Personality Compass* 6(11), pp. 773-791.

Vitale, Jennifer, Kosson, David S., Resch, Zachary, & Newman, Joseph P. (2018). "Speed-Accuracy Tradeoffs on an Affective Lexical Decision Task: Implications for the Affective Regulation Theory of Psychopathy". *Journal of Psychopathology and Behavioral Assessment* 40(1), pp. 412-418.

ROBINSON AND SELF-CONSCIOUS EMOTIONS
Appreciation Beyond (Fellow) Feeling

Irene Martínez Marín

Uppsala University

1   Introduction

In this paper, my focus will be on some central aspects of Jenefer Robinson's influential work 'Deeper than Reason', more specifically on the role our emotional responses play in art appreciation and the value attributed to the ensuing emotional experience. Whereas Robinson argues (1) that bodily responses, and our awareness of these bodily changes, can provide us with information relevant to the appreciation of artworks, and (2) that in some cases affective empathy is *necessary* to artistic understanding, I want to raise two concerns about whether this position holds for artworks conveying self-conscious emotions.[1] Such emotions are of particular interest in this context since, in the first place, self-conscious emotions can, in fact, be experienced without moving us physiologically in a full-fledged sense. These higher cognitive emotions, also known as non-primary or intellectual emotions—including guilt, shame, embarrassment, pride, or nostalgia—are not automatically triggered

---

[1] Even though higher cognitive emotions have traditionally received considerably less attention from emotion researchers than the so-called basic emotions such as joy, fear, and sadness, there has been a significant increase in research on higher cognitive emotions in recent years. See, for example, the edited collection by Tracy, Robins, & Tangney 2007 and Stocker 2010.

but require self-evaluation. And, unlike the basic emotions, they are also believed to lack stereotypical expressive or behavioural features.[2] Second, self-conscious emotions are unavailable to affective empathy since they are known for involving self-directed cognition.[3] As we will see, it is this tight connection to the self—where the 'self' is both the subject and the particular object of the emotion—which makes it difficult to take the emotional perspective of another person.

My strategy will be to question the weight Robinson places on bodily feelings and affective empathy by bringing Peter Goldie's theory of emotions into play. To that end, I will discuss a particular self-conscious emotion that Robinson's model is unable to accommodate within her two-step theory of emotions: autobiographical or reflective nostalgia. I take this kind of nostalgia to be *not* a mere deliberate fantasy about the past (e.g. the longing for a non-lived period of history like a retro-nostalgia for the decade of the fifties), but an emotion concerned with the irrevocability of one's own past.[4] The artworks I propose to look at in order to develop my claim are the avant-garde films by Jonas Mekas, 'Reminiscences of a Journey to Lithuania' (1972) and 'As I Was Moving Ahead Occasionally I Saw Brief Glimpses of Beauty' (2000).[5] These autobiographical films, also known as 'film-diaries', will enable us to discuss the possibility of a model of art appreciation that includes emotions that do not require bodily feelings. At the same time, Mekas' characteristic style of presenting a collage of images about his

---

[2] Tracey & Robins 2007, pp. 6–7.

[3] For more on the object of self-conscious emotions, see Tracey & Robins 2004.

[4] See Boym 2001, Van Tilburg, Wildschut & Sedikides 2018 for a characterisation of nostalgia as a self-conscious emotion.

[5] Although Robinson's artistic examples are mainly focused on her areas of expertise—literature and music—her theory has the potential to be expanded to other art forms. Here I propose film, a medium classically focused on emotions, as a candidate that can benefit from this discussion. See Smith 2017, for a recent analysis of Robinson's work on emotions in a philosophy of film framework.

everyday life while offering his own poetic observations on what the viewer is watching will allow me to show why *fellow-feeling* is not the appropriate way of engaging with artworks of this type—what I shall discuss below as 'intellectual artworks'. I will suggest that the nostalgic view from which Mekas' film-diaries are narrated presents problems for perspective-shifting, a necessary process in empathetic responses, according to Robinson. The reason for this lies in the impossibility of sharing the particular object of the nostalgic emotion expressed in these experimental films—something desirable and unrecoverable from Mekas' past.[6] The memories presented by the Lithuanian filmmaker are to be understood as a personal self-evaluation of the director's past, and, as an invitation to reflect on one's personal story, not as an emotional episode to be felt on Mekas' behalf. I will conclude that the value of artworks expressing such self-reflective states is, first and foremost, an epistemic one. This is a value bound up with the aesthetic features of the artwork but not reducible to its artistic value (at least if the latter is narrowly construed). Note that this paper should not be perceived as a rejection of feeling as an important part of our artistic experiences, but as a reconsideration of its role and function.

2   Emotional Understanding

How much do we need to feel in order to be able to appreciate aesthetically?

This section highlights the intricate relation between cognition and feeling in our emotional responses to art. I will show that certain emotions towards artworks do not seem to operate in the way Robinson describes. The emotions I have in mind do not, it seems, occur as the result of a non-cognitive affective appraisal of our

---

[6] Howard 2012, p. 641.

environment, as Robinson suggests, but instead appear to call for some kind of conscious self-evaluation at the very moment of their experience. My argument will reveal that Robinson's analysis suffers from being developed primarily with emotional bodily feelings as a model and does not sufficiently take into account the distinction between being emotional, on the one hand, and having an emotion, on the other hand. For that reason, what I want to put forward here is a view of art appreciation inspired by Goldie's theory of emotions,

> where there are emotional feelings of a kind that can be directed immediately towards objects in the world beyond the bounds of the body: these feelings are bound up with cognition and perception, and are not the mere effects of cognition or perception.[7]

One of Robinson's main claims in 'Deeper than Reason' is that in some cases our emotional responses are means towards artistic understanding. To use Robinson's own words, "cognition without emotion simply does not do a good job".[8] On this account, in order to gain an appropriate understanding of an artwork – that is, not only of the particular emotional state expressed but of the overall point of view of the work—we need to be able to attend to the bodily feelings related to the emotional experience of the artwork. As Robinson puts it:

> Emotional understanding is in the first instance a kind of bodily understanding: my affective appraisals of characters, events, and situations are automatic and instinctive, and they immediately

---

[7] Goldie 2009, p. 232.

[8] Robinson 2010, p. 80.

> produce physiological and behavioural responses
> that reinforce these emotional appraisals.[9]

This defence of the importance of bodily feelings in our artistic experiences is in tune with the portrayal Robinson gives of emotions outside of art appreciation. Her model, in a nutshell, is characterised by identifying an emotion as a non-cognitive appraisal of our environment—that is, an automatic and instinctive evaluation—which *causes* a bodily response in us (e.g. physiological changes, tendencies to act, and expressive gestures).[10] This bodily reaction helps the subject to focus attention upon those things in the environment that matter to her. Robinson adds a second step to her model to ensure the evaluative appropriateness of our bodily response when she claims that this affective appraisal that is experienced subjectively as a feeling subsequently has to be modified by cognitive monitoring, or reflection.[11] In a similar sense, emotions for Robinson are a useful tool for artistic criticism because they serve the double function of, first, alerting us to what is of significance and, second, preparing us to critically evaluate these initial responses. For example, my weeping for a character like Anna Karenina works as an inarticulate understanding of the events read about, and it is by evaluating this initial response that I develop my aesthetic appreciation of the novel.[12] In addition, it is when we experience these artistic emotions and reflect upon them that we also come to gain a deeper understanding about the artwork. As Robinson declares:

> If I laugh and cry, shiver, tense, and relax in all the
> "right" places, then I have in some sense

---

[9] Robinson 2005, p. 127.

[10] For a complete description of the account, see Chapter 3 in Robinson 2005, pp. 57–99.

[11] Ibid., pp. 75–9.

[12] Ibid., p. 123.

understood the story. But if I want not just to have a rich emotional experience while reading the novel but also to give a critical account—a "reading" or an "interpretation"—of it, this requires reflecting on my emotional experience. I need to reflect upon my affective appraisals, figure out what they were, what it was in the story that provoked them, and whether they were justified.[13]

However, for Robinson not every artwork needs to be emotionally understood in order to be appropriately appreciated. There are some artworks that she excludes from the process of emotional understanding, namely those she describes as being about 'ideas'. These artworks, she states, demand to be experienced on an 'intellectual' level, not on an emotional level.[14] In her own words:

Different kinds of artworks have very different goals. Some are mainly concerned with design or form. Others deal mainly with ideas and demand to be appreciated primarily on an intellectual level. I am not arguing that all artworks have to be experienced emotionally if they are to count as art.[15]

Against Robinson, I believe that these intellectual artworks she refers to should also be experienced emotionally. The only difference is that the emotions we engage with in these cases tend to be what are commonly referred to as higher cognitive emotions. Now, the heart of the matter here is Robinson's refusal to call 'emotions' those emotional feelings that do not seem to involve awareness of any

---

[13] Robinson 2010, p. 77.

[14] Robinson 2005, p. 102.

[15] Ibid., p. 102.

particular bodily change or state. An important part of Robinson's motivation for prioritising bodily feelings is the fact that her neo-Jamesian account is built on a base of solely the basic emotions (anger, fear, sadness, happiness, surprise, disgust). William James, in 'What is an Emotion?', famously describes self-conscious emotions as non-standard or purely cerebral, and gives hardly any attention to them, because emotions without bodily feelings are, for him, 'only judgements'. The main difference between standard emotions and non-standard, or 'intellectual' emotions, is that "the bodily sounding—board, vibrating in the one case, is in the other mute".[16] Robinson, following James' lead, also gives a problematic and, in some sense at least, counterintuitive characterisation of 'cold' emotions when she argues that, for example, a lack of bodily feelings during reflection on our remembrances, in the case of nostalgia, prevents nostalgia from having the status of an emotion. In Robinson's view:

> If I wish I were young again and in the Paris of yesteryear, then we may say that I am nostalgic for Paris, even if I am not making any non-cognitive appraisal and I am physiologically unmoved. I am not experiencing an emotional response to (my thoughts of) Paris. We can call my state an "emotion of nostalgia" if we like, but this is misleading because there is no "emotionality" or emotional upheaval.[17]

My main worry, then, is that, by identifying emotions with individual bodily feelings felt at a specific time, Robinson's model seems to be excluding from the process of art appreciation, a set of emotions that are familiar to us in life and in the arts. In other words, this model

[16] James 1884, p. 202.
[17] Robinson 2005, p. 95.

excludes those emotions that are partly constituted by cognitive processes that unfold over time and which require attention to the personal history of a person; such as my case study, nostalgia, but also grief, shame, envy, pride, etc. Furthermore, the cognitive monitoring or reflective, second-step, which is of vital importance for Robinson's account of the epistemic value of emotions, is something I find problematic since it seems to go against Robinson's refusal to afford higher cognitive emotions the status of really being emotions. One could think that these emotions, especially the self-conscious type, would hold a privileged position in Robinson's account, since they do not need to be subsequently cognitively recognised in order to warrant the appropriateness of the bodily feeling; but as we have seen, this is not the case. What we need, then, is an account of emotions and of art appreciation utilising a unified concept of emotion that does justice to the particularities of both the basic and the higher cognitive emotions. One way of giving a better characterisation of cognitive complex emotions is by adopting Goldie's useful distinction between bodily feelings and *feeling towards.*[18] Goldie's theory, like Robinson's, posits a close identification between emotion and feeling, but contrary to Robinson's model, with the notion of *feeling towards* Goldie can make sense of feelings that do not involve any characteristic physiological change. This notion would also explain distinctive emotional episodes where bodily feeling is involved and yet the feeling is not always present. In Goldie's words:

> I think it is possible to do justice to the importance of feelings in emotional experience without needing to claim either that emotions must involve feelings at all times, or that, for each type of emotional experience, there will be a

---

[18] See Goldie 2009.

distinctive phenomenology which is uniquely identifiable with it.[19]

In what follows, with the help of Jonas Mekas' naturalistic films, I try to show how the notion of *feeling towards* allows for a different understanding of the role of emotions in our appreciation of artworks, especially when assessing the value of artworks that deal with ideas. In 'As I Was Moving Ahead Occasionally I Saw Brief Glimpses of Beauty', which is made of fragments of homemade films recorded over a period of about thirty years, we see the everyday lives of Mekas' family and friends. Here, no strong bodily feelings are elicited; instead, we find pleasure in the contemplation of past events, in the ambiance of the film, in ordinary but meaningful situations like weddings, picnics, or family reunions, all of them depicted in an intimate and subtle way.[20] This film sets the spectator in a reflective mood by inviting her to experience a kind of nostalgic feeling that goes beyond the boundaries of the body. Mekas' films are an example of artworks that might have associated feelings, but which do not have associated bodily feelings or behavioural dispositions.[21] When engaging with them we come to judge them as evocative, poignant, or consoling, and we do this by identifying precise artistic choices, and without necessarily being physiologically moved by them. For example, we recognise the calm piano that accompanies Mekas' own voice, the collage-montage, or the exploitation of silence as elements that infuse this work with

---

[19] Goldie 2000, p. 69.

[20] Ruoff 1991, p. 14.

[21] Jesse Prinz discusses how there are some emotions that do not have observable behavioural dispositions; nostalgia being one of these. "It may be difficult to link some emotions with specific behaviors (e.g. hope, pride, aesthetic pleasure, confusion, nostalgia, and so on). It may be that some named emotions can impact behavior in a variety of different ways, and it may also turn out that some emotions are not associated with highly specific biologically prepared behavioral responses" (2012, p. 202).

nostalgia. Likewise, the themes of the film–childhood memories, immigrant identity, the reconstruction of a personal self–contribute to setting the spectator in a self-evaluative state. Whereas for Robinson *bodily feelings* are directed towards the body and the awareness of these bodily changes help the appreciator to scan and detect the important aspects of the artwork, *feeling towards* can help the appreciator to take certain aspects present in the artwork under the 'appropriate light'. Hence, from a Goldiean view the relevance of feelings in appreciation is not about being causally moved by features of an artwork, but has to do with acquiring the 'right' phenomenology required for correctly judging the aesthetic features present in the work. For example, in Mekas' film, where "the restless handheld camera and the rapid montage never holds enough on any image to fix it in our mind", the viewer will only perceive these images as memories if she is able to pick up the phenomenal character of the nostalgic emotion in place.[22] So, *contra* Robinson, I would suggest that the emotionality of the artistic experience is revealed in the way we perceive the artwork. This means that our emotional responses towards artworks are bound up with cognition and perception, and are not the mere effects of cognition or perception.[23] In other words, having an emotional experience of an artwork does not exhaust our understanding of the artwork, but what is relevant about emotions is that they can provide us with a deeper understanding of the relevant aesthetic features present in an artwork. And from there, once we have the appropriate phenomenal experience it is more likely that we acquire an appropriate judgement of the artwork.

I agree with Robinson that in our engagement with certain artworks we tend to scream or at least flinch when a monster appears on screen, or to weep at the end of a sentimental film. That said,

---

[22] Rouff 1991, p. 16.

[23] For Goldie's discussion on how non-cognitivism gets 'phenomenology' wrong, see Goldie 2009, pp. 235-6.

emotional bodily feelings are not the only feelings we can experience. Mekas' films show that in artistic contexts we can also be drawn to experiment with subtle or more cognitively complex emotions. In these cases, what we will typically be engaging with are things grasped with the mind, such as thoughts and memories. It is the automaticity of responses to art such as tears, screams, or laughs, that leads Robinson to propose that the comprehension of our responses, and of the artwork itself, comes after the evaluation of their appropriateness. The reason for this is that no degree of bodily feeling can by itself reveal what our emotions are about, and that is why Robinson's model needs a second step, lest she merely be discussing a subjective feeling.[24] The point I am making here is that emotions like the nostalgia expressed in Mekas' 'film-diaries' do not require this second moment because they are by nature partly constituted by reflection. I do not mean to say with this that we do not also reflect upon our self-conscious emotions and come back to analyse what made us respond in such a way. We can attend to what we felt when experiencing a reflective emotion and evaluate the appropriateness of our emotional response to the artwork, but, as I understand it, for Robinson, cognitive monitoring is a necessary step to justify not only the emotion's appropriateness with the artwork but the emotion itself.

My central idea in this section can be described as follows: whilst artworks typically require an emotional response, not every artwork requires an emotional bodily response—think of abstraction, conceptual art, experimental film, or post-modern literature. These types of artworks, which do not elicit bodily responses in us, are normally signified as 'cerebral' or about 'ideas directed to the mind'. I believe that these tags are problematic since they leave intellectual emotions out of the art appreciative process. Therefore, Robinson might consider adapting her account if she wants to be able to accommodate artworks involving higher

---

[24] On this see Goldie 2000, p. 58.

cognitive emotions, but also, I claim, if she wants to be able to do justice to a general insight that has arisen from recent philosophy of emotions, namely the thought that cognition and feeling are not distinct or easily separable.

The reflective emotions I have been discussing might not be vividly felt in an automatic way, but in this section I hope to have shown how sometimes in the arts it is emotion without cognition that does not do a good job.

3   Affective Empathy

Which emotional perspective should we adopt in art appreciation?

Recently, empathy has been accorded a central role in explaining our emotional involvement with fictional characters and events presented in a narrative.[25] These empathetic responses are thought to be an intimate form of identification in which we come to make sense of another person's emotions by imaginatively engaging with the person's point of view. In this section, I challenge Robinson's account and push her to take into consideration the limitations of perspective-taking as a tool for art appreciation when applied to self-conscious emotions. Indeed, on many accounts empathy is seen as a valuable tool in the proper appreciation of artworks—including film.[26] Robinson, for her part, defends the view that 'high level' affective empathy can be helpful for understanding the point of view from which an artwork is shaped, with this being especially true for novels, plays, and films that purport to depict 'real life' events.[27] Also,

---

[25] One of the most important volumes on empathy is the edited collection by Coplan & Goldie 2011.

[26] See Smith 1995; Gaut 2010; Vaage 2009.

[27] Robinson 2005, p. 106.

it is through *fellow-feeling* that we gain access to key elements of appreciation like the theme, vision, and unity of the work.[28]

So, if we want to understand a character that we care about, like Anna Karenina, what we need to do is to imagine what is like to be in Anna's situation, and we do this not only to understand her despair but also to acquire a fuller understanding of the work.

That said, the presumption that such affective empathy is possible might be seen as problematic when we apply it to self-conscious emotions, which are bound up with a strong sense of self. So, while, for example, nostalgia typically involves a process of self-evaluation and of self-awareness—you are nostalgic about *your* memories—fear and pity do not seem to be necessarily concerned with the self, but with external objects, such as green slimes or characters like Anna Karenina, respectively. And, even though Robinson does not want to generalise by asserting that 'every emotional response to art is empathetic', her notion of art appreciation does not provide us with a clear picture of which emotions or affective aspects of an artwork are to be empathetically understood and which not. I suspect that for Robinson's account, those characters or events that we feel positively towards count as an invitation to be felt in an empathetic way (feeling Anna's own sorrow), while those who we feel negatively towards are not to be emotionally understood in an empathetic manner (revulsion for Macbeth).[29] This difficulty of feeling empathy for characters like Macbeth captures our imaginative resistance to adopting perspectives that go against our values or worldviews. For example, Alex Neill maintains that the success of empathic imagination is

---

[28] For Robinson, "fellow feeling is fellow *feeling*, not just an intellectual recognition that someone is in trouble, say, but a 'gut reaction' or compassion, an emotional or bodily response, a response consisting in autonomic and motor changes" (2010, pp. 80–81).

[29] Robinson 2005, pp. 108–12.

dependent on how much the other is *like myself*.[30] However, in the case of reflective emotions—where the self is part of the formal object of the emotion—the limit is not a problem of imagination or of possessing different psychological dispositions. I believe that the problem lies in the impossibility of accessing the particular object of another's self-conscious emotion. Here there is no external perceptual recognition or imaginative access to the target of someone else's emotion. In the case of nostalgia, we fail to intentionally imagine the other's recollections in an empathetic way because the yearning associated with this emotion is inseparable from the subject's self-evaluation of her own past. Furthermore, in order to experience nostalgia, we not only need to be aware of the event that is remembered by someone, we also need to have the correct phenomenal experience of the event. This is partly due to the fact that in order to experience nostalgia it must be the case that one has the feeling of being in a specific space, transported by one's own memories. So even if nostalgia can be triggered by external objects perceptually recognisable by all, it is the personal connection with this object to someone's autobiography that sets off the emotion. It is this narrative sense of self that makes self-conscious emotions peculiarly resistant to empathy.

Along similar lines, authors like Noël Carroll, Peter Goldie, and Peter Lamarque believe that the difficulty of perspective-shifting lies in the different informational and evaluative positions between the audience and the characters involved.[31] Goldie understands that this problem is not a contingent one that has to do with imaginative limits, but a conceptual one, because perspective-shifting usurps the 'full-blooded notion of first-personal agency that is involved in deliberation'.[32] Therefore, putting it more crudely, self-conscious emotions seem to work as anti-empathetic states that need to be

---

[30] Neill 2006, p. 254.

[31] See Carroll 2011; Goldie 2000, 2011; Lamarque 2011.

[32] Goldie 2011, p. 303.

experienced from one's personal perspective. [33] Let's go back to Mekas' example to get a better sense of what I am suggesting.

In 'Reminiscences of a Journey to Lithuania', made in the summer of 1971, we find the story of a displaced person who decides to go back to his home country after escaping from the Nazi occupation during the Second World War. This is the personal story of Jonas Mekas, a man in search of a home, who is in the process of reconfiguring his identity when the emotion of nostalgia appears. He feels profoundly nostalgic about his hometown, the Lithuanian village of Semeniskiai, and the family he had to leave behind. He longs for insignificant but meaningful things, like the taste of the water he used to drink from a well, or the old women that reminded him of 'sad autumn birds' when he was a child. From time to time, we hear the filmmaker's voice reflecting on these scenes. And it is through Mekas' delicate and painful descriptions of his own past experiences that we come to understand that his manner of valuing this document is very different from a spectator's appreciation of the nostalgic work, even for someone like Mekas' brother, who accompanied him in the process of making the film.[34] This asymmetrical emotional response towards the film is noted by Tomkins when he declares that:

> [a]ccording to Adolfas, the trip was an intensely emotional experience for his brother, who broke down and wept several times when he was called on to say something before a gathering. Adolfas, three years younger than Jonas and completely at home in America, apparently was not subject to the same emotions. But for Jonas the trip seemed

> to confirm his long-standing suspicion that he has
> not yet found any place of his own in the world.[35]

The impossibility of empathising with Mekas occurs not because his psychological states are alien to us—we all have nostalgic memories and the capacity to recollect those *brief glimpses of beauty*, but, contrary to what Robinson seems to be holding, because the artwork's perspective is not something that we can, and more importantly should, attempt to match. What, then, is the appropriate emotional response to Jonas Mekas' work? Although the embodiment of nostalgia in a personal narrative makes it inaccessible to us, at least if it is a question of our supposedly feeling the same way as the nostalgic subject, that does not mean we cannot be sensitive to it at all. So, even if the images we find in Mekas' film-diaries do not have the same meaning for, or connection with, the spectator, his work is still valuable to us because we are able to see his artistic intention of retelling the past as something familiar. It is just that our emotional responses will not be a *pure* autobiographical nostalgia like the one expressed by Jonas Mekas himself.

Nevertheless, there are times in the viewing of Mekas' films in which we can experience some kind of analogous but self-directed autobiographical nostalgia. This can be explained without appeal to empathy, but through memory association. This is an idea shared by Robinson when she states that we possess emotional memory.[36] According to Robinson, we store past emotional situations that help us to understand similar episodes as the ones presented in a novel or film. In Mekas' case, the content of the film can trigger a nostalgic memory in us, and this is a pleasurable experience for the spectator because we tend to feel pleasure in what is familiar to us. It is this feeling of pleasure in the familiar but, more importantly, the capacity Mekas has for presenting the world as broken and beautiful which

---

[35] Tomkins 1973, p. 32.

[36] Robinson 2010, p. 74.

connects us to his story in an intense and unique way. So, the nostalgic experience of Mekas' work does not reside in its capacity to be *fellow felt* (or experienced in a bodily way), but in how the aesthetic experience involving an acknowledgment of Mekas' nostalgia can inform our knowledge of our own past.[37]

An example is the moment in 'As I Was Moving Ahead Occasionally I Saw Brief Glimpses of Beauty' where we hear the filmmaker's voice inviting us to create this intimate connection between his life and our own:

> I am looking at these images, now, many, many years later. I recognise and remember everything. What can I tell to you, what can I tell to you. No. No. These are images that have some meaning to me, but may have no meaning to you at all. Then, suddenly, this being midnight, I thought: there is no image that wouldn't relate to anybody else. I mean, all the images around us, that we go through our lives, and I go filming them, they are not that much different from what you have seen or experienced [...] From what you have seen or experienced. All our lives are very, very much alike.[38]

If Robinson's theory specifies that art appreciation sometimes requires the appreciator to adopt the perspective from which a story is told, then it needs to give space to those emotions that require a strong sense of self and that are not readily imagined by another person. While I believe that she would urge us to try to imagine what it would be like to have lived Mekas' life, I believe that the epistemic merit of Mekas' film resides instead in its capacity to help us to

---

[37] My thanks to Jeremy Page for urging me to pursue this.

[38] Mekas 2013, p. 13.

understand our *own* personal past. And, as art appreciators, what we ought to assess is not only this invitation to reflect about the way we think about the past, but the way in which emotions are conveyed and crafted in Mekas' experimental films.

In this section, I have used Mekas' autobiographical films—and our responses to them—to challenge Robinson's assertion that affective empathy has the potential to play an important role in our appreciation of artworks. I have focused on the self-conscious emotion of nostalgia, which, I argued in section one, Robinson's account should accommodate as an emotion if she wants to do justice to our complex emotional life. As noted above, many commentators have suggested problems with the notion of affective empathy, and my contribution here consists in showing that self-conscious emotions present a specific and unique challenge to accounts like Robinson's because these emotions are tied to the subject's sense of self and autobiographical self-understanding, something which we necessarily cannot share. The intimacy that Mekas reveals with his work is a personal invitation, or guide, to experience the world in a very specific way, with tenderness whenever childhood memories of his daughter are portrayed, wonder and amusement when images about their youth in Greenwich Village during the sixties appear on screen, or even with indignation when sad memories about WWII and his life as a displaced person are remembered. I believe that a possible path for Robinson to explore would be to think of our experience of self-conscious emotions expressed in the arts as yielding an epistemic merit related to an enriched understanding of how emotions work. Asking us to share another's emotions and the feelings based on them might be asking for too much.

4   Conclusion

In this paper, I have shown that we cannot explain all emotional cases using the two-step model defended by Robinson. Her theory of

emotions is problematic for those artistic cases where understanding and value do not reside in the capacity of the audience to bodily feel or imagine the same emotion expressed. As we have seen, not all emotions are either bodily feelings or instances of empathy. I have focused on the nature of self-conscious emotions, and in particular on the expression of autobiographical nostalgia in Jonas Mekas' 'film-diaries', in order to show how emotional feelings, in art appreciation, can be better understood if we adopt Goldie's notion of *feeling towards*, instead of Robinson's *bodily feeling*, especially in the case of 'intellectual' artworks. I have also tried to assess the limitations of perspective-shifting in Robinson's affective empathy when applied to reflective emotions that are tied to a strong sense of self. In Jonas Mekas' work, the epistemic value of his films—and our appreciative response to them—comes not from bodily feeling the same nostalgic emotion expressed in his highly personal work, but from how the aesthetic experience involving an acknowledgment of his nostalgia can inform our knowledge of our own past. It is in this recognition that we understand Mekas' exploration of private and subjective memories as something universal.[39]

irene.martinez@filosofi.uu.se

ABOUT THE AUTHOR   Irene Martínez Marín is a Ph.D. candidate in Philosophy at Uppsala University. She is currently writing a thesis on the relation between emotions, reasons, and value within our aesthetic

judgements. She is the coordinator of the Nordic Network for Women in Aesthetics.

References

Boym, Svetlana. (2001). *The future of nostalgia*. New York: Basic Books.

Carroll, Noël. (2011). "On Some Affective Relations between Audiences and the Characters in Popular Fictions." In: *Empathy: Philosophical and Psychological Perspectives*. Amy Coplan & Peter Goldie (Eds.). Oxford: OUP, pp. 162-84.

Howard, Scott. A. (2012). "Nostalgia". *Analysis* 72, pp. 641-650.

Gaut, Berys. (2010). "Empathy and Identification in Cinema". *Midwest Studies in Philosophy* 34(1), pp. 136-157.
Goldie, Peter. (2000). *The Emotions: A Philosophical Exploration.* Oxford: Oxford University Press.

—. (2009). "Getting Feelings into Emotional Experience in the Right Way". *Emotion Review 1*(3), pp. 232–239.

—. (2011). "Anti-Empathy". In: *Empathy: Philosophical and Psychological Perspectives*. Amy Coplan & Peter Goldie (Eds). Oxford: Oxford University Press, pp. 302-17.

James, William. (1884). "What is an Emotion?". *Mind* 9(34), pp. 188-205.

Lamarque, Peter. (2011). "On Keeping Psychology Out of Literary Criticism". In: *The Aesthetic Mind: Philosophy and Psychology*.

Elisabeth Schellekens & Peter Goldie (Eds.). Oxford: Oxford University Press, pp. 299-312.

Mekas, Jonas. (2013). "As I Was Moving Ahead Occasionally I Saw Brief Glimpses of Beauty". *Cinema Comparat/ive Cinema* 1(3), pp. 10-17.

Neill, Alex. (2006). "Empathy and (Film) Fiction". In: *Philosophy of Film and Motion Pictures*. Noël Carroll & Jinhee Choi (Eds.). Malden: Blackwell, pp. 247-359.

Prinz, Jesse. (2012). "Emotion". In: *The Cambridge Handbook of Cognitive Science*. Keith Frankish & William M. Ramsey (Eds.). Cambridge, UK: Cambridge University Press, pp. 193-211.

Robinson, Jenefer. (2005). *Deeper than Reason*: *Emotion and its Role in Literature, Music, and Art*. Oxford: Clarendon Press.

—. (2010). "Emotion and the Understanding of Narrative". In: *A Companion to the Philosophy of Literature*. Garry L. Hagberg & Walter Jost (Eds.). Malden: Blackwell, pp. 71-92

Ruoff, Jeffrey K. (1991). "Home Movies of the Avant-Garde: Jonas Mekas and the New York Art World". *Cinema Journal* 30(3), pp. 6-28.

Smith, Murray. (1995). *Engaging Characters: Fiction, Emotion, and the Cinema*. Oxford: Clarendon Press.

—. (2017). *Film, Art, and the Third Culture: A Naturalized Aesthetics of Film*. Oxford: Oxford University Press.

Stocker, Michael. (2010). "Intellectual and Other Nonstandard Emotions". In: *The Oxford Handbook of Philosophy of Emotions*. Peter Goldie (Ed.). Oxford: Oxford University Press, pp. 401-23.

Tomkins, Calvin. (1973) "All Pockets Open. Profiles: Jonas Mekas". *The New Yorker*, pp. 38-39.

Tracy, Jessica L., & Robins, Richard W. (2004). "Putting the Self into Self-Conscious Emotions: A Theoretical Model". *Psychological Inquiry* 15(2), pp. 103-125.

Tracy, Jessica L., Robins, Richard W., & Tangney, June Price (Eds.). (2007). *The self-conscious emotions: Theory and research*. New York: Guilford Press.

Vaage, Margrethe Bruun. (2009). "The Role of Empathy in Gregory Currie's Philosophy of Film". *British Journal of Aesthetics* 49(2), pp. 109-128.

Van Tilburg, Wijnand AP, Wildschut, Tim, & Sedikides, Constantine. (2018). "Nostalgia's Place Among Self-conscious Emotions". *Cognition and Emotion* 32(4), pp. 742-759.

RESPONSE TO CRITICS

Jenefer Robinson
University of Cincinnati

I would like to start my responses with grateful acknowledgments to all those who responded to the invitation by 'Debates in Aesthetics' to comment on my work, and especially to those who submitted the very accomplished papers that were accepted and which I will be discussing in this commentary. I am also very grateful to the editors of 'Debates in Aesthetics', Ryan Doran and Shelby Moser, for inviting the critics to write about my work and for giving me the opportunity to respond. It has been both fun and stimulating to grapple with these papers. I think I have learned a lot. I have certainly been pressed to defend my positions on various issues and to justify some of the more outrageous claims I have made.

1

Vanessa Brassey[1] gives a rich and thoughtful response to a paper of mine called 'The Missing Person Found: Expressing Emotions in Pictures'[2], in which I develop what she calls a 'personalist' theory of pictorial expression, in contrast to the 'impersonal' theory defended by Dominic Lopes in 'Sight and Sensibility'.[3] Brassey's paper is full of insights and I cannot address all of them here, so I will focus on three points which are the most interesting and important from my point of view: (1) the difference between 'pictorial expression', meaning expression by (whole) pictures, and 'skilled depictions of emotions', including expression by figures, scenes and designs; (2) an ambiguity

---

[1] Brassey 2019, pp. 15-29.
[2] Robinson 2017.
[3] Lopes 2005.

in the notion of 'design expression'; and (3) the idea of an *implied painter* as the locus of expression.

(1) Brassey characterises my view of pictorial expression as a 'personalist' view, meaning that for me 'pictorial expression' is always *transitive*, i.e., pictorial expression is always an expression of someone's actual mental state or disposition, whereas according to Lopes' 'impersonalist' view, expression in pictures takes the form of what he calls figure, scene, or design expression. Figure expression is "an expression that is wholly attributable to a depicted person or persons"[4]; scene expression is "an expression that is attributable at least in part to a depicted scene and is not wholly attributable to any depicted person"[5]; and design expression is "an expression that is wholly attributable to a picture's design or surface and not to any figure or scene it depicts".[6] In contrast, I argue that although figure and scene expression can contribute to the expression in a (whole) picture, they are not themselves expressions of emotion, but merely, in Brassey's phrase, "skilled depictions" of figures and scenes *as* (looking) sad, anxious, light-hearted etc.[7] A genuine pictorial expression of emotion, on the other hand, is a genuine expression of emotion in a person, normally the painter.

Brassey comments that Lopes seems to think that I "[assume] that viewers sustain an FE figure expression line of thought in the absence of figures", and that once we drop this assumption—which we should, on Lopes' view—we can "adopt an impersonal theory", such as Lopes endorses.[8] It is not entirely clear to me what Brassey means by a "figure expression line of thought"[9], but I think she is accusing Lopes of interpreting me 'qua personalist', to be taking 'figure expression', i.e. the expression of emotion by depicted figures,

---

[4] Lopes 2005, p. 51.

[5] Ibid., p. 52.

[6] Ibid., p. 57.

[7] Brassey 2019, p. 18.

[8] Ibid., p. 20.

[9] Ibid.

as in Daumier's 'Fatherly Discipline', to be "the paradigm case of pictorial expression"[10], and that's why for me there is a missing person problem when it comes to scene expression (and presumably design expression too). Figure expression can be transitive insofar as a depicted character in a painting wears an expression of emotion that reflects the emotion the character is depicted as actually experiencing: if, in Daumier's 'Fatherly Discipline', the father is not expressing his own actual state of irritation but simply putting on an irritated face, the painting loses its point.

But Brassey perspicaciously points out that pictorial expression on my view is *not* modeled on figure expression. On the contrary: figure, scene, and design expression are all what I call 'expressive elements' in a picture: they contribute to what the picture expresses *as a whole*, which is what I mean by 'pictorial expression'. As she writes, "Lopes' error is to conflate his categories of expression"—figure, scene, and design expression—"with a picture's expression of an emotion E".[11] Pictorial expression is indeed transitive on my view but not because it is a version of figure expression. As she rightly points out, "personalists do not argue that simply seeing depicted figures expressing emotions is sufficient for seeing a picture as an act of expression".[12] She notes that a row of smiling faces generated by the Flickr algorithm is a good example of what's wrong with this claim. What a picture expresses *as a whole* is a function of what the figures in the picture express (in Lopes' sense), what the scene or scenes in the picture express (in his sense) and, above all, *how* the depicted figures and the depicted scenes are unified into an expressive whole by the overall design or composition of the picture. In the Kokoschka self-portrait that I discuss, for example, there is a semi-realistic depiction of Kokoschka's anxious facial expression, posture and gesture (holding his hand to his mouth). This is what Lopes calls figure expression, but that's not what gives the *picture* its

---

[10] Ibid.

[11] Ibid., p. 21.

[12] Ibid.

expression. There are also *design elements* that reinforce the expression of anxiety in the depicted Kokoschka: The "agitated passages of paint, the awkward perspective and the vague, unstable use of space" are what link up the depicted figure with the expression of anxiety by the picture as a whole.[13]

This is a particularly interesting example insofar as the depicted figure is the same person as the painter: in painting his own self-portrait in this fashion, Kokoschka paints his gesture, his posture, his facial expression and so on in such a way that the picture as a whole expresses his own feelings of insecurity and anxiety. In R. G. Collingwood's terms, his painting explores and makes concrete the emotions he is experiencing and thereby allows the viewer to feel what it is like to be the person depicted in the picture.[14] Lopes' three types of (what he calls) 'expression' come together in the picture to express the emotion of the painter, in the rich Romantic way that I espouse, which in this case is also the emotion of the person depicted.

(2) Brassey thinks there is "an unresolved tension" in my response to Lopes, which is illustrated by my description of Kokoschka's 'Self Portrait with Hand to Mouth'.[15] I say that the painting depicts Kokoschka as "looking anxious and insecure" ('figure expression' à la Lopes), but that the picture as a whole also "conveys anxiety and insecurity in the agitated passages of paint, the awkward perspective and the vague, unstable use of space", which I describe as 'design expression' (à la Lopes).[16] Brassey suggests this means that I think that design expression is "solely responsible for giving us" expression by the implied painter of the whole painting), whereas, as we just saw, I also say that figure, scene *and design expression* are

[13] Robinson 2017, p. 262.
[14] See Collingwood 1963.
[15] Brassey 2019, p. 22.
[16] Robinson 2017, p. 262.

only 'expressive elements', not genuine artistic expressions of emotion.[17]

Brassey has picked out an important ambiguity in my view. I should have been more careful when I talked about how a painting is unified by its 'design'. The trouble stems from Lopes' narrow definition of design expression as "an expression that is wholly attributable to a picture's design or surface and not to any figure or scene it depicts".[18] He cites Jackson Pollock's non-representational drip paintings as an example. On Lopes' view, it seems that figure, scene, and design expression are *independent modes* of what he calls 'pictorial expression'. In 'Fatherly Discipline', for example, he claims that "the narrative comprises nothing but what the figures express"[19] but that certain of its design elements—notably the "coiled lines used to render the faces and hands of father and child"[20]—reinforce figure expression. Likewise, it is true that the Kokoschka painting has 'design elements', namely colour, brushwork, and use of perspective, that reinforce its 'figure expression'. But my notion of 'design' is not as a set of design or formal elements, such as colour, line and use of space, but rather the way the whole painting expresses an emotion by means of its *composition*. By this I mean, using Lopes' terms, that it is the way that the figures, scenes, and design—the 'expressive elements'—in a painting work together to express an overall emotional state that unifies the painting.

It is only in this sense that design expression is "solely responsible for giving us" expression by the implied painter of the whole painting. [21] I am grateful to Brassey for pointing out this interesting ambiguity which has large consequences for my view and helps to indicate why I think it is preferable to Lopes' view. Indeed, I am reminded of my long ago paper on 'Style and Personality in the

---

[17] Brassey 2019, p. 22.

[18] Lopes 2005, p. 57.

[19] Ibid., p. 55.

[20] Ibid., p. 57.

[21] Brassey, 2019, p. 22.

Literary work', in which I argued that artistic expression is expression in an artist or implied artist, and that 'style elements'—like expressive elements—cannot even be detected as such until we know what the overall artwork expresses. The rough brushwork in Kokoschka's self-portrait is part of its overall expression, for example, whereas the rough brushwork in my portrait of myself is simply the result of incompetence.

Brassey's point is at the heart of my view that pictorial works of art can sometimes express emotions in the way described by R.G. Collingwood: they do not merely depict or describe emotions; they explore and articulate the artist's emotions in such a way that viewers can recreate for themselves the emotions expressed.[22] To conclude this section: at the beginning of her paper Brassey invites me to clarify "the relation between (a) what a picture expresses and (b) what is depicted as expressing in the picture" but she has done a good job of doing this herself.[23]

(3) Finally, I come to perhaps the most contentious issue raised by my view, which is the idea that, although the emotions expressed by a picture as a whole are typically those of the artist who expresses his or her emotions through the picture, in cases where this seems implausible, i.e., where the personality expressed in the work does not match the personality of the artist in daily life—Brassey cites "Tolstoy's ability to write with compassion while lacking compassion in his personal relationships"—the emotions expressed should be attributed to an *implied* painter, the painter as he or she seems to be from the evidence of the painting in its historical and cultural context.[24] This idea raises hackles partly because the implied painter isn't actually a painter at all and lacks any 'psychological reality', as Richard Wollheim once put the point.

---

[22] See Collingwood 1963.

[23] Brassey 2019, p. 15.

[24] Ibid., p. 25-26.

My main goal in the fairly long discussion of this issue in 'The Missing Person Found: Part One'[25] was to question those who put so much weight on the possibility of the implied author being different from the actual author. Broadly speaking, I wanted to point out that there is a much more intimate connection between actual and implied author than this objection assumes. On the other hand, it is no easy matter to describe exactly what the relationship is and in any case my suspicion is that different paintings and different painters will need to be treated differently in this respect. I acknowledge that implied painters may be suspect in some cases, because the apparent author of a painting seems to have so little in common psychologically with the actual author, but I guess I would want to insist that in most cases extreme skepticism is not warranted. Even if the implied author is not identical to the actual author, they typically have a great deal in common psychologically.

Lopes, of course, wants nothing to do with implied authors (or even authors!), since his is an 'impersonalist' account of expression, but Brassey points out that this has unfortunate consequences for him insofar as his "categories [of expression] can only account for meaning derived from the formal (visible) qualities of the picture surface", so that "meaning is restricted to what goes on *inside* the represented pictorial world".[26] In contrast, Brassey points out that if we accept the idea of an implied painter, we can make sense of the "different levels at which we understand a picture".[27] Brassey explains her idea by reference to Edward Hopper's "'Nighthawks' which she says illustrates how "a naïve painterly style" is "put to use by a mature, skilled artist who adopts a naïve style".[28] She cites Kendall Walton who "suggests that this kind of multi-level impression of an implied painter operates at a 'deeper level' than just taking it that the work is painted by a naïve illustrator and can lead

---

[25] Robinson 2017.

[26] Brassey 2019, p. 23.

[27] Ibid., p. 22.

[28] Ibid., pp. 22-23.

theoretically to the viewer experiencing a 'richness and elusiveness of style' in some paintings that 'derives from an obscure partial awareness of a multi-level situation of this kind'".[29] The point is a nice one and interestingly it once again links the discussion of expression to the notion of style. I am very happy to accept this idea.[30]

2

The rest of the papers are about literature and film. I turn first to the paper by Mary Beth Willard on the paradox of fiction, which she describes as a commitment to the following three mutually inconsistent propositions.[31]

1. We have genuine and rational emotional responses towards fictional characters and situations.

2. We believe that fictional characters and situations do not exist in reality.

3. In order for us to have genuine and rational emotions towards a given character or situation, we must believe that they exist in reality.[32]

In 'Deeper than Reason' I gave a somewhat dismissive account of the paradox of fiction: given my view of what emotions are and how they operate, I argued that the so-called problem of the paradox of fiction dissolves. If emotions are triggered by people, events and situations which are of great significance to our survival and/or well-being, then our emotion systems respond to anything that seems significant in this way. The trick to getting readers to respond emotionally to works of literary fiction is to get readers interested in, curious about, and

---

[29] Brassey 2019, p. 23; Walton 1976, p. 52.

[30] See Walton 1976.

[31] Willard 2019, pp. 30-44.

[32] Ibid., pp. 30-31.

often sympathetic towards the characters and to get personally involved in their problems and difficulties. We feel sympathy for Anna Karenina and that's why we weep for her even though we know she doesn't exist. Similarly, we respond emotionally to what happens to characters on-screen in fictional movies.[33] What I say is that "[i]t does not matter to my emotion systems (fear, sadness, etc.) whether I am responding to the real, the merely imagined, the possible or the impossible".[34] Always I respond very fast and 'instinctively' to something fearsome, offensive, contaminated or whatever, and subsequent cognitive monitoring—including realising that the object of my emotion is imaginary, for example—modifies my response so that my incipient action tendency may not issue in any actual action.

Willard's main focus is on whether emotional responses to fiction are *irrational*.[35] My own view is that, strictly speaking, feeling emotions for fictional characters and events *is* irrational, even though it happens all the time. But Willard claims that I "shouldn't concede irrationality, or at least not on the grounds that the characters do not exist", because "existence turns out *not to matter* all that much to whether [...] our emotional responses toward fiction are rational".[36] This may be true in the sense that the existence of fictional characters and events is not germane to how we respond emotionally to them qua fictions. Often we respond to them *as if* they existed, even if we know they don't (and didn't). But as a philosopher, rather

---

[33] Willard says that "on Robinson's view, becoming emotionally involved in a movie requires nothing more than instinctively responding to the visual stimuli" (p. 32.). This makes it sound as if I think movie-watching is nothing but paying attention to changing patterns of light and colour, which is not what I mean. The "visual stimuli" are the filmed characters, scenes and situations.

[34] Robinson 2005, p. 145; Robinson 2010, p. 86.

[15] She also talks a little about whether they are genuine. I assume she is referring to the debate about whether the emotions evoked by fictions are genuine insofar as they meet the same conditions as 'life' emotions, as opposed to the quasi-emotions posited by Kendall Walton and others.

[36] Willard 2019, p. 32 (my italics).

than as a reader of fictions, we can still ask whether such emotional responses are rational or not, and on the face of it, it does seem *cognitively* irrational to believe at one and the same time that 'Anna Karenina is having a hard time' and that 'Anna Karenina does not exist'.[37]

The 2010 paper that Willard cites was 'adapted and abbreviated' from 'Deeper than Reason', chapters 4-5.[38] One section of chapter 5, which I omitted from the 2010 paper, is headed 'Is it Irrational to Respond Emotionally to Fictions?' and there I argue that, although from "a strictly cognitive point of view it is irrational to have wants and goals with respect to Anna Karenina and her ilk, it is not 'emotionally irrational' and it is certainly not maladaptive".[39] Here I am drawing on the work of Patricia Greenspan, who argues that in a case of mixed feelings, when my good friend has won a prize that I would have liked to win myself, it is in her terms 'basically rational' or 'emotionally rational' to feel both happiness and unhappiness, and that feeling *both* emotions, even though this would be inconsistent on strictly cognitive grounds (given the way Greenspan has set up the case), "on a standard of rationality that evaluates emotions according to their behavioral consequences— which takes into account, for instance, the social value of identification with others—ambivalence might sometimes be *more* rational than forming an 'all things considered' emotion that resolves

---

[37] In deciding whether emotions in general or emotions for fictions in particular are rational or irrational, much depends on what your conception of rationality is. In what is still one of the best extended discussions of 'the rationality of emotion' Ronald de Sousa, in his eponymous book, distinguishes rationality of belief from rationality of desire and both from rationality of emotion. *Cognitive* rationality aims at truth. Strategic rationality—or rationality of desire—aims at goodness. Emotional (or axiological) rationality depends upon the 'paradigm scenarios' within which a given emotion term was learned. Fear is rational if the situation I am in is sufficiently like the fearsome situations in which I learned the meaning of 'fear'.

[38] Robinson 2005.

[39] Ibid., p. 147.

the conflict".[40] I grant that emotions can be rational in Greenspan's sense, which is more-or-less equivalent to 'adaptive', without being cognitively rational.

According to my view of emotion, emotions are *processes*. Emotional responses are bodily responses triggered by crude appraisals [e.g. This is GOOD or BAD for me] that sometimes operate on simple stimuli, at other times on cognitively complex stimuli. Suppose I am frightened by the possibility of another recession. In this case, my fear presupposes that I have some cognitively complex beliefs or thoughts about economics. However, what makes my reaction a *fear* reaction is its registration of THREAT in bodily changes of a fairly specific sort (including physiological and behavioural changes, states of action readiness and so on), which fasten my attention on the fact of a threat. If this is right, then the bodily registration of THREAT occurs *prior* to any cognitive assessment of the economic situation, and cognitively it is *arational*. However, after the crude 'bodily appraisal' of THREAT, there is invariably 'cognitive monitoring' of the situation, which is itself subject to norms of rationality.[41]

However, none of this is really germane to Willard's main point, which is that "whether our emotional responses towards fiction are rational depends on our compliance with [...] the *norms of aesthetic engagement*" and that "the paradox of fictional emotions has always, at its core, been a question of aesthetic appraisal".[42] Interestingly, when Colin Radford published the paper that started the paradox of fiction industry, 'How Can We be Moved by the Fate of Anna Karenina?'[43] the companion piece to his, in a supplementary volume of the 'Proceedings of the Aristotelian Society' for 1975, was a paper by Michael Weston in which he points out that "it is the

---

[40] Greenspan 1980, p. 238.

[41] For an explication and defence of my evolving view of emotion, see Robinson 2004, Robinson 2005, and Robinson 2018.

[42] Willard 2019, p. 32.

[43] Radford 1975.

recognition of Mercutio as part of a work of art that I find missing in Doctor Radford's treatment".[44] The idea is similar in spirit to Dr Willard's. As she says, most discussions of the paradox of fiction have followed Radford in adopting the assumption that "if it is irrational to weep over purely imagined states of affairs, then it is also irrational to respond emotionally to fictional characters and situations, which after all are imagined states of affairs, the production of which has been outsourced to an artist".[45] My weeping about a possible but non-actual situation in which my fit and healthy husband has died of a wasting disease is 'on a par' with my weeping for Anna Karenina's tragic suicide.

Willard argues that this is a mistake. She perceptively notes that the crucial difference between these two cases is that there are *norms of aesthetic engagement* governing our interactions with 'Anna Karenina'—and hence also, presumably, with Anna Karenina, whereas there are no such norms governing our counterfactual imaginings about our friends and relations. This is an important and often neglected point, and Willard is quite right to emphasise the essential role that aesthetic norms play in the debate about the rationality of responding emotionally to fictions: what it's rational to feel about Anna Karenina depends a very great deal on her role in the fiction in which she occurs, something that has no application when I'm musing sentimentally about my husband.[46]

However, I have some doubts about the details of Willard's proposal. She describes norms of aesthetic engagement as norms governing "the responsibilities appreciators of art incur when they engage works of art"[47] and suggests that they include "norms that govern the range of licensed emotional responses to the artwork".[48]

---

[44] Weston 1975, p. 84.

[45] Willard 2019, p. 33.

[46] In the discussion of Konrad et al I note the double nature of characters as both human beings and elements in a design.

[47] Willard 2019, p. 34.

[48] Ibid., p. 32.

"Roughly, [these norms] say, 'to experience the work correctly, appreciators should respond in this way'".[49] The two aesthetic norms most relevant to her discussion are, she says, "(1) Being open to engaging with fiction", and "(2) Exhibiting the appropriate emotional response".[50]

The first proposed norm seems innocuous, but I think there are problems with the second. The norm "Exhibit appropriate emotional responses [to a fiction]" is meant to rule out such inappropriate responses as laughing at 'Schindler's List', failing to be shocked by the violence in a "gritty film noir"[51]; or responding with terror to any dramatic tension in a work (as in her young son's reaction to some action movies). As she rightly points out, none of these responses will be blocked by the assumption that the characters and situations in question don't exist. The problem, however, is a logical one. One of Willard's main points is that "only emotions that are sanctioned by the norms of aesthetic engagement are appropriate to have toward fictional characters and situations".[52] However, the norms she is referring to here include the norm "Exhibit appropriate emotional responses [to a fiction]".[53] The claim is tautological: "only emotions appropriate to fictional characters and situations are sanctioned by the norm: 'exhibit appropriate emotional responses to fictional characters and situations'".

Another problem is that there is rarely one and only one appropriate emotional response to a particular fictional characters or situation.[54] I defend a 'reader-response theory' of literary interpretation which allows for diverse interpretations partially as a

---

[49] Ibid., p. 34.

[50] Ibid.

[51] Ibid., p.33.

[52] Ibid., p. 41.

[53] Ibid., p. 35.

[54] What counts as an 'appropriate' response is an issue I return to in the discussion of Quixote Vassilakis' paper.

result of the different emotional responses different readers have to the characters and events in a novel.

3

Next I discuss the papers by Eva-Maria Konrad, Thomas Petraschka, and Christiana Werner (henceforth Konrad et al)[55] and by Quixote Vassilakis[56], both of which address my claim that appropriate emotional responses to realist novels are important to understanding and interpreting (those kinds of) novels. Konrad et al are happy to accept this relatively weak claim, but they reject my stronger and more controversial claim that in some cases a novel "cannot be adequately understood" without emotional involvement with the characters and events of the novel.[57] In their elegant and well-argued paper, they question three of my arguments, (1) 'the trigger argument', namely, that out of all the information that a novel provides, it is often our emotional responses that make *salient* the information most important to understanding the novel; (2) 'the empathy argument', namely, that "understanding character is relevantly like understanding real people, and that understanding real people is impossible without emotional engagement with them and their predicaments"[58]; and (3) 'the data argument', namely, that "we use our emotional responses towards [characters] as data in arriving at an interpretation of [a] character".[59]

(1) They agree with me that readers's emotional responses are often 'sources of salience' that trigger recognition of particularly important events or scenes in a novel, but deny that "being triggered by [one's] emotional reactions" is the only way for a reader to

---

[55] Konrad et al 2019, pp. 45-59.

[56] Vassilakis 2019, pp. 60-73.

[57] Robinson 2005, p. 107; Konrad et al 2019, pp. 46 & 49.

[58] Robinson 2005, p. 126; Robinson 2010, p. 78.

[59] Robinson 2005, p. 116; Konrad et al 2019, p. 55.

become aware of significant episodes in a text".[60] They are surely right about this. In particular, as they point out, there are structural features of novels and plays that also help readers pick out the most significant events and scenes. They note that a crux in a play might occur in the very middle of the play, for example, as in the meeting between Mary and Elizabeth in Schiller's 'Maria Stuart'. And very often there is what has been called a 'tension arc' in a play or novel, making it more likely that tension will gather over the first part and reach a climax in the middle or towards the end of the play or novel. There is no need for our emotions to get involved.

On the other hand, a tension arc is designed to arouse our curiosity and feelings of suspense as well as the relief that follows the release of tension. And as for the meeting between Mary and Elizabeth we have been primed to expect it as an emotional climax in the play. So the emotions might be playing a role even when it appears that purely formal considerations are what determine salience. However, I agree with Konrad et al that it is unlikely that emotional triggers are the *sole* sources of salience in a literary work.

(2) More controversially, I argue that understanding characters in a realist novel is *necessary* to understanding the great realist novels of F. R. Leavis' Great Tradition, that "understanding character is [...] relevantly like understanding real people", and that "understanding real people is impossible without emotional engagement with them and their predicaments".[61] Konrad et al focus on readers's *empathic* emotions for characters, and they have an ingenious argument designed to show that emotional responses to fictional characters are very different from emotional responses to people in real life. They suggest that on the one hand we might adopt a realist theory of character, in which case characters really exist but only as constructs of words, and on the other hand for anti-realists about characters, fictional characters are not real at all. Whichever alternative we adopt, however, it turns out that fictional characters

[60] Konrad et al 2019, p. 55.
[61] Robinson 2005, p. 126.

lack mental states: constructs of words don't have mental states and neither do phantom 'characters'. Hence interacting with characters is radically unlike interacting with real people.

In response, I would point out that engaging with a realist novel involves treating characters as having two aspects: we are encouraged to treat them as real people and simultaneously as constructs of words fitting into a larger structure. In 'The Philosophy of Literature', Peter Lamarque talks about viewing characters from two perspectives, the internal perspective from which they are (mostly fictional) people in a set of situations, and the external perspective in which they are elements of the structure of the work.[62] What I'd like to emphasise is that if we did not respond to characters *as if* they were actual people, the point of a novel would be lost. If characters were *nothing but* constructs of words, they would not appeal to us emotionally nor would they get us engaged in issues of love, death, remorse, or any of the other great themes of literary novels. So although I would plump for the view that fictional characters are indeed 'not real' in a metaphysical sense, nevertheless, novel-reading, at least where the great realist works of fiction are concerned, requires that we treat them *as if* they were real.[63]

(3) Finally, the data argument.[64] Konrad et al object to my claim that certain data or information about characters can be gleaned only through our emotional reactions to those characters. My examples include *gentle amusement* for Lambert Strether, the main protagonist of James' 'The Ambassadors' and *compassion* for Anna Karenina. As in their first argument, Konrad et al argue that there are other ways in which we can learn that Strether is mildly amusing and that Anna is in a desperate situation. This may well be

---

[62] Lamarque 2009, pp. 201-202.

[63] In Robinson 2005, chapter 4, I defend the view that we should respond to fictional characters as if they were human beings, and in chapter 7 I draw attention to the 'double function' of both plot and character, e.g. p. 215. Both, I argue, are also "important [principles] of *structure* or form" (ibid.).

[64] Konrad et al 2019, pp. 54-57.

true, but I think there is an important difference between 'understanding' in the abstract, and understanding with your gut, as it were. You can tell me that Strether is mildly amusing, and in the past maybe your views on such matters have been reliable, but unless I feel amused myself by Strether, why should I believe you that he is in fact amusing? Moreover, I won't really understand either Strether or Anna as 'people' unless I approach them *sympathetically*. Otherwise, Strether may seem like nothing but a fuddy-duddy and Anna may come across as nothing but a foolish narcissist. In other words, treating characters as if they were real is not sufficient for understanding them as we understand our close friends: care and concern may also be necessary. Konrad et al also argue that empathic responses to characters will provide new data to a reader only when that reader has never previously experienced the emotion in question. If I have suffered as Anna Karenina suffers in the eponymous novel, I can simply recall my own past experience in order to understand Anna's present experience. They therefore argue that "empathic emotions are only necessary for those empathisers who have not felt the affective state in question before".[65] I would like to respond in two ways to this claim. First, one of the things that novels can do is to broaden our emotional repertoires and to get us to feel in ways we have never felt before. Anna Karenina's situation is described in vivid, painful detail in Tolstoy's novel. Even if we remember having gone through a difficult divorce or being abandoned by a lover or having lost a child in a custody dispute, our remembered feeling cannot be identical to Anna's because our situation in the contemporary Western world is so different from hers. Hence the data we glean from our own memories about what it's like to be in a state of despair about such things may fail to illuminate Anna's very individual situation. Secondly, as a matter of empirical fact, we have great difficulty *remembering feeling* a certain way in the past. It is much easier to remember *that* such and such

---

[65] Ibid., p. 52.

happened to me. (Luckily, mothers typically forget what it felt like to give birth, for example.) Hence on more than one count it is unrealistic to think that recalling my own experiences will be a good source of data collection when I try to understand what a character is going through.

Quixote Vassilakis also takes me to task for overstating my case for the importance of emotional engagement to the understanding of literary works, especially given my understanding of what an emotional response *is*, namely, a response to a perception or thought of something of great significance to me or mine that is *registered in bodily changes*. But Vassilakis takes his argument in a different direction from Konrad et al. He is focused mainly on the possibility of variation among acceptable interpretations. He notes the ubiquity of individual differences in readers's responses to narratives as well as the ever-present possibility of inappropriate emotional responses to narratives.

(1) Like Konrad et al, Vassilakis argues that it's perfectly possible to understand another person's emotional states—whether actual or fictional—without any kind of emotional bodily response, but, unlike Konrad et al, he bases his argument on the presumed fact that there are certain individuals—namely, psychopaths—who have a "severe emotional deficit" but also a "potentially 'enhanced' [...] ability to understand the emotional cues of others".[66] (2) He makes the perceptive point that "bias is built into emotion" and that it can lead to misunderstanding and even confabulation about how a narrative unfolds.[67] (3) Finally, he points out that readers can and do have inappropriate responses to fictional worlds and claims that I "cast reader error as *incidental and largely correctable*" by after-the-fact cognitive monitoring, whereas in fact the ubiquity and persistence of idiosyncratic emotional response is a central feature,

---

[66] Vassilakis 2019, pp. 61 & 65.

[67] Ibid., p. 68.

not a rare anomaly, of emotions.[68] These are important and interesting criticisms which I will now attempt to answer.

(1) The argument from psychopathy strikes me as odd. First, psychopaths appear to figure out other people's emotions by some kind of calculation. In this respect they are like autistic people who can tell that someone is (probably) happy by the fact that the sides of their mouth are pushed up, not because of any empathic emotional response to a smile. But, although psychopaths may be better than the rest of us at divining clues to a person's mental states from the way they look and act (although I think the data are not clear on this point), in general they are hopeless at figuring out what other people feel, because they are very self-centered and totally lacking in empathy. More importantly, unlike (most) autistic people, psychopaths are stunningly lacking in *care and concern* for other people. Moreover, whatever the truth about psychopathy, it is clear that novels are not by and large written for psychopaths—a comparatively small group of very strange and scary people—and the existence of psychopaths does not seem to me to be relevant to principles about how to read a novel in a normal way.

(2) Vassilakis rightly points out that "bias is built into emotion" and he perceptively points out that one reason why is that "emotion necessarily elicits action tendencies, which can be difficult to reverse—or even to be aware of", something he thinks I should emphasise more.[69] This is also true of other physiological reactions that comprise emotional reactions. He rightly realises that this creates difficulties for theories of interpretation that stress emotional engagement with a work, especially those that advocate *immersing oneself* in a novel and becoming so "deeply emotionally involved with a novel" that they "construe fictional worlds as real", at least—I assume—for the period while they are reading.[70] If readers are too immersed—are completely absorbed in reading from the internal

---

[68] Ibid., p. 69 (my italics).
[69] Ibid., p. 68.
[70] Ibid.

perspective—their interpretations may be warped. I agree. As I mentioned in the discussion of Konrad et al, interpretations of literary narratives require taking an external perspective as well as an internal perspective and going back and forth between the two. All this seems to me very sensible.

(3) Finally, Vassilakis presses me about how I deal with inappropriate emotional responses. He claims that on my view "readers must respond to certain great literary works in accordance with particularly-defined emotional patterns in order to properly value them".[71] But this is not what I say. As I observed earlier, I defend a reader-response theory of literary interpretation, according to which different readers may arrive at different but equally coherent interpretations of a novel or other literary work. In chapter 4 of 'Deeper Than Reason', I remind my readers of Monroe Beardsley's principles of congruence and plenitude, which provide criteria of correctness for interpretations i.e., "make your interpretation of the parts of the text consistent" and "make your interpretation fit as much of the text as possible".[72] These are very broad criteria, however, and it is not hard to see how they can be met by quite different interpretations.

In general, I argue that when we are talking about great writers and their great works, the best interpretations are going to be those that adhere most closely to what the author probably intended. However, I was careful to point out (in chapter 6) that what sort of person the implied author is and what they probably intended is in the last resort up to the reader. 'The meaning' of a novel is determined by the *interaction* of author and reader. So I reject the charge that in my view 'reader error' is "incidental and largely correctable".[73] A difference of opinion about a small passage can lead to major differences in the interpretations of two different but equally sensitive readers. Vassilakis is quite right to chide me for not

---

[71] Ibid., p. 61.

[72] Robinson 2005, pp. 428-9.

[73] Vassilakis 2019, p. 69.

emphasising this enough. He is also right to highlight "the ubiquity and persistence of idiosyncratic emotional response" as "a central feature, not a rare anomaly, of emotions".[74] All I can say is that I heartily agree.

4

Finally, I come to Irene Martínez Marín's[75] paper, on self-conscious emotions. Martínez Marín attacks my Jamesian theory of emotion, which claims that bodily changes are a *sine qua non* of episodes of emotion. She goes to the heart of the Jamesian position when she takes me to task for failing to distinguish between *being emotional* and *having* or experiencing *an emotion*. For me, all emotion episodes, whether initiated by a fleeting perception of a rattlesnake or a considered judgment that I'll be bankrupt in six months, involve bodily changes. This is the crucial difference, it seems to me, between a *mere* judgment or thought and an emotion. However, I do acknowledge explicitly that emotion *words* are sometimes used to refer not to a genuine emotion but only to a related judgment or thought: "I am very proud of you" may simply report my approval of you, without any implication that I am currently in an emotional state of pride.[76] Similarly, if the shop assistant says "I regret that the jacket is not available in red", we do not think that he is reporting on his current emotional state.

Martínez Marín wants to make a sharp distinction between "higher cognitive emotions" or "non-primary or intellectual emotions", such as "guilt, shame, embarrassment, pride, or nostalgia", and "basic emotions", such as anger and fear, because she thinks that, unlike the latter, the former do *not* require any bodily involvement.[77]

---

[74] Ibid.

[75] Martínez Marín 2019, pp. 74-94.

[76] Compare Bedford 1956-7.

[77] Ibid., p. 73.

Consequently, she thinks that according to my view "higher cognitive emotions [cannot have] the status of really being emotions".[78] But this is not my view. For me, 'higher cognitive emotions' require bodily involvement in order to be genuine emotions, just like any other emotions. Here too I follow James, who argues that even in the case of 'subtle emotions' such as aesthetic emotions, 'the bodily sounding-board' is at work. There are no "pure cerebral emotions", he says.

> Unless in them there actually be coupled with the intellectual feeling a bodily reverberation of some kind, unless we actually laugh at the neatness of the mechanical device, thrill at the justice of the act, or tingle at the perfection of the musical form, our mental condition is more allied to a judgment of right than to anything else. And such a judgment is [...] a cognitive act.[79]

He goes on:

> But as a matter of fact the intellectual feeling hardly ever does exist thus unaccompanied. The bodily sounding-board is at work, as careful introspection will show, far more than we usually suppose.[80]

He thinks we should distinguish between "the keen perception of certain relations being right or not", and "the emotional flush and thrill consequent there upon", and claims that "these are two things, not one".[81] This seems to me perfectly plausible.

---

[78] Ibid., p. 80.

[79] James 1884, pp. 201-202.

[80] Ibid.

[81] Ibid. I suspect that Martínez Marín has misread James and perhaps also me on this point. James does not claim, as she accuses him of doing, that the main difference between 'standard emotions' and 'non-standard' or 'intellectual'

However, this response, I think, ignores Martínez Marín's most pressing concerns. To me her most interesting contribution is the challenge to my view she thinks is represented by the 'film-diaries' of Jonas Mekas. (1) She worries that "[my] model seems to be excluding from the process of art appreciation a set of emotions that are familiar to us in life and the arts" namely, "emotions that are *partly constituted* by cognitive processes that unfold over time and which require attention to the personal history of a person", such as "nostalgia, [...] grief, shame, envy, [and] pride", and that have no obvious bodily signature.[82] She also argues that "the nostalgic view from which Mekas' film-diaries are narrated cannot be grasped *empathically* because it is impossible to "[share] the particular object of the nostalgic emotion expressed in these experimental films".[83] (2) Nostalgia, in Martínez Marín's evocative phrase, is "an emotion concerned with the irrevocability of one's own past".[84] It involves not simply remembering my past life but *reflecting* upon it, so that in this sense nostalgia is indeed "partly constituted" by reflections.[85] But my question would be: why call nostalgia an *emotion* if it is primarily a set of reflections? What I have consistently claimed in my work is that *long-term* emotions—what Goldie might call 'narrative'

---

emotions, is that "the bodily sounding-board, vibrating in the one case, is in the other mute" (p. 80). He is claiming that the bodily sounding-board is *always* "vibrating" in a state of emotion, whether "standard" or "non-standard". It is 'mute', however, for the critic who has become jaded and incapable of the "flush and thrill" of genuine aesthetic delight. On the one hand there is the critic who is delighted by the elegance of the proof or 'flushed and thrilled' by the beauty of the music, and on the other hand the jaded critic who *recognises* the beauty perhaps, but does not feel the emotional flush or thrill that he did in his youth.

[82] Martínez Marín 2019, p. 80 (my italics). She thinks that such emotions are better accounted for on Peter Goldie's theory of emotions. See Goldie 2000. I do not have space to defend this claim, but I have always found Goldie's key concept of 'feelings towards' as very unclear.

[83] Ibid., p. 76.

[84] Ibid., p. 75.

[85] Ibid., p. 84.

emotions[86]—are not properly regarded as *emotions*, rather than *mere* judgments or thoughts*, unless they involve episodes of emotion with bodily components*. It seems to me that if I am in an emotional state of nostalgia my reflections about my past will themselves evoke emotions: very roughly, happiness at the memory of something loved, and sadness at the fact that this happiness is irrecoverable. Nostalgia is a *bittersweet* emotion.[87]

As for bodily symptoms of these emotions, as far as I know, it is not clear whether or not there are specific bodily changes in nostalgia,[88] but some of Martínez Marín's other examples clearly do have bodily symptoms: shame and guilt involve withdrawal or hiding behaviours, for example, and grief is expressed in characteristic postures, gestures, vocal and facial expressions. Yet I am loath to say that the long-term emotion of nostalgia is *nothing but* a disposition to experience episodes of happiness about the past mixed with sadness about the past. As I am sure Martínez Marín would agree, this description does not seem to capture the *phenomenology* of long-term nostalgia.

My (untutored) impression is that Mekas is trying to create a *mood* of nostalgia, by creating films that are expressive of nostalgia and inviting spectators to *feel* nostalgia for their own pasts as they watch these flickering images. We can trace these effects to formal and expressive aspects of the films such as their 'amateur' appearance, characteristic of home movies, the unclear temporal sequencing, their impressionistic, dreamy quality, the lack of developed characters, and, more generally, the lack of narrative

---

[86] For Goldie's view of emotions as having a narrative character, see Goldie 2000, especially chapters 1 and 2.

[87] Notice that a *long-term* emotion is not the same as a *higher cognitive* emotion: I can have a brief surge of pride for my son that lasts a few seconds as well as a long-term emotional state of fear or anger such as fear of snakes or anger about Brexit.

[88] Both happiness and sadness have been studied, but it seems unlikely that nostalgia would get half its bodily symptoms (say) from happiness and the other half from sadness: this would pose a problem if only because the two sets of bodily changes seem to be incompatible.

drive. Martínez Marín has identified a way in which art and emotion interrelate that is different from anything I have written about (although I have things to say about *music* that may be relevant) and I am grateful to her for introducing me to these haunting films. Clearly, they raise issues that I have not thought about enough.

(3) Martínez Marín also suggests that 'affective empathy' for the narrator of Mekas' diary-films may not be possible because we can never "[access] the particular object of another's self-conscious emotion".[89] For example, in nostalgia, "we fail to intentionally imagine the other's recollections in an empathetic way because the yearning associated with this emotion is inseparable from the subject's self-evaluation of her own past".[90] This may be true, but it's also true that empathy is always a matter of degree, and although it may be hard for me to recreate in imagination Mekas' nostalgia for the village of his youth, other attempts at empathy for situations closer to home are equally difficult to achieve. William Ickes in particular has shown that empathic accuracy is, in general, just not very great.[91] Recall, too, my discussion of Konrad et al. on the relevance of my memories to my ability to empathise with another person.

Finally, I cannot help ending this comment on a plaintive note. Throughout 'Deeper than Reason' and most of my other work on art and emotion, I explicitly say that I am focusing on canonical and/or *realist* works in the various arts: 'Anna Karenina', 'The Raft of the Medusa', Beethoven's symphonies etc., and that I am precisely *not* discussing avant-garde, experimental, post-modern or other non-realist works.[92] Mekas is a very interesting case, but not the sort of thing that I was aiming to analyse. What Martínez Marín's discussion reveals is that there is a lacuna in my view—how should we think of long-term higher cognitive emotions? —that I might have recognised

---

[89] Martínez Marín 2019, p. 86.

[90] Ibid.

[91] See e.g. Ickes 1997.

[92] And I barely mention movies!

and tried to fill. But there again, the book is already 500 pages long, and, more importantly, my critics have demonstrated that they and no doubt others are well-equipped to take on some of these issues for themselves!

Thank you again to all my critics for really interesting and perceptive papers.

robinsjm@ucmail.uc.edu

ABOUT THE AUTHOR   Jenefer Robinson is professor emerita at the University of Cincinnati where she was full professor of philosophy from 1985 until 2016. She writes mainly on topics in aesthetics and philosophical psychology, especially the theory of emotion. Her book, *Deeper than Reason* (OUP 2005) applied recent advances in emotion theory to issues in aesthetics, such as the expression of emotion in the arts, how music arouses emotions and moods, and how the emotional experience of literature and music in particular can be a mode of understanding and appreciation. More recently she has written a series of papers on empathy and the arts, including papers on architecture, painting, literature [in progress] and music. Jenefer was President of the American Society for Aesthetics from 2009-2011.

References

Bedford, Errol. (1967). "Emotions". In: *Essays in Philosophical Psychology,* Donald Gustafson (Ed.). London: Macmillan, pp. 77-98.

Collingwood, Robin George. (1963). *The Principles of Art.* Oxford: Clarendon.

De Sousa, Ronald. (1987). *The Rationality of Emotion.* Cambridge, MA: MIT Press.

Greenspan, Patricia. (1980). "A Case of Mixed Feelings: Ambivalence and the Logic of Emotions". In: *Explaining Emotions*, Amelie Rorty (Ed.). Berkeley: University of California Press, pp. 223-250.

Goldie, Peter. (2000). *The Emotions: A Philosophical Exploration*. New York: Oxford University Press.

Ickes, William. (1997). *Empathic Accuracy.* New York: Guilford Press.

James, William. (1884). "What is an Emotion?". *Mind* 9(34), pp. 188-205.

Lamarque, Peter. (2009). *The Philosophy of Literature.* Malden, MA: Blackwell.

Lopes, Dominic. (2005). *Sight and Sensibility: Evaluating Pictures*. New York: Oxford University Press.

Radford, Colin. (1975). "How Can We Be Moved by the Fate of Anna Karenina?". *Proceedings of the Aristotelian Society*. Supplementary Volumes 49, pp. 67-80.

Robinson, Jenefer. (2004). "Emotion: Biological Fact or Social Construction?". In: *Thinking about Feeling*, Robert C. Solomon (Ed.). New York: Oxford University Press, pp. 28-43.

—. (2005). *Deeper than Reason*: *Emotion and its Role in Literature, Music, and Art*. Oxford: Clarendon Press.

—. (2010). "Emotion and the Understanding of Narrative". In: *Routledge Companion to the Philosophy of Literature*, Garry Hagberg & Walter Jost (Eds.). Oxford: Wiley-Blackwell, pp. 69-92.

—. (2017). "The Missing Person Found, Part One". *British Journal of Aesthetics* 57, pp. 249-267.

—. (2018). "Emotion as Process". In: *The Ontology of Emotions*, Hichem Naar & Fabrice Teroni (Eds.). Oxford: Oxford University Press, pp. 51-70.

Walton, Kendall. (1976). "Points of View in Narrative and Depictive Representation". *Noûs* 10, pp. 49-61.

Weston, Michael. (1975). "How Can We Be Moved by the Fate of Anna Karenina?". *Proceedings of the Aristotelian Society*, Supplementary Volumes 49, pp. 81-93.

'YOU TALKING TO ME?'
Conversations on Art and Aesthetics

Hans Maes
University of Kent

In May 2017, my book 'Conversations on Art and Aesthetics' appeared. It contains conversations with, and photographic portraits of, ten prominent philosophers of art. They are Noël Carroll, Gregory Currie, Arthur Danto, Cynthia Freeland, Paul Guyer, Carolyn Korsmeyer, Jerrold Levinson, Jenefer Robinson, Roger Scruton, and Kendall Walton. The book has two main aims. One is to provide a broad and accessible overview of what aesthetics as a subfield of philosophy has to offer. The other is to stimulate new work in this area of research. In this brief paper I'd like to say a bit more about this second objective. Current research is rarely conducted or communicated in the form of conversations, so the question arises: how can a book like mine fit with and feed into a research culture which is very much dominated by the format of the journal article?

The first thing to note is that, despite the obvious differences in presentation, there are also strong similarities between the discussions that take place in philosophy journals and the discussions laid down in my book. The same basic sequence—X defends a claim, Y formulates objections, X responds to objections— is really at the heart of both. Moreover, it's not too much of a stretch to see the debates that take place in philosophy journals as ongoing conversations between scholars. Looked at it this way, it is not the incongruity but precisely the continuity between the two formats that appears striking. In addition, the conversational format has

some distinct advantages over the more familiar format of the journal article. I'd like to highlight six ways in particular in which this collection could prove a unique and useful resource for further research.

First, in today's academic culture where scholars are prompted to publish separate essays, rather than present grand philosophical systems, it is easy to lose sight of the underlying ideas and overarching themes that hold their work together. The conversation format has made it possible for me to ask authors directly about the overall coherence of their work. And some of the answers I received were surprising. Levinson, for instance, begins by saying that contextualism—the idea that the context of creation is crucial in determining the identity, art status, and meaning of a work of art—is the central thread running through his work. But when I ask him what distinguishes his views from other contextualist views, he mentions how he tends to foreground experience and value more than other analytic aestheticians—a response I had not anticipated given that Levinson is probably best known for essays that barely touch upon issues relating to experience and value (such as 'What a Musical Work Is'[1] and 'Defining Art Historically'[2]). Another interesting contrast comes up in my conversation with Guyer, who has devoted much of his career to the study of one of the most systematic thinkers in history, Immanuel Kant, but who reveals that he has not attempted to make a systematic contribution to contemporary aesthetics himself and that he is in fact a strong supporter of non-reductionist, pluralistic theories of aesthetic value. When I met up with Carroll and Danto, I put the question to them in

---

[1] Levinson 1980.
[2] Levinson 1979.

terms borrowed from Isaiah Berlin's famous essay 'The Hedgehog and the Fox'[3], which divides thinkers into two categories. Hedgehogs relate everything to a single, universal principle, whereas foxes rely on multiple experiences and entertain a vast variety of ideas without seeking to fit them all into, or exclude them from, any one grand system. But while one might expect Danto to own up to being a hedgehog and Carroll to being a fox, they both resist this easy categorisation and go on to explain why their work cannot be pigeon-holed in any straightforward way.

Incidentally, the more holistic approach of these conversations not only allowed me to probe the overall coherence of an author's work, but also to bring to light certain tensions or inconsistencies in their thinking. This is nowhere more evident than in my conversation with Danto. For example, while Danto is adamant that beauty is as obvious as blue and that we spot it immediately when it is present in a work, he also recounts in some detail how he came to appreciate the beauty of Bernini's 'Santa Teresa' only very gradually. Or consider the idea that art does not always have to be beautiful. On the one hand, Danto calls this one of the great conceptual clarifications of the twentieth century. On the other hand, he also acknowledges that a lot of medieval art is not, and was not meant to be beautiful. From a methodological perspective, readers may find it amusing to see how, after faulting Wollheim for refusing to go along with an argument from indiscernibles, Danto himself manifests a similar reluctance when I invite him to think about a painting that would be indiscernible from Motherwell's 'Elegy to the Spanish Republic'.

Second, all of my conversation partners have left their mark

---

[3] Berlin 2013.

on philosophy of art and aesthetics, but some of them have also done significant work in other areas of philosophy (or outside of philosophy). This work is not always acknowledged in the professional journals of our discipline. So I have taken the opportunity in this book to ask them about some of their other writings. Scruton, for example, is a notable conservative philosopher and talking to him about the possible connections between his social and political reflections and his academic work in aesthetics was quite instructive. In my encounter with Carroll we briefly discussed his experience as a critic and screenwriter as well as his book on Buster Keaton.[4] And I begin my conversation with Levinson with some reflections on his not-too-well-known essay on sexual perversion.[5]

Third, in research articles there is seldom room to elaborate on the provenance of one's theories, even though knowledge of the early influences on an author is often helpful in understanding the views they ultimately arrive at. So, I hope the reader will find it as illuminating as I did to hear how, say, Currie was influenced by Imre Lakatos, David Lewis, David Armstrong and later on by Walton and Levinson. Or how Stanley Cavell's teaching and thinking had a lasting impact on Guyer. The book may throw up some further surprises in this respect: Robinson, who is perhaps the most scientifically oriented of all the philosophers I spoke with, acknowledges her debt to F.R. Leavis, the literary scholar who was notoriously dismissive of science in the so-called 'two cultures' debate. Carroll, who has been a vocal critic of some French philosophy in the past, talks about the influence that the French phenomenologist, Merleau-Ponty had on

---

[4] Carroll 2007.
[5] Levinson 2003.

him when he was writing his Ph.D. And although twentieth century philosopher Nelson Goodman is no longer the central figure in aesthetics he once was, it is intriguing to see how his name still pops up in half of the conversations in this volume. Equally intriguing, I should say, are the passages where authors reveal how some of their ideas were not so much influenced but definitely anticipated by others. For instance, it turns out that one of Guyer's key insights about the relation between analysis and psychology in aesthetic theory can also be found in the work of a now largely forgotten female philosopher, Ethel Puffer. Walton, on his part, even admits to being beaten to the punch by a fictional character.

Fourth, what authors do *not* write about can potentially be as revealing as what they do end up writing about. Hence in some of my conversations I have tried to address what might be considered blind spots in the author's oeuvre. With Korsmeyer that meant talking about the definition of art; with Danto and Levinson it meant talking about the aesthetics of nature; and with Walton I spoke about architecture and dance. Furthermore, the conversational format gave me a chance to query not just individual omissions, but also lacunae in the discipline as a whole. For instance, if you look at the leading aesthetics journals you will find many essays on beauty in art and nature, but very few that deal with the beauty of human beings (notwithstanding the fact that outside of academia the term 'aesthetics' is most commonly used to refer to cosmetics, beauty treatments, and bodybuilding). Why is that? Or why has so much been written about particular art forms, especially music, and virtually nothing about other art forms such as sculpture? And what about philosophical texts or philosophers of the past who have fallen into obscurity? As I put the question to Guyer, might there be any hidden gems out there just waiting to be rediscovered?

Fifth, collecting the answers of ten leading philosophers in one volume allows one to develop the sort of overview that can be hard to maintain given the constant stream of research articles. And that in turn can bring to light various unsuspected contrasts and convergences between these philosophers. Danto and Scruton, for example, could not be further apart in their appreciation of contemporary art (one considers Warhol's Brillo Boxes a work of genius, the other dismisses it as a corny joke). But it turns out they do share a strong scepticism regarding the academic professionalisation of philosophy and the relevance of science for aesthetics. Conversely, Robinson and Currie are both eager to forge closer links between scientific and philosophical investigations, but they are increasingly at odds, so it transpires, about the cognitive value of art and literature.

Where possible I have asked authors to comment directly on some of the disagreements that emerged. So, I asked Scruton what he thinks about Danto's idea that Warhol's Brillo Boxes are the culmination of the history of art and I asked Robinson what she thinks about Currie's reasons for doubting that we learn anything significant from the novels she so admires. I also asked Robinson to comment on her disagreement with Korsmeyer regarding the notion of aesthetic disgust and her differences with Levinson regarding musical expressiveness.

In gaining a sense of where these prominent figures stand on important issues, one also gets a better idea of the direction in which the discipline is headed. Take the question that is often assumed to be at the very heart of what analytic philosophy of art is about: the question of the definition of art. In reading these conversations it becomes abundantly clear that the question has lost much of its urgency and importance in recent years. Many of the philosophers I

spoke with simply declare to have no interest in the topic and even those who have written extensively about it in the past, like Levinson and Danto, exhibit an unmistakable weariness when the subject is broached. Questions around aesthetic and artistic value, by contrast, have become much more central now. And readers of this book will be able to track exactly how the battle lines are drawn in discussions about value (with pluralists, such as Carroll and Guyer, pitted against monistic theories of different stripes, including cognitivism, championed by Robinson and Freeland, and aestheticism, championed by Scruton).

This brings us to the sixth and last advantage, namely that these conversations present an excellent occasion to reflect on the discipline of aesthetics itself—something for which the main research journals do not always allow space. What are the future challenges and opportunities for the discipline? Is there genuine progress in philosophy in general and in aesthetics in particular? (Most of my conversation partners believe that there is, though Korsmeyer and Freeland offer some caveats.) Does one need to study the history of aesthetics if one wants to do research in this area? Where do the analytic and continental approaches differ most and is there a possibility of mutual enrichment? (Almost everyone thinks the latter is the case, though there is also the acknowledgement that the divide may have widened in recent years.) How important is style in philosophy and does writing about aesthetics itself need to be aesthetically rewarding? Can aesthetics be relevant for art practice? (Carroll and Danto believe so, but Levinson and Guyer are not so sure.) How, if at all, can aesthetics benefit from current scientific research? How promising are emerging subdisciplines such as experimental aesthetics and everyday aesthetics?

The answers to these questions, and to the other questions I raise, will help to paint a picture of the state of aesthetics today. And that picture, I would like to add in conclusion, is not at all a grim one. In fact, the opposite is true. The field is thriving and expanding, constantly producing new theories and charting unexplored territory: from the culinary arts to video games, from musical chills to 3D cinema, from experimental aesthetics to aesthetic disgust. On the one hand, as you would expect from a flourishing field of study, research is becoming more and more specialised with increasingly sophisticated answers to the most fundamental questions as well as a growing body of work focusing on more and more specific topics. On the other hand, as I hope will be evident from this forthcoming collection, all this research activity has not made aesthetics into an esoteric or exclusive field of study, accessible only to a small elite of experts and isolated from other disciplines or from everyday concerns. To the contrary: aesthetics was and is a perfect 'hub field', as one of my conversation partners rightly pointed out. That is to say, it's a central area from which you can do almost any kind of research in philosophy and which maintains close ties with cognate disciplines such as musicology, film theory, art history, psychology, and narratology. Moreover, since any credible philosophy of art and aesthetics must take its cues from our everyday engagement with aesthetic phenomena and works of art, esoteric tendencies have little chance to develop.

So, if 'Conversations on Art and Aesthetics' can help to make our prospering and accessible branch of philosophy even more appealing to a wider audience, whilst also making a modest contribution to its research culture, I shall consider my time well spent.

Postscript: There is one more advantage to this sort of project that I should mention: as early career philosophers know all too well, writing a philosophy dissertation or paper can sometimes be a lonely affair. The inner dialogue that we constantly engage in when we

consider potential objections and try to think of smart replies sometimes reminds me of the troubled and isolated Travis Bickle character in 'Taxi Driver' ("You talking to me? Well, I'm the only one here") From that perspective, too, having more conversations with actual people might not be a bad thing.

H.Maes@kent.ac.uk

ABOUT THE AUTHOR   Hans Maes is Senior Lecturer in History and Philosophy of Art at the University of Kent. He is the author of *Conversations on Art and Aesthetics* (Oxford University Press, 2017) and editor of the essay collections *Art and Pornography* (Oxford University Press, 2012) and *Pornographic Art and The Aesthetics of Pornography* (Palgrave MacMillan, 2013).

References

Berlin, Isaiah. (1953). *The Hedgehog and the Fox*. London: Weidenfeld & Nicolson.

Carroll, Noël. (2007). *Comedy Incarnate: Buster Keaton, Physical Humor, and Bodily Coping*. Malden, MA: Blackwell-Wiley.

Maes, Hans. (2017). *Conversations on Art and Aesthetics*. Oxford, UK: Oxford University Press.

Levinson, Jerrold. (2003). "Sexual Perversity". *Monist* 86(1), 30-54.

—. (1980). "What a Musical Work Is". *Journal of Philosophy* 77(1), pp. 5-28.

—. (1979). "Defining Art Historically". *British Journal of Aesthetics* 19(3), pp. 232-250.

AESTHETICS NATURALISED
An Interview with Murray Smith

Angelo Cioffi

1   Introduction

This interview has been long in the making: we started it on the 6th of February 2017. I met Murray in his office at the University of Kent, and we spent an intense but pleasant afternoon drinking coffee and talking film, art, and aesthetics. In preparation for the interview, I sent Murray a list of questions I wanted to discuss with him; but, in all honesty, I only managed to get through a third of all the questions I had in mind, and that's because the conversation developed in ways I didn't anticipate, and brought new questions to the table. So, our initial plan for the conversation looked like a jazz score, with some general indications, but with room for improvisation. Ironically, despite having taught with him on his undergraduate module 'Sound and Cinema', among the things I didn't manage to ask Murray were things about his work on film sound and music.[1]

Since our meeting, we have been in contact via email, editing the interview. As a result, the fruits of our improvisation have been manipulated in post-production, and what you see here is an edited version of our conversation rather than just a transcript of the recording. At the time of our meeting, 'Film, Art, and the Third Culture' ('FACT')[2] was about to be published, so a large chunk of this interview is devoted to an analysis of the main arguments presented in that book. Indeed the relationship between aesthetics and science

---

[1] Smith 2002, 2006b.

[2] Jerrold Levinson's 'poetically licensed' acronym for Smith 2017a, in Levinson 2018b ('FACT is a Fact of Both Art and Life'). Levinson's paper also appears in slightly modified form as a review, Levinson 2018a.

is the central theme of the interview, as this relationship remains relevant even when considering topics that have long been among Murray's research interests, such as the role of empathy and emotions in our engagement with films. We have amended the tenses of a few sentences: on the occasion of the original interview, Murray referred to the twentieth anniversary of the first Society for Cognitive Studies of the Moving Image conference as a future event, but as it is now squarely in the rear view mirror, you will read about the conference in the past tense. In addition, Murray held a Laurance S. Rockefeller Fellowship at Princeton University's Center for Human Values for 2017–18, and during the editing of the interview he added some hints of his project there.

The first section of the interview, 'Aesthetics Naturalised', is an introduction to 'FACT', as we talk about the aims of the book and the story behind it. In the second section, 'Aesthetic Experience Triangulated', we delve deeper into the arguments that bind together science and the study of art and aesthetics—in particular, we clarify the role the hard sciences may play in the study of aesthetic experience. In the third section, 'Spectatorship', I ask Murray to reappraise his first—and extremely influential—monograph, 'Engaging Characters' ('EC'), in light of his latest work.[3] In this section, Murray also shows how a naturalised account of aesthetic experience may help provide a solution to the paradox of horror and the paradox of fiction. In the final section, 'Film and/as Philosophy', Murray clarifies his viewpoint on the relationship between film and philosophy, and ends with advice for young academics working in the fields of film and aesthetics.

---

[3] Smith 1995a (revised edition forthcoming 2020).

2   Aesthetics Naturalised

Angelo Cioffi: Let me start with your most recent book, 'Film, Art, and the Third Culture'. What is the 'third culture' you refer to in the book's title?

Murray Smith: The idea comes from a debate that was initiated in the late 1950s by C.P. Snow, who was a Cambridge physicist, but also a novelist, as well as a government minister for a period. So, he was a kind of a polymath figure. Snow wrote an essay—"The Two Cultures"—and delivered a version of it as a lecture in 1959.[4] Snow's essay was essentially a complaint that, as he put it, the culture of the sciences—by which he meant the natural sciences—and what he called 'literary intellectuals'—that was his expression, but really that was a stand-in for the humanities as a whole—were moving apart from one another, that there was a widening gulf between those two domains of intellectual life, those two parts of the academy. So, that became known as the 'two cultures debate'. The 'two cultures' is really a very central debate in the public sphere throughout the 1960s, petering out in the 1970s. Now, within that debate, Snow uses the expression the 'third culture' at one point in his essay, as a way of referring to a kind of intellectual culture that transcended the two cultures. The 'Third Culture' of the title of my book, then, refers to that ideal, the ideal of not being bound into an intellectual culture that sees a simple and rigid divide between what we do in the humanities and what we do in the sciences.

That is where the 'Third Culture' phrase originally comes from, and it has been picked up before, so I'm not the first subsequent author to adopt the phrase and the idea of a third culture. In particular, it was used by John Brockman (the well-known founder of the 'Edge' website). Brockman is a sort of intellectual

---

[4] The first published piece by Snow on this theme appeared in 1956. Generally speaking, though, the start of this debate is dated from 1959, when Snow delivered a revised and expanded version of his essay as the Rede Lecture at Cambridge.

entrepreneur, who works with high-profile, mostly scientific academics. In the mid-nineties Brockman published a book called 'The Third Culture', which was a large collection of interviews with figures like Steven Pinker, Richard Dawkins, Daniel Dennett, and others of that ilk.[5] Now, with that venture Brockman is doing something a little different to what I am doing with the idea of the third culture. What Brockman is really interested in are scientists who are willing and feel able to speak about questions and issues that are traditionally dealt with by the humanities. For example, evolutionary theorists, who have things to say about how culture works; along with the figures I've mentioned, Lynn Margulis and Stephen Jay Gould also appear in the book. So, in Brockman's version of the third culture, the humanities do not have much of a role to play; the idea is that certain scientific disciplines are now sufficiently well developed, they now have enough momentum, so that they can begin to say things about culture without reference to traditional debates. So, that's Brockman's version of the third culture, where, so to speak, the humanities are swallowed up by the machine of science. My version is, I would say, more interdisciplinary, arguing not that we can or should jettison all of the traditional techniques that have been developed in the humanities over decades—indeed, centuries—but that we should complement them and integrate them with the knowledge and methods that come out of science.

Let me try to round off the answer to this first question: I am saying that there are at least these two ways in which the idea of the third culture has been adopted. One way is the Brockman version, in which science colonises the territory of the humanities, and the other one is mine, which suggests that it is a fruitful project to try to integrate the traditions of humanistic and scientific methods.

AC: The book is mainly focused on aesthetics, but do you think that an integration of humanistic and scientific methods can also be fruitful in other domains of knowledge?

---

[5] Brockman 1995.

MS: I am focused on aesthetic questions, but the aesthetic questions are usually a narrower form, a specific form, of a broader question. Now of course, the more specific question will often bring very particular things into play, but I still think, in general, that a lot of the arguments that are specifically about aesthetics in the book will have an echo in other domains, at a high level. So, to take a quick example, one could adopt a third cultural perspective on morality. And I am not talking about morality just as it enters aesthetic experience— that is, the debate about the relationship between aesthetic and ethical value. Let's just say we want to understand ethics and morality on their own terms. And that's our primary focus. There is a whole approach to ethics and morality, parallel to the third cultural approach to aesthetics, which treats them as part of our evolved, natural behaviour, perfectly amenable to scientific enquiry.

AC: This would include the work of Jonathan Haidt and Jesse Prinz, for example?

MS: Indeed. As a general rule, to answer your question, yes, the focus in this book is clearly on aesthetics but in many, many ways this is a version of an argument that can be run with respect to other specific domains. And I have to say that this was one of the challenges in writing the book, that a lot of the time what I am trying to digest and distil is an understanding of some more general area. For example, empathy is not a uniquely aesthetic phenomenon, right? So, the task is, number one, let's understand empathy as a general phenomenon, and of course that is a very complex and controversial area on its own terms. And let us understand it in a naturalised spirit—that is, against the backdrop of relevant scientific knowledge; that's the second part of the task. And then when all of that's done, let's think about how all of this has implications for the aesthetic deployment of empathy.

So, I suppose that is a feature of the way I approach aesthetics—that at every moment I am trying to say, aesthetics *is* a

particular thing; I am not one of those theorists or philosophers who think that aesthetics has been eliminated from our theoretical vocabulary, and that somehow it is just an archaic concept. I think it still picks out something fundamental and real in our experience. But part of the naturalised approach to it involves saying that it is not something inexplicable, something mysteriously distinct from ordinary forms of emotional response or experience. It is particular, but not mysterious; distinctive, but not ineffable. So that is why there is always an effort to relate what is going on in the domain of aesthetics with kindred things that happen in other parts of our lives. Another example, I guess, would be suspense. We might think of suspense as something we mostly experience in artistic and aesthetic contexts—in relation to narratives. But it wouldn't be weird to talk about being in a state of suspense in an ordinary context, when you are waiting for some important result that is about to come through and you are hopeful but fearful about what the outcome is going to be. That is at least very much akin to what we call suspense. So there again, I think we have a relationship between something that seems to be especially relevant in the domain of aesthetic experience but is connected with ordinary experience. One of the principles or general strategies of the book—and I can't remember if I talk about it (laughs) but it strikes me talking to you now—is that that is another way in which aesthetic questions can be made as concrete and as naturalistically tractable as possible.

AC: Your response echoes the subtitle of your book, which sounds more specialist than the main title: 'A Naturalized Aesthetics of Film'. Perhaps this can be taken as a statement of purpose: what is the aim of the book?

MS: I probably should have given the book a subtitle more like '*towards* a naturalized aesthetics of film' (laughs), because one thing that is certainly true of this project is that, you know, every time I would do a spell of work on it, I felt that as the clouds cleared, the

mountain range that I'd set out to climb seemed to get ever higher. If ever there was a project where eventually I reached the point where it was a case of—"this is either going to drive me insane, or I am going to die before I complete this project. So I have to find a way of wrapping up what I have discovered at this point"—this is that project. And that is why I say maybe it should be called 'towards a naturalized aesthetics of film', because to realise the project in its full form is not a project for one person. It's really a research programme, meaning it is a proposal for a whole different way of approaching film in particular, and aesthetics in general, which, if it has a justification—if it is a worthwhile endeavour—it is not for one person alone to realise. Another way of putting this would be to say that what the book tries to be is not so much a realisation of the research programme as a kind of philosophical defence of a research programme that I think is already coming into being. I am not inventing naturalised aesthetics, but rather recognising something that I think has happened around me across my academic career. I have been part of it, but just a part of it. So the book is an attempt to recognise what that thing has been and to give it some shape and to justify it. And when I say 'justify it' I mean also to try and identify its limits, its character, what it can do, what it can't do, what it can claim to do.

AC: And the expression 'naturalised aesthetics', where does that come from?

MS: I am not sure how far it goes back. It probably goes back many decades, but however far it goes back I think its origin must be as an echo of an expression used by the philosopher W.V.O. Quine, in a famous essay from the late 1960s called 'Epistemology Naturalized'. According to Quine, knowledge is an empirical phenomenon, amenable to scientific—and in particular psychological—enquiry. Quine writes: "Epistemology, or something like it, simply falls into place as a chapter of psychology and hence of natural science. It

studies a natural phenomenon, viz., a physical human subject".[6] 'Naturalism' had been a term of art in philosophy for many decades before that—it goes back at least to the early twentieth century. But Quine's work in general, and that essay in particular, was certainly an important landmark and polemical moment in suggesting how far the claims for a naturalistic approach to philosophy could be pushed. So, a couple of decades later you begin to hear aestheticians talking about naturalised aesthetics as an echo of naturalised epistemology. That is part of the explanation of the subtitle and the story of the book.

 AC: And what actually is a 'naturalised aesthetics'?

MS: One important thing that I talk about at the beginning of the book is that naturalism, for most people in the humanities—not philosophers, but just about anyone else in the humanities—would usually be understood as referring to an artistic style or a tradition of art; if we think of the novel it's going to be Zola, or Dreiser in an American context. Or a filmmaker like Ken Loach maybe. It's related to realism ... naturalism has some particular nuances, but it is about rigorously capturing the way the world actually is. That is what people understand by 'naturalism' in the context of art and art theory. However, in philosophy, naturalism really picks out something quite distinct. There might be interesting connections to make between naturalism as a philosophical stance and naturalism as a style of art (though I do not make these connections in the book). So, what does naturalism mean philosophically? Naturalism essentially means an approach to philosophical questions that says that the methods of the sciences have been the most successful knowledge-generating approach to the world that we have invented, and it therefore behoves us, when we think about any question at all, to approach that question against the backdrop of a scientific

---

[6] Quine 1969, p. 82.

understanding of the world. Now that is quite a broad definition, so again to restate it very simply, naturalism is a stance in philosophy that is oriented towards a scientific approach to the world.

To take what is going to seem like an absurd example, but just to pump some intuitions about why this would be an attractive and indeed an important way of proceeding: if I put forward some kind of theory that seems to rely on the idea that the world is flat, or disregards the fact that for many centuries now we have had an understanding of the topography of the earth which holds that it is spherical, people would think that I am crazy, because I would be flouting a pretty fundamental and almost universally shared item of knowledge about the world. So, you could think of naturalism as if it was basically generalising over that principle. If this example strikes people as plausible, the principle is: you should at least seek to make any theory you have about a specific aspect of the world not conflict with firmly established knowledge that is already in place about the world in general. That is a broad-brush idea of what naturalism is.

Let me say a couple more things here. First of all, there is a strong parallel between naturalism in philosophy and the third culture proposal. Really the main title and the subtitle of the book are doing exactly the same thing. The main title is using an idiom which has been used in the public sphere, is a little better known and which tries to flag up the relevance of the two cultures debate, and the idea of a third culture, for some major academic and intellectual debates of the last fifty years. The subtitle is pointing us towards a more localised debate in philosophy, but a very central debate, which I think is essentially the same or very closely-related to the two cultures debate. What these two ideas—the naturalistic stance in philosophy, and the third culture—share is a focus on the question: what is the purview of science? How much weight should we put on scientific knowledge and scientific method when we seek to understand the world as a whole? Are there domains of experience where, so to speak, we put science on the backburner and we just proceed without it, or is that a mistake? Within philosophy, there are

approaches which one can think of as non-naturalistic, which hold that there are certain domains of enquiry, including aesthetics, including morality, where science has no grip, because the phenomena are not apt for scientific enquiry. That's the non-naturalistic perspective. So, the main title of the book and the subtitle are mainly echoes of one another, but addressed to slightly different audiences. One thing I am trying to do is to put these things together, to show how they are related. I think this is part of the role of philosophy—as Wilfrid Sellars famously put it, "to understand how things in the broadest possible sense of the term hang together in the broadest possible sense of the term".[7]

AC: Well, then let me ask you, what is the role of philosophy in the context of naturalised aesthetics? What role can philosophy play in the analysis of, say, our aesthetic experience of films, given that such an approach seems to put so much weight on science? In other words, what is left for us?

MS: One thing philosophy can and does bring to the table is a kind of synthetic approach, whereby what you are looking to do is connect the insights of a broad range of disciplines. So, it is a feature of the modern intellectual world that it is increasingly specialised. We all talk about interdisciplinarity, but what we don't say is "hey, why don't we just *merge* film studies and media studies and cultural studies, why do we not just merge them all?". On the contrary, things always generally tend towards further sub-specialisation rather than the merging of fields towards more unified and larger academic disciplines. So, one of the jobs for this ancient breed we call 'the philosopher' is to look for the underlying shared principles across apparently disparate domains and different academic disciplines, but also for points of conflict and incoherence which may go unnoticed unless somebody is charged with looking for these things. This takes

---

[7] Sellars 1962, p. 37. The passage is used as an epigraph for Chapter 1, and is discussed on p. 21, of Smith 2017a.

us back to the passage from Sellars I mentioned a moment ago, which I think is a beautiful expression of what makes philosophy different from any branch of empirical enquiry. So, I suppose, though I haven't really thought to put a description of the book in these terms before, you can think of what I am doing in the book in that spirit, in the way I am trying to talk about *both* evolutionary theory and neuroscience. It is not a discovery that these two areas of research are somehow related, but it is true that they are both very specialised disciplines, and each of them is broken down into sub-disciplines. Thus, I think there is a job to be done—to say ok, we've got these two trends in the sciences, and it is true for both neuroscience and evolutionary theory that these are areas where many of the participants are very interested in making statements about how their research has relevance for the way culture works. So, I think that there is a task to be done there. We have these trends, we have these bodies of scientific research which have these ambitions, let us try and sort out what is going on here, let us see how these things relate to one another and in turn how it all relates to the kinds of traditional work that we do in aesthetics and in other domains like the one I originally come from—film studies—which take as their foci particular artforms.

AC: Before we delve into the theoretical standpoints you develop in the book, I would like to ask you about the story of the book itself: why did you decide to write this book, and how did you develop the project?

MS: I was in graduate school in Film, in Madison, Wisconsin, from 1985–91. In that particular department, it wasn't regarded as unusual … but let's put it positively: it was regarded as a perfectly respectable research project which took as part of its methodology that it would engage with scientific, generally, and specifically psychological, research. This was the moment when David Bordwell was really launching the research programme which became known as

cognitive film theory. I should note that David had some forerunners, so he wasn't exactly the first person to have had that notion, but he was really the person that put everything together in some crucial books and essays in that period. So, from a sociological point of view, he was the person with both the institutional and the intellectual power to bring things together in a sufficiently cogent way that he was able to make a dent in the way film was studied; and he achieved that partly through his graduate students, including me. So, what I am getting at is that through David and through some other people, an intellectual climate was created—at least where I was based … let's call it a micro-climate (laughs), though I am not sure how far I recognised it was only a micro-climate! —a climate in which you could freely engage in what I am now calling a naturalistic approach to aesthetics, and film in particular or, alternatively, a third cultural approach. Though I wouldn't have used these expressions at the time. That was the period when I wrote my doctoral thesis on character, which would become 'Engaging Characters'. It'll be evident to anyone who reads that book reasonably closely that it draws quite extensively on cognitive science, psychology, and other empirical domains, for example anthropology.

So, that was that period, and that was that project; and then I was released from Madison, back into the wider academic world, and in particular I came back to Britain. I was brought up short by the fact that the intellectual micro-climate that I had been living in for several years really was a very different climate to the one that I returned to in Britain and, in general, I guess I came up against the fact that much of the humanities was still either actively hostile to the interventions by scientists into humanistic questions or, if not actively hostile, it was indifferent to a naturalistic approach, not interested in it. But whichever of those terms you use, the point was: I came up against the fact that the ground was a lot less fertile for the kind of approach that I had been schooled in, at graduate school, than I had anticipated. I have only gradually come to this realisation, but to some extent 'FACT' is the culmination of a very long

experience lasting some twenty-five years, the first phase of which was the period of research in graduate school which concluded with the book 'EC'. That phase involved an unself-conscious immersion into what I can now call a naturalistic approach to those questions. And although 'EC' has been successful enough—it has its fans as well as its detractors (laughs)—in spite of all of that I felt a pressing need, which maybe grew as the years went by, to offer a kind of defence and justification of what I had mostly taken for granted at the time that I was working on that first book. So, that's the broad story of the new book, which is to say that it is a self-conscious justification of a kind of approach to research and to aesthetics in particular, which I adopted really early on in my career but have increasingly felt the need to spell out and justify for my own sake, but also as a project of independent value.

AC: During these years, do you think this method has been spreading in film studies or aesthetics in particular?

MS: I think to some extent it has. Take a couple of symptoms: again around the late 80s, early 90s, you see cognitive film theory gradually coming into being as a new approach to the study of film and it begins to create various institutional structures, one of the more important of which has been SCSMI.[8] Now, that's still quite a modestly-sized academic society, but it has been around now for about 20 years (in fact 2017 was the twentieth anniversary of the first ever meeting of the society). So, I would say that it has grown to some degree, it has consolidated itself and there are parallel developments in literary studies, for example, but I wouldn't say it has massively spread. I don't think you can claim that there's been exponential growth year on year. Indeed, if there had been exponential growth, by now we would be talking about naturalised aesthetics being the dominant approach, and it just clearly isn't. And it's the same with

---

[8] The Society for Cognitive Studies of the Moving Image. Murray is a founding advisory board member of SCSMI, and served as its President from 2014-17.

cognitive film theory—it is still a minority approach, albeit a significant one.

AC: Whatever the current, actual state of affairs, though, yours is still a normative claim, right? That is, you think this is how aesthetics should be done?

MS: ... Well ... Yes. (laughs) I mean, I would need to say more about how the approach I am talking about fits into the broader scope of aesthetics. I was saying earlier that the project is a defence and a justification of naturalised aesthetics, but part of that is about delimiting it, right? So it's not like I want to say: "this method will answer every single question you might have about a work of art, or an experience of natural beauty, and is your one-stop shop or all-encompassing solution". But to come back to your question about it being a normative claim: yes, it is a normative claim, so I agree with you there, but *what* that normative claim is, I may want to spell out a bit more carefully. You know the claim would be: for certain kinds of questions, which I think are central questions, the naturalised method is the best, and, you might even say, the only respectable method. But there might be other questions that people can legitimately ask, where naturalism does not have any claim to be necessarily the best approach to take.

AC: Do you have an example of one of these questions?

MS: I think that if you are talking about the aesthetic evaluation of a single work, it is not that I think that you cannot learn anything from a naturalised approach to aesthetics, but it is going to be very indirect. So, those are the kinds of questions, which I think are going to be the most remote from the method I am talking about, and in general, I am very wary of suggesting that a naturalised approach can answer evaluative questions, normative questions. Because in order to be a scientifically oriented and informed method, it has to hold evaluation and judgement at bay. You can't allow preferences and

judgements of that sort to start colouring the answers you give to what you are posing as empirical questions. How would I like it to turn out to be? That's the whole point about science: it tries to bracket all of those evaluative questions in order to accurately describe and explain the world. So, I don't think a naturalistic philosopher can have it both ways. You can't say that this is the strength of this approach to philosophy and then also say, "and it can answer all these normative questions", because the normative questions have a different character.

AC: Let me get back to something you said earlier about naturalism as requiring a scientific approach to the world. In your book, you mark a difference between scientific knowledge and scientific method. How is this difference relevant to naturalism?[9]

MS: I don't think that distinction is registered often enough. There's an important difference between taking note of *what* science seems to have discovered about the world (scientific knowledge), and *how* science goes about investigating the world (scientific method).

I suppose the way to look at it is this: I would say the first obligation for a naturalistically-inclined researcher is to think about how the question they are posing, the kind of evidence they are drawing upon, the conclusions they are reaching ... how these sit against the backdrop of already recognised knowledge of the world, much of which would be scientific knowledge of the world. Now this does not mean we ought to *slavishly* follow current scientific orthodoxy, because of course what science tells us about the world often changes. One-decade butter is bad for you, the next decade it might be good for you after all (laughs). But taking account of current, relevant scientific wisdom is a kind of pressure or a constraint. That does not yet imply that we are obliged to adopt scientific methods, however. It is a further step for a naturalistically-

---

[9] Smith 2017a, pp. 24-37.

inclined researcher not only to be vigilant and aware of the best scientific knowledge relevant to their inquiry, but to adopt scientific methods.

To narrow the claim down a bit, what I am saying is: if you identify yourself as a naturalistic philosopher, the first thing you do is to be vigilant about your background assumptions concerning whatever domain you are doing research in. So, if you are asking questions about, let us say, empathy in relation to film viewing, I am not going to reach for some arcane theory which may no longer stand as an accepted theory of mind, and think that that is a perfectly legitimate thing to do. I am going to feel some obligation to orient my enquiry to what are regarded as reasonably well-established ideas about the mind. So that is the softer constraint, that is 'naturalism 101'. 'Naturalism 301' is when you start to say, "well, we might be philosophers but, insofar as there is an empirical aspect to the question we are asking, maybe we should actually get our feet wet and engage in some data-gathering".[10] That's where you are beginning to adopt the scientific method. The fashionable name for that trend in philosophy is XPhi, Experimental Philosophy. And there are small subsets of aestheticians who are doing experiments. So that is well established. I should add that I am only on the cusp of really doing that myself, in other words, collaborating with scientists to run experiments; I have done a little bit on it with eye tracking and I might be doing some more with Vittorio Gallese in relation to suspense, using EEG techniques ... But again, to characterise 'FACT' correctly, the book cannot lay claim to what I am calling 'naturalism 301' (laughs).

So, that hopefully gives you an idea of why there is an important difference between scientific knowledge and method. A really humdrum example I use in the book concerns painkilling drugs. We all walk around with what we take to be reasonably reliable knowledge about how painkillers work. Now painkillers are

---

[10]American terminology: 101 is the most elementary course in an academic programme, 301 is a more advanced course, 501 still more advanced, and so on.

absolutely a product of the modern scientific medical world, but as lay individuals, we haven't used scientific methods to prove to ourselves that ibuprofen and paracetamol work the way that we take them to; we've just accepted this knowledge because we live in a scientifically-informed society, which embeds, so to speak, scientific knowledge in so many of its technologies and structures. Like it or not, we're already implicated in a lot of scientific knowledge in that way. To the extent that you take painkillers and you assume that they are going to work reliably in a certain kind of way is just to have accepted a certain body of scientific knowledge. A different step would be to say, "you know what, I am a bit uncertain about the claims that are made for aspirin, I am not so sure that aspirin really is a painkiller, I am going to run some tests!!" (laughs) And of course there are people who are sceptical about certain medicines—see all the stuff about vaccination—and the truly scientific answer to that scepticism is, well, run some tests then! Now, of course that is not easy to do, unless you are within the scientific establishment and you've got all the personnel and the equipment to do so. But the principle is what I am getting at. I am saying that it is one thing to accept, as most of us do in an everyday way, the deliverances of scientific knowledge, and it is another thing to go to naturalism 301— where I actually get engaged in some tests of these elements of scientific knowledge, and investigate them for myself. And of course, we can't all do that all of the time.

3   Aesthetic Experience Triangulated

AC: You endorse what you call a form of 'Cooperative Naturalism'. Could you explain this position, and clarify how is it different from what you call 'cherry-picking'?[11]

---

[11] See the discussion in Smith 2017a, pp. 1-3, which also considers 'Autonomism' (the view that the study of human behaviour should remain wholly independent of the study of the physical and non-human animal world) and 'Replacement Naturalism'

MS: Let me start with the second concept and then work my way backwords to the first one. 'Cherry-picking' is just an everyday expression which I use to refer to academics in the humanities, including philosophers, who will draw upon or allude to scientific discoveries, but do so in an *ad hoc*, unsystematic fashion. They pick scientific cherries when it suits them to embellish and garnish some claim that they wish to make. I think a very good example of this is our friend Gilles Deleuze, less so actually in the books on cinema, but Deleuze's work on mathematics was the subject of a major critique by the physicist Alan Sokal.[12] So, Deleuze would be a good example. He will often appeal to some particular scientific insight or discovery, but without really any more general effort to think about how that discovery fits with his theory in general, how that particular piece of scientific research came out of a larger body of research. Another metaphor I use in my book is "the magpie theft"[13] of scientific ideas. "Oh that's a shiny looking thing, let's grab that scientific claim and we can make something of this", we can use it for rhetorical purposes. So, 'cherry-picking' is where a researcher from the humanities pays this kind of instrumental, short-term, very unsystematic attention to scientific discovery.

Now that contrasts with what I am endorsing as the right way forwards, 'Cooperative Naturalism', where the idea is that if you are researching a question or a domain in the humanities, you should be alert to whether this research touches upon discoveries and knowledge which have been acquired in any other domain, and in particular—bearing in mind that naturalism says, "we tend to learn most about the world through those disciplines which have adopted a scientific method"—why not look to the sciences, to see if there is

---

(the view that the study of human behaviour should, and in time will, be entirely subsumed by the natural sciences).

[12] Sokal & Bricmont 1998. See also Smith's 2010 comments on the 'Sokal Hoax', triggered by Sokal 1996.

[13] Smith 2017a, p. 3.

anything relevant in those specific disciplines? So, much of the book, as you know, talks about various aspects of emotional response to film and to the arts, and that would be a very good example. If you were going to write about emotions in films, there are at least two other ways you could approach this, non-naturalistically. You could either say, we actually have a pretty rich everyday vocabulary for emotions—this is sometimes referred to as 'folk psychology'—so we just stick with that, we just do an investigation of how emotions enter our experience and judgements of films, using nothing more than our folk psychology. And I think that that would take you a certain distance. That's one possibility. Another one is that you might locate some arcane, possibly outdated, body of theory about the emotions, for example, a psychoanalytic one (laughs) …

AC: (laughing) … just a random example …

MS: … and you might say, yes, this looks kind of interesting, on its own terms. Let's not worry too much about whether it holds up to empirical enquiry, testing, replication, and so forth; it just looks interesting, so let's run with this, let's use this as our model of what the emotions are and set it against a body of films and see where it takes us. I am not even saying that that approach has no potential value at all, but I am not sure how it has any *knowledge-generating* value. In other words, it may have a kind of aesthetic value, founded on its ingenuity. A lot of theory works in this way; it's less about the degree to which the theory persuades you that it is insightfully and illuminatingly telling you things about some part of the world, or how the world works. Rather it is a thing unto itself. It's a kind of invented world.

AC: It would be, though, like building up a theory out of the claim that the earth is flat?

MS: Exactly. There is a kind of theorisation where, I think, really the kind of value that it's seeking to fulfil has more to do with the inventiveness of the theoretician, irrespective of what it claims to be studying. But as soon as you hold such a theory up to a more empirical standard it doesn't look very strong. Such work is more like aesthetic performance than empirical investigation. So again, to line up our options here: at the extremes we have autonomism, and replacement naturalism; and we have cherry-picking, which is just an opportunistic use of individual scientific discoveries. Then we've got what I'm endorsing, cooperative naturalism, which is an effort to assess systematically what one is trying to say about an aesthetic question against the backdrop of what is more generally known about human psychology.

So, with all of that in the background, let us just think about what claims can be made for the kind of cooperative naturalism which I am endorsing. In effect, I am saying: "ok, let's take our folk psychology, and put that alongside what various scientific disciplines are telling us about the nature of emotion". And that is going to range over various types of psychology, for example, the very famous work conducted now over several decades by Paul Ekman and his associates about facial expression. [14] It's going to include neuroscience, but also anthropology, and possibly parts of sociology, so there are social sciences which have an important role in this debate as well. But what we are trying to say is: let us see if we can get a more nuanced and deeper understanding of the way emotions work by pressing on beyond folk psychology into the kinds of things that have been discovered scientifically about the emotions.

AC: Let's turn our attention to aesthetics *per se*, and in particular to aesthetic experience. You define aesthetic experience as a particular kind of experience, one that is not merely *had*, but that is *savoured*.[15] So, aesthetic experience is characterised by, or perhaps is a type of,

---

[14] Ekman & Davidson 1994.
[15] Smith 2017a, p. 57.

self-consciousness. I would like to ask you if you could position your own definition within the wider debate on aesthetic experience in philosophical aesthetics. What kind of features make an experience aesthetic? This is a vexed question in the debate on aesthetic experience, since it is related to the unclear distinction between features of the experience itself and the features of the objects that elicit the experience. Does the naturalised account of aesthetic experience offer a solution to this quarrel?

MS: Well, there are several things to say here. First off, I am committed to the idea that there is such a thing as 'aesthetic experience'—that it isn't a myth which ought to be eliminated from our theories. And second, leading on from this, I hold that aesthetic experience can't be reduced to something more basic like attending to particular features of an object. In my view, we need an account which treats aesthetic experience as a distinctive, multi-layered, and complex experience. And yes, as you say, the self-consciousness of aesthetic experience is central to this complexity. This is also the foundation for the idea that aesthetic experience is something that matters to us, something that we value. After all, you can only 'savour' something that you're positively disposed towards! This third point isn't one that I particularly stress in the book, but it is the focus of the project I'm now working on at Princeton.[16]

As for whether a naturalistic approach can illuminate the nature of aesthetic experience: absolutely! I don't take the view that consciousness in general, or aesthetic experience as a type of conscious experience, somehow eludes or transcends empirical investigation. As Daniel Dennett puts it, these phenomena are puzzles, but not mysteries. The primary idea in relation to this point

[16] This point is qualified by Smith 2019, p. 132 ('Proust Wasn't a Neuroscientist'), his contribution to 'Is Psychology Relevant to Aesthetics?', a symposium on both Smith 2017a and Nanay 2016 ('Aesthetics as Philosophy of Perception'). The first fruits of Murray's research project at Princeton will appear as 'Human Flourishing, Philosophical Naturalism, and Aesthetic Value', forthcoming in Corrigan 2020.

in 'FACT' is that we can get a much better grip on the idea of aesthetic experience—often cast in very abstract terms and consequently subject to scepticism—by drawing on psychological research on aspects of aesthetic experience or closely-related phenomena. Here I draw on work by Diana Raffman in relation to 'nuance ineffability' in music perception, Jenefer Robinson's extensive and rigorous engagement with scientific research on emotion, and Bence Nanay's strategy of bringing ideas from perceptual psychology to bear on aesthetic experience.[17] The idea is to soberly insist that aesthetic experience is no more and no less than another facet of human mental life. Moral psychology is widely regarded as an aspect of morality that can be studied empirically. Think of a naturalised approach to aesthetic experience as equivalent to that. In fact, a few years back Elisabeth Schellekens and Peter Goldie hosted a conference on 'aesthetic psychology' that issued in their edited book 'The Aesthetic Mind'. Those enterprises are very much in the spirit of naturalised aesthetics.[18]

AC: As you acknowledge, many believe that consciousness in general cannot be the subject of scientific study, nonetheless you hold that this specific sub-species of consciousness, the one involved in aesthetic experience, can be analysed with a scientific method. You propose to take into account three different levels of analysis we have at our disposal to study mental phenomena (the phenomenological level, the psychological level, and the neurophysiological level), and argue that we can 'triangulate' aesthetic experience. That is to say, you aim to explain aesthetic experience by bringing together evidence from different spheres of knowledge—an example of consilience, if I understand it properly. Can you sketch for us your account of a triangulated aesthetic experience?[19]

---

[17] Nanay 2016; Raffman 1993; Robinson 2005.
[18] Goldie & Schellekens 2011.
[19] Smith 2017a, pp. 57-82

MS: Triangulation is the idea that our best hope of understanding the human mind—and particular aspects of our mental life, like aesthetic experience—is to be open to the various kinds of evidence we have at our disposal: evidence from phenomenology (introspection, self-report), psychology, and neuroscience. So this is a kind of methodological pluralism, compared with approaches like hardcore functionalism (which holds that we have little to learn from neuroscience) and some varieties of phenomenology (which reject any empirical investigation of the mind). And yes, insofar as triangulation seeks to integrate these different types of evidence, it is indeed a form of consilience, characterized by William Whewell (who invented the idea) as the 'jumping together' of different bodies of evidence.

The primary examples of triangulation at work in the book concern empathy and suspense. To take the case of suspense: suspense is a distinctive kind of affective state, and most of us think we can recognise it when we feel it. So, there's our phenomenological evidence. There is also a well-established psychological account of suspense: on this theory, suspense arises when we don't know how a story will turn out, and where we hope for a good outcome but fear a bad one. So, according to that theory, suspense won't arise in situations where we know the outcome of a story—on repeat viewings, say, or films narrating well-known historical events. And yet the evidence of experience—of the phenomenology of watching such films—suggests that we do or at least can experience suspense in such contexts. That gives rise to the problem of 'anomalous suspense', the apparent experience of suspense where the orthodox theory says it shouldn't arise.[20]

This is where neuroscience comes into play: I argue that one way of adjudicating between what introspection tells us ("this feels like suspense!") and what the psychological theory says ("sorry, it

---

[20] The problem of 'anomalous suspense' is explored in ibid., pp. 69-72.

can't be suspense") is to look at the neural correlates of suspense. If we find that the profile of brain activity for a subject engaging with a story in the classic suspense condition (that is, in ignorance of the story's outcome) is identical in the relevant respects with the profile of a subject engaging with a story in the 'no suspense' condition (that is, knowing the story's outcome), then we have reason to revise or reject the orthodox account of suspense, because we have a new source of evidence suggesting that suspense does or can arise even when we know a story's outcome. By the same token, if those two profiles of brain activity look notably different, then we'll have reason to think that our phenomenology is misleading, and that what's going on in the 'suspense' and 'no suspense' conditions are really different. That shouldn't surprise us; introspection is fallible. What also comes out from this example, then, is the idea that, although at the outset we take all three types of evidence seriously, further downstream in the process of triangulation, we may decide that a given piece of evidence is misleading or needs to be reinterpreted. This is where, so to speak, triangulation bares its teeth.

4   Spectatorship

AC: Before, while you talked about the history of the overall project behind 'FACT', you mentioned that in 'Engaging Characters' you were already using a naturalised approach to the study of film. Can you trace a sense of continuity between the two books?

MS: Well, for one thing there is a good amount of continuity in terms of a focus on emotion. That is another reason to think about emotion as a compelling example of a phenomenon which seems to demand a naturalised or third-cultural approach. More specifically, both 'EC' and the new book are heavily concerned with emotional response to art in general, fiction as a sub-type of art, and film fiction as a sub-sub-type. And there is a methodological connection: both are works of naturalistic theory or philosophy. 'EC' is a largely unself-conscious

piece of naturalistic theory, while the new book—and unfortunately it may not be better for it—is burdened with self-consciousness. That's also one of the explicit goals of the new book—to reflect on the underlying principles that gave birth to the first book.

AC: In 'EC', you deal with crucial psychological aspects that shape our aesthetic experience of film, but I also see a continuity in the way you treat spectatorship. In 'EC' (and in a couple of essays you published around that time—I refer to your debate with Richard Allen)[21], you outline a new conception of spectatorship that significantly diverged from what had been the predominant theories in film studies, which saw spectatorship as an illusion or a dream-like experience. In contrast, your new conception underlines the role of consciousness in our experience of fiction. Can you sketch for us your account of spectatorship? How does it relate to the naturalised method you propose in 'FACT'?

MS: It is very important to remember that at the time of the genesis of that project there really was, and I think one can use this kind of language, a reigning view of film spectatorship, which was largely psychoanalytic in inspiration and in which the key term of art was 'subjection'. On this view, to be a spectator is to be subjected to the ideology of the film. In the most sophisticated versions of that theory, the ideological values of the film are manifested in the narrative and the visual structures of the film. What I've just given as a capsule description of the theory can be laid out with a great deal of nuance. But what I am trying to stress was central to that vision of spectatorship was the idea (you know the word 'subjection' says a lot) that individual human spectators are more subjects than agents.[22] And spectators have very little flexibility of response, very

---

[21] Smith 1995b, 1998; Allen 1998.

[22] This contrast might be more precisely stated in terms of a contrast between 'patients' and 'agents', but Murray is adopting the language of contemporary film theory here.

little awareness of the nature of their response, very little ability to reflect upon the nature of their response, all of which simply seemed a caricature of the way, not just film spectatorship, but any kind of engagement with art, or any kind of human engagement with the world, works.

So, you are absolutely right that, put at its simplest, the goal was to say: "Ok, let's put two things back at the centre of our understanding of film spectatorship: one is agency; the other one is consciousness". So, when we go to see films, we are acting in a certain way, and much of that agency takes an at least partially-conscious form. Now, that's the broad thrust and the motivation for a new theory of spectatorship. That said, it is important to stress that in the alternative theory that I articulate, there is still a recognition that there are aspects of film spectatorship that have an involuntary character and over which we have no real control, so to speak, once we've made the decision to engage with a film. In relation to these aspects of film experience, we are passive recipients, being worked on by the film, rather than agents. I am thinking of things like the mere fact that we perceive motion; we have no control over that level of our physiology. But in my view, it would be crazy to regard that as somehow compromising our freedom. We have made a decision to engage with this experience—with the wonders of the moving image; it is not as if this has been imposed upon us. Another example of what I mean, in terms of the more passive aspects of spectatorship, would be reflex responses. This is something that the new book talks about a lot: that there is an aspect of our emotional response—low-level startles, shocks, chills, shivers in the spine—which arises from the way in which films are engineered to work on us in a very directly causal fashion. The larger theory that I am putting forward insists, however, that this is just one dimension of film spectatorship. And we shouldn't think of these visceral, more passive aspects of film spectatorship as somehow necessarily in tension with the more conscious, more intellectual, more reflective dimensions of film spectatorship, any more than we would think of those two things

being somehow in problematic tension with one another in our ordinary engagement with the world. So, the fact that you are cooking dinner and you jump when you realise you've accidentally put your hand on the stove—that's not in problematic tension with the fact that while you are cooking dinner you are also thinking about the lecture you are going to give tomorrow, or for that matter thinking about the recipe you're using, right? Or maybe that crucial part of the puzzle in that final chapter of your dissertation finally gets worked out, as you are stirring the spaghetti sauce. What I am saying is: we are multileveled embodied agents. And that is true across all the domains that humans act and exist within.

Now, there are various kinds of opponents I have here, not just one type of opponent. So, against the psychoanalytic school of thought regarding film spectatorship, I want to stress the cognitive, conscious, and reflective aspects of spectatorship, and say that however it is that we experience a movie as it unfolds—and that may include a lot of reflex behaviours over which we exercise very little immediate control—nevertheless in the longer run, and considering the experience as a whole, we are perfectly capable of reflecting on all of that. And of course, what else is a discipline like film studies if not systematic reflection on the nature of spectatorship? I am insisting that we only have things like film studies because of this capacity, a more basic capacity to reflect on our experiences. So against the psychoanalytic school of thought I am insisting on the active, cognitive and reflective aspects of spectatorship. But in the new book, another school of thought which I am equally opposed to downplays or denies altogether the involuntary, reflex aspects of film and aesthetic experience. My boogie man in the book is Raymond Tallis, so this is mostly in the chapter on neuroscience.[23] Tallis, I should say, does not write about cinema—he generally writes about philosophical issues and to the extent that he writes about the arts it is mostly about literature. And there is a beautiful irony to this story

---

[23] Smith 2017a, pp. 82-105.

which I will come to in a moment. The key thing to emphasise is that Tallis objects to neuroscientific and evolutionary accounts of human experience because he thinks that such accounts reduce us to biological machines. Essentially, he thinks that neuroscientific and evolutionary explanations entirely wipe away any recognition of things like consciousness, our capacity to reflect, our capacity to understand the nature of our actions, and so forth. Now I just think that that is another kind of caricature, a caricature in the following sense: I am with Tallis in recognising that we have all of those capacities and that they are an important part of what makes us human; and these capacities are surely central to what separates us from most of the rest of the animal world. But the idea that you can really understand human behaviour, or the little slice of it that is watching movies, while disregarding things like reflex reaction and basic physiology—that strikes me as an equally reductive perspective. Do you see what I am saying?

AC: Yes, but then how do you reconcile the fact that, as you've argued, aesthetic experience requires conscious reflection on the very kind of experience we are having (we 'savour' the experience, as you put it) with the idea that many of the mechanisms that are involved in our experience actually happen at the level of the sub-personal and the cognitive unconscious?

MS: The way I have been talking about it so far would lead you to think that the reflection I am referring to can happen only after the fact. So you go watch your movie and while you are watching the movie you are jumping around in your seat as the shocks bear down upon you, and you are jamming in the popcorn and gulping down the sugared water ... so you're basically a bag of nerve-ends and reflexes and synaptic firing and there is nothing much going on beyond those physiological reactions. And then, after the fact, you are away from the heat of battle and you can reflect on the nature of the experience. Now clearly, that is part of the picture, in the sense

that I think that for any artform what happens when you are away from engaging with the work is a massively important thing. That is true for literature as much as it is for film. When you are attending to an artwork, you are not at the point in the process where you are reflecting on how you are experiencing whatever it is you are experiencing, or possibly learning from that work … that is, you are not, at that moment in the process, engaging in any kind of reflection on how the artwork might have implications for life itself. So, one thing that makes reflection on our experiences of artworks away from our engagement with them important, is that that's when we think about and reflect upon how our experience of a work might have implications for the way we live our lives, how we learn things about the world, and so forth. But I also want to say that this happens to some degree in the experience itself. And this comes back to your question about reconciling the importance of the underpinning sub-personal components of aesthetic response with the overall character of aesthetic experience, which, as you point out in your question, has this reflexive, self-conscious character where we not only experience a work but we savour our experience of the work. If such reflection just happened after the experience of the work, it would be a much weaker claim. So, I think that part of what is happening, certainly when you have a very powerful and rewarding—a successful—aesthetic experience, is that you are having it and you know you are having it, you know you are gripped by this. And part of that feeling of being gripped and compelled by a work of art is the recognition that "wow, this is really holding my attention, this is really fascinating". But it is important to note that this state of self-consciousness represents the apex of aesthetic experience; I don't mean to suggest that for something to count as an aesthetic experience, the second-order layer must be in evidence throughout the duration of the experience. For certain stretches of an aesthetic experience, and perhaps for the entirety of very simple aesthetic experiences, we may simply be engrossed in whatever the object offers up to our senses and imagination. In 'FACT' I say that

we savour, and don't merely have, aesthetic experiences, when such experiences 'go well'.[24] So what I am really describing is a kind of ideal prototype or exemplar.

Now this complex response—the reflexive, 'double-ordered' character of aesthetic experience—can cash out in a number of ways, depending on the individual work and the genre in which it is situated. Take horror, for example. As everyone in aesthetics knows, through Noël Carroll's work especially, one of the many paradoxes we can talk about in the aesthetic domain is the paradox of horror—which is a modern equivalent of an age-old problem, the paradox of tragedy—where the puzzle is: "how can we explain the fact that we seem to be attracted to something which on the face of it is something that repels us?". It is in the nature of horror as ordinarily understood that it's aversive, it's something that we want to avoid. But, in the context of horror fictions, including horror movies, we (or at least those of us who are horror movie aficionados) seem to be actively attracted to this horrific subject matter. Now, part of what I am saying is that the reflexive character of aesthetic experience gives us a partial explanation of things like the paradox of horror, in the following way: that one stage of engaging with a horror movie for sure might involve those moments of repulsion and disgust and shock, moments where many of us are even going to turn our eyes away from the screen or at least wince. We are going to have emotional reactions which, considered locally, are unpleasant things to experience. But they are contained and framed within a larger kind of project of engaging with this work which—let us put it this way—we calculate, we gamble, will be a rewarding experience, as a whole experience. So that, if the experience of a horror movie was nothing other than a series of disgusting, horrific, localised shocks then it would be very hard to see a solution to the paradox of horror.

---

[24] Ibid. For more on the question of whether second-order 'savouring' is strictly necessary for an experience to count as aesthetic, see Paisley Livingston 2018, 'Questions about Aesthetic Experience', and Murray's 2018 response, pp. 71-5 and 116-19.

But once we introduce into the argument the idea that there is a second-order dimension to our experience, a solution comes into view. We have first-order responses, including, in the case of horror, reactions of disgust and shock. But we also experience second-order responses, in which we reflect on the way our first-order responses to the film are unfolding, the way they are evolving. And that dimension of self-awareness is central to our ability to take pleasure and find something rewarding in things which are, at a first-order level, just repulsive. So, again to tie that back to your question, many of the first-order disgust reactions that I am referring to will have a sub-personal character. They will just be working on some specific physiological mechanism. So, a certain kind of sound, or a certain kind of visual image just will generate a reaction of disgust from us, or a loud unexpected sound will trigger a 'jump scare', a startle response. So that is the sub-personal end of the experience. But that is not mutually exclusive with the reflective end of the experience.

Let us put it this way: to be a human agent is to be this complex of different orders of response, from the very low-level physiological responses to the highest-level, most reflective, most integrated responses. In other words, I am not saying that any human agent is a perfectly consistent and integrated entity, but what I am saying is that there is a part of what it is to be a human agent which involves an attempt to make sense of oneself, as a consistent being.[25] And that is never complete, it is never perfect, but it is an ambition. It goes beyond aesthetics, but it is highly relevant to aesthetic experience, and that's how I reconcile the sub-personal dimension with the conscious reflective character of aesthetic experience.

AC: But then I have another question.

MS: You mean I haven't answered it perfectly? (laughs)

---

[25] On this point, see Smith 2017a, pp. 81-2.

AC: Let's say that it was a perfect answer, but then we have another problem. You solve the paradox of *horror* by pointing at the reflective dimension of our aesthetic experience, yet, the paradox of *fiction* may pose a different challenge: how can *fictional* events and entities elicit *actual* emotions?[26] After all, why are we moved, if we are conscious of the fact that we are attending to a fictional representation? In relation to the paradox of fiction one might even say that perhaps the illusion-based conceptions of spectatorship that you criticise may be in a better position to face the challenge posed by this paradox. According to illusion-based theories of spectatorship, our emotional responses to fictions could be explained either because we take a fiction to refer to real events and persons, or because we mistake the fictional representation for an actual event. Your view of spectatorship as a conscious experience needs to resort to different kinds of arguments to solve this paradox, for it implies that when we apprehend fiction films we never cease to attend to the fact that fiction films are representations built upon conventions. So, how do you solve the paradox?

MS: The way I tend to see this is that there are essentially three kinds of solution that can be offered to the paradox of fiction. As you say, there is the *suspension of disbelief* solution, which essentially says that when we are in the heat of the moment, immersed in and engaged by the work of art—a film or another work of art—we actually lose awareness that it is merely a fiction. And as you are saying there are various different versions of this, but interestingly, this is generally regarded as the longest-standing attempted solution to the paradox of fiction; it goes back at least as far as the Romantics. The phrase 'suspension of disbelief' is from Coleridge and perhaps one can understand why historically it's been a favoured solution, because it is very neat. It simply says, "ok, we can only have emotions when we take ourselves to be responding to something actually

---

[26] See Doran & Moser 2019, Willard 2019, and Robinson 2019 in this issue for an outline and discussion of the nature of the paradox of fiction.

happening, or to have happened at least". So, if we just observe the way in which people become very absorbed in fictions, maybe we can say, within that frame, they've lost their ability to discriminate the real and the fictional. That's one possible solution. The second possible solution is the one most famously articulated by Kendall Walton, which denies that the kind of affective responses we have to fictions are emotions in exactly the same sense as the emotions we experience in real contexts; rather, fictions prompt 'quasi-emotions'.[27] The more I contemplate this question, the more I think that Walton's 'quasi-emotion' solution may actually be the right solution.

AC: So you think that we do not actually feel real emotions when we engage with fiction ...

MS: Yes, except that this is an easily misunderstood solution. It is very important to understand that Walton is not saying that the responses we have lack an affective character, that they cannot be intense, and that they do not in many respects resemble straightforward emotions. To come back to our horror example: when you are sitting there, gagging, bouncing around in your seat, shocked, appalled, weeping, or whatever your specific responses are, Walton is not saying that you are not deeply moved; he is making a very technical point about the difference between those responses and the very similar responses you would have, let us say, if you were on the edge of a motorway, and you were witnessing very similar scenes of actual carnage. In this context, the possibility of intervening in the scene would indeed be a real possibility. Walton's position can easily be misunderstood to be saying that the affective responses we have to fictions are somehow weak, dilute, not very powerful; but the qualifier 'quasi', in the expression 'quasi-emotion',

---

[27] Walton 1978.

has nothing to do with the strength of the feeling. It has to do with the way the response relates to one's beliefs. That's the nub of it.

Now let me run through the third kind of solution. This is the so-called 'thought theory', associated with Noël Carroll and Peter Lamarque.[28] It points out that there are lots of contexts where we appear to have emotional reactions to things which we don't believe to be taking place. And that is not unique just to our experience of fictions. So, we can contemplate things which might happen to us, and sometimes we might contemplate things which might plausibly happen to us. And contemplating things which *might* happen to us, *actually* helps us to plan. So, if you are a person of a very cautious sensibility, then contemplating what a period of unemployment would be like might guide your actions in a very different way to the way somebody who is much more inclined to risk-taking would find their actions steered if they imagined what it is like to be unemployed for a period. The imagination, the power of the imagination, is not something which is only narrowly pertinent to the arts. This is another way of stating a point that came up earlier in our discussion, when I was stressing that part of the perspective of the new book is to say: "look, aesthetic experience is distinctive, but it is not entirely disconnected from the rest of ordinary experience". This is a very good example of that, right? We create fictions, which build on our capacity to imagine things, things which are not true of the world. But where does that capacity come from? Well it comes first of all from the fact that our minds are such that we can plan and anticipate, we can shape our future actions by delimited acts of the imagination representing possible future states. That is what planning is. So, fictions, in a grand way, work on the same basis.

That, I think, all argues very much in favour of the thought theory of emotional response. And that theory is not something which Carroll invented just for the case of horror; it fits with a more general feature of our emotional life, that we do not just have emotions in response to events that have actually happened. We also

---

[28] Carroll 1990; Lamarque 1981.

have emotional responses to things which plausibly might happen, and we also sometimes have such responses to things which never will happen, or never could happen, and that is when emotions tend to be called phobias: when they become entirely irrational. So, I think that, coming back to your original question, about how the model of spectatorship in 'EC' copes with the paradox of fiction, when it looks like the old 'illusion theory' of spectatorship has the advantage of being much better prepared to solve the paradox of fiction, my answer is: well, there are at least two alternatives to the suspension of disbelief theory, both of which are better candidates. The one that is historically and—probably as I see it right now—my own favoured response is the 'thought theory' …

AC: But you said the 'make-believe theory' before …
MS: (laughs) I did, yes.

AC: So, in the span of five minutes you changed your mind? (laughs)

MS: Yes, in the span of five minutes (laughs) … No, let me clarify this. What I am saying is, I have always thought that the thought theory was the solution to the paradox of fiction, and it lines up best with the theory of spectatorship in 'EC' … but a few nights ago I read an essay by Stacie Friend (laughs), which was a defence of Walton's theory of quasi-emotions, which made me think that maybe there is more to be said in defence of his account.[29]

AC: Ok, so your reply to the paradox has changed over time? (Indeed, I did remember that you favoured the 'thought theory').

MS: Well, in all honesty I would say that it always seemed to me that the thought theory and the 'quasi-emotion theory' are the leading

---

[29] Friend 2003. See also Friend's 2018 critical commentary on Smith 2017a, Smith's 2018 response, especially p. 125, and note 51, p. 258, in Smith 2017a.

contenders because the suspension of disbelief theory just has too many costly and implausible implications.

AC: But I would also like to ask you about the role of the physiological dimensions of emotions in the make-believe theory of emotion, because you stress the relevance and importance of the sub-personal and automatic reactions. Do they still play a role in the make-believe theory of emotion?

MS: You are quite right that another part of my perspective is—and this connects with what I was saying a few minutes ago about the multileveled nature of human agency—that when we are watching a movie certain things are happening on a primal physiological level, where discriminations between what's real and what's merely a representation don't come into it, *at that level*. So, if we are strictly talking about the sub-personal level, again, when you jump at a loud unexpected sound, your body reacts just as it would do if you were walking down the high street and a car backfires, or a firecracker goes off. It's the same phenomenon. And your body goes: HUH! In both of those contexts, you exercise no deliberative control over those responses. By the way, the firecracker too is a kind of representation, even though it is also an actual explosion: a miniature, controlled explosion mimicking a larger, uncontrolled explosion. But whether it's a movie explosion, a firecracker, or an actual explosion, which you may one day have the misfortune of witnessing, your body reacts in the same way to all three of those events. But of course, our reaction isn't just at the bodily level. At the same moment that our body jumps, we also cognise what's happening and a few milliseconds later, or certainly a few seconds later, we've already gotten a more sophisticated, conscious (and usually more accurate) understanding of what it is that's happened to us. Combining this personal-level response with the lower-level, sub-personal reactions is an important part of what I term a 'thick explanation'—one that

attempts to capture the various levels and layers of embodied, emotional, human cognition.[30]

Let's take the example of 3D experience. We go to see a 3D movie and it's a strange experience for us; 3D is still a sufficiently new form of cinema that when you sit there and you put your glasses on, and these objects are floating around somewhere between you and the screen—it's a weird visual experience. 3D experience in the movies has almost nothing to do with the three-dimensional experience that you and I are having right now—that is, our experience of navigating a three-dimensional spatial world. Actually, that picks up on another theme of the new book. The idea of *expansionism*. But we'll come back to that … (phone starts ringing) … sorry that's my wife calling but I am going to ignore her (laughs).

AC: I will make a note of that …

MS: Yes, you can keep that on record for posterity … So let's return to the example of 3D experience. You are having this weird visual experience, which is not like any other visual experience you are likely to have outside a movie theatre. And you don't really have an understanding of how on earth the technology is working on your perceptual system to create this bizarre set of sensations, this experience you are having. But at the same time that that is happening, you know full well that you have signed up and paid to have this 3D experience, and you're watching yourself having this experience. That comes back to the self-conscious part of the aesthetic experience—the idea that we don't merely have aesthetic experiences, but that we savour them. And I do not necessarily mean that at every moment you are having this second-order experience, but I am saying, globally, it is a feature of your experience with the film or work of art. You can think of it as an intermittent thing. During certain parts of the experience, you are just going to be focused on first-order visual or narrative experience. But there are

---

[30] On thick explanation, see Smith 2017a, pp. 51-4.

other moments when you will be having a second-order, more reflective kind of experience.

Now let me connect this back to my discussion of the paradox of fiction. Here's an idea which I'll just float as a final hypothesis. One thing that differentiates the kind of emotional experience that we have when we are watching a movie from ordinary emotional experience is that it is partly characterised by a second-order, self-conscious dimension. Now that lines up pretty well with the make-believe emotion theory. So, when I am using my imagination to contemplate future possible courses of action and I am scaring myself with the thought that I might become unemployed, that doesn't necessarily have a second-order dimension to it. The focus here is straightforwardly practical and action-oriented: what am I going to do? Well, that's a really scary thought, and I am not going to let that happen. I would rather work in McDonalds than be unemployed, some people might say, on the basis of such imaginings. So, the thought theory may not discriminate—or at least it may not discriminate very well—between the kinds of emotional experience we have when we engage with artworks, and ordinary emotional experiences prompted by the kind of imagining characteristic of everyday planning. Whereas the quasi-emotion theory captures this difference. It's also really important to remember that that is a small piece of Walton's much bigger make-believe theory of the representational arts, and you have to look at it in the context of the more general theory to understand the strength of the quasi-emotion theory. In this context, it may be a more powerful theory than the thought theory.

And there is a further, interesting difference in this respect: Walton is a system builder. He's got this overarching, very detailed, very nuanced, very ambitious general theory of the representational arts. Carroll by contrast is a piecemeal theorist, he likes to bite one problem off at the time. I wouldn't say that he disregards the systematic perspective, because, you know, the guy has written so much and so broadly of course there are connections in his overall

pattern of thinking. But what I am saying is that he is less concerned with how his answer to any one question connects with a set of related questions and coheres with them. That's a methodological preference. But I think that a piecemeal perspective is an incomplete perspective; ultimately, as a theorist, you need to look at how all the 'pieces' do or do not fit into a larger picture. Sellars' adage again: how does it all hang together? Walton is really—programmatically—interested in that.

5   Film and/as Philosophy

AC: With Tom Wartenberg, you have edited a special issue of 'The Journal of Aesthetics and Art Criticism'.[31] The theme of the issue was film's capacity to convey philosophical meaning or to philosophise. On that occasion, you published an essay where you sounded sceptical.[32] More precisely, you argued that films cannot function like thought experiments because they prompt different kinds of imagination, related to their different purposes and contexts of appreciation. A thought experiment requires *hypothetical* imagining, where the possibility of some counterfactual is imagined in a *spare* and *abstract* way; whereas films require *dramatic* imagining, which involves *elaborating* and *ramifying* the bare counterfactual in one or more ways.[33] Such a difference in the details we are required to imagine seems to point at the fact that a philosophical thought experiment and an artistic thought experiment are 'geared toward different tasks', or so you argued. In addition, you held that the relationship between narrative and argument was impressionistic and undertheorised, implying that narratives could not be taken to make philosophical arguments. Do you still stand by these

---

[31] Smith & Wartenberg 2006.

[32] Smith 2006a.

[33] The distinction is made by Moran 1994, pp. 105-6.

arguments, or has your position on the issue changed?[34] In the same essay you say that there is a difference between knowledge in general and philosophical knowledge more specifically, but then can you clarify what is this philosophical knowledge?

MS: I do think that there is a great deal of work out there which seeks to forge a very close tie between film (and art in general) and philosophy far too quickly; I am very sceptical of the most strident claims in this area. But I have tried to make sense of these arguments and the impulse behind them. I argue that the 'film as philosophy' thesis really amounts to a strong claim about the (potential) cognitive value of film—we can learn things by watching films, and perhaps learn from films in a unique way, specific to film. Few people would deny at least the first part of this claim—that we can gain knowledge by watching films—but by casting this in terms of films 'doing philosophy', the ante is greatly upped. In my view, if we put it in these terms, we're implying that the knowledge we can derive from films meets especially exalted standards, benefitting from the kind rigour and depth of reflection that we expect of philosophy. To my mind, that just mischaracterises the way most art works and what it seeks to achieve, in two senses. First, it tends to obscure the ways in which films and other types of art matter to us non-cognitively—that is, aesthetically—by furnishing us with what I refer to in 'FACT' as "adventures in perception, cognition, and emotion";[35] complex, multimodal experiences which we value for their own sake, that is, independently of any further value they may have. This is the sense in which films are generally 'geared toward different tasks' than works of philosophy. And second, by assimilating films to the category of philosophy too quickly or too completely, we also mischaracterize the *way* in which they can be sources of knowledge.

---

[34] Both Smith and Wartenberg have pushed this debate further in a recent book edited by Katherine Thomson-Jones 2016. On a related topic, see also Smith 2017b.
[35] Murray does not use this exact wording in Smith 2017a, but the idea is discussed on pp. 138-141.

What films lack in argumentative rigour, empirical adequacy, and reflective maturity—the three marks of specifically *philosophical* knowledge—they make up for in terms of imaginative vividness and particularity.[36]

Thinking about the relationship between film (as an art) and philosophy takes us back to the relationship between philosophy and science that we touched on earlier. This is a question of interest because one might think that a naturalised aesthetics is nothing more or other than a scientific, empirical aesthetics; that a naturalised aesthetics is one that is wholly absorbed into science. But that's not my view. As in other areas of philosophy, while a naturalistic stance is one that aligns itself with the sciences, it's not reducible to science in general or any particular science. I mentioned above three hallmarks of philosophical excellence: argumentative rigour, empirical adequacy, and reflective maturity. What marks out philosophy is its combination and self-conscious pursuit of these ideals. It's not that these standards are entirely absent from the sciences, of course; but they do not combine to occupy centre stage as they do in philosophy.

AC: To wrap up, a last question. Do you have any advice for young academics working in philosophy and/or film studies?
MS: There's an old Hollywood adage, usually attributed to John Ford, who said that he'd survived in the film industry by alternating films 'for the studio' with more personal projects. Something like that applies to surviving in academia, I think. On the one hand, it's very important to be intellectually honest, not least with oneself—to pursue the questions one finds important, and the answers one finds most plausible. On the other hand, it's important to be aware of, and not get run over by, the realities of the environment in which one works, and that means everything from government and university

---

[36] Murray characterises philosophy and philosophical knowledge in these terms in 'Film, Philosophy, and the Varieties of Artistic Value', his contribution to Thomson-Jones 2016.

policies to the fads and fashions of academic disciplines. Somehow one has to balance strategic awareness of the opportunities and dangers which present themselves, with authenticity, truth to oneself. And, as the man said: it ain't easy!

angeloe.cioffi@gmail.com

ABOUT THE AUTHOR   Angelo Cioffi received a Ph.D. in History and Philosophy of Art from the University of Kent where he worked as an assistant lecturer in the departments of Film and History of Art. His research is focused on the cognitive value of political cinema; his *Philosophical Theories of Political Cinema* is forthcoming with Routledge in 2020.

ABOUT THE INTERVIEWEE   Murray Smith is Professor of Film and co-director of the Aesthetics Research Centre at the University of Kent. He was President of the Society for Cognitive Studies of the Moving Image from 2014–17, and a Laurance S. Rockefeller Fellow at Princeton University's Center for Human Values for 2017–18. He has published widely on film, art, and aesthetics. In addition to the recent *Film, Art, and the Third Culture: A Naturalized Aesthetics of Film*, his publications include *Engaging Characters: Fiction, Emotion, and the Cinema* (Clarendon); *Trainspotting* (BFI); *Film Theory and Philosophy* (co-edited with Richard Allen) (Clarendon); *Contemporary Hollywood Cinema* (co-edited with Steve Neale) (Routledge); and *Thinking through Cinema* (co-edited with Tom Wartenberg) (Blackwell).

References

Allen, Richard. (1998). "Film Spectatorship: A Reply to Murray Smith". *Journal of Aesthetics and Art Criticism* 56(1), pp. 61-63.

Brockman, John. (1995). *The Third Culture.* New York: Simon & Shuster.

Carroll, Noël. (1990). *The Philosophy of Horror: Or, Paradoxes of the Heart*. New York: Routledge.

Doran, Ryan, & Moser, Shelby. (2019). "Art and the Emotions: Debating the Work of Jenefer Robinson". *Debates in Aesthetics* 14(1), pp. 1-14.

Ekman, Paul, & Davidson, Richard. J. (Eds.) (1994). *The Nature of Emotion: Fundamental Questions*. New York: Oxford University Press.

Friend, Stacie. (2018). "Collaboration in the Third Culture". *Projections* 12(2), pp. 39-49.

——. (2003). "How I Really Feel About JFK". In: *Imagination, Philosophy and the Arts*, Matthew Kieran & Dominic McIver Lopes (Eds.). London: Routledge, pp. 43-62.

Lamarque, Peter. (1981). "How Can We Fear and Pity Fictions?". *British Journal of Aesthetics* 21(4), pp. 291-304.

Livingston, Paisley. (2018). "Questions about Aesthetic Experience". *Projections* 12(2), pp. 71-5.

Levinson, Jerrold. (2018a). Review of *Film, Art, and the Third Culture* by Murray Smith. *British Journal of Aesthetics* 58(3), pp. 336-41.

——. (2018b). "FACT is a Fact of Both Art and Life". *Projections* 12(2), pp. 60-70.

Moran, Richard. (1994). "The Expression of Feeling in Imagination". *The Philosophical Review* 103(1), pp. 75-106.

Nanay, Bence. (2016). *Aesthetics as Philosophy of Perception*. Oxford: Oxford University Press.

Quine, W. V. (1969). "Epistemology Naturalized". In: *Ontological Relativity and Other Essays*. New York: Columbia University Press, pp. 69-90.

Raffman, Diana. (1993). *Language, Music, and Mind*. Cambridge, MA: MIT Press.

Robinson, Jenefer. (2005). *Deeper than Reason: Emotion and its Role in Literature, Music, and Art*. Oxford: Clarendon Press.
—. (2019). "Response to Critics". *Debates in Aesthetics* 14(1), pp. 95-121.

Schellekens, Elisabeth, & Goldie, Peter. (Eds.) (2011). *The Aesthetic Mind: Philosophy and Psychology*. Oxford: Oxford University Press.

Sellars, Wilfrid. (1962). "Philosophy and the Scientific Image of Man". In: *Frontiers of Science and Philosophy*, Robert G. Colodny (Ed.). Pittsburgh, PA: University of Pittsburgh Press, pp. 35-78.

Smith, Murray. (2020). "Human Flourishing, Philosophical Naturalism, and Aesthetic Value". In: *Cinema and Well-Being*, Timothy Corrigan (Ed.). New York: Oxford University Press.

—. (2019). "Proust Wasn't a Neuroscientist". In: "Is Psychology Relevant to Aesthetics?" Symposium on Murray Smith, *Film, Art, and the Third Culture* and Bence Nanay, *Aesthetics as Philosophy of Perception*, featuring commentaries by Sherri Irvin and Elisabeth Schellekens. *Estetika* 1/2019 LVI (New Series: XII), pp. 125-38.

—. (2018). "Film, Art, and the Third Culture: A Response". Response to symposium commentaries on 'FACT' by David Davies, Laura Di Summa-Knoop, Joerg Fingerhut, Stacie Friend, Vittorio Gallese,

Jerrold Levinson, Paisley Livingston, Rainer Reisenzein, Jane Stadler, Katherine Thomson-Jones, and Dominic Topp. *Projections*, 12(2), pp. 111-36.

—. (2017a). *Film, Art, and the Third Culture: A Naturalized Aesthetics of Film*. Oxford: Oxford University Press.

 —. (2017b). "Experience and Explanation in the Cinema". In: *Film as Philosophy*, Bernd Herzogenrath (Ed.). Minneapolis: University of Minnesota Press, pp. 306-330.

 —. (2016). "Film, Philosophy, and the Varieties of Artistic Value". In: *Current Controversies in Philosophy of Film*, Katherine Thomson-Jones (Ed.). Abingdon, Oxfordshire: Routledge, pp. 182-202.

 —. (2010). "Film Theory Meets Analytic Philosophy; or, Film Studies and *L'Affaire Sokal*". *Cinema: Journal of Philosophy and the Moving Image* 1(1), pp. 111-7; http://cjpmi.ifilnova.pt/1-contents.

—. (2006a). "Film Art, Argument, and Ambiguity". *Journal of Aesthetics and Art Criticism* 64(1), pp. 33-42.

 —. (2006b). "The Sound of Sentiment: Popular Music, Film, and Emotion". *16:9* 4(19); http://www.16-9.dk/2006-11/side11_inenglish.htm

—. (2002). *Trainspotting*. London: BFI Modern Classics.

 —. (1998). "Regarding Film Spectatorship: A Reply to Richard Allen". *Journal of Aesthetics and Art Criticism* 56(1), pp. 63-65.

 —. (1995a). *Engaging Characters: Fiction, Emotion, and the Cinema*. Oxford: Clarendon Press.

—. (1995b). "Film Spectatorship and the Institution of Fiction". *Journal of Aesthetics and Art Criticism* 53(2), pp. 113-127.

Smith, Murray & Wartenburg, Thomas E. (Eds.) (2006). *Journal of Aesthetics and Art Criticism* 64(1), special issue on film as philosophy. Reprinted as *Thinking Through Cinema: Film as Philosophy*. Malden, MA: Blackwell.

Snow, C. P. (1956). "The Two Cultures". *New Statesman*, 6 October, pp. 413-14.

Sokal, Alan. D. (1996). "Transgressing the Boundaries: Toward a Transformative Hermeneutics of Quantum Gravity". *Social Text* 46/47, pp. 217-52.

Sokal, Alan, & Bricmont, Jean. (1998). *Intellectual Impostures*. London: Profile.

Thomson-Jones, Katherine. (Ed.) (2016). *Current Controversies in Philosophy of Film*. Abingdon, Oxfordshire: Routledge.

Walton, Kendall L. (1978). "Fearing Fictions". *Journal of Philosophy* 75(1), pp. 5-27.

Wartenberg, Thomas E. (2006). "Beyond Mere Illustration: How Films Can be Philosophy". *Journal of Aesthetics and Art Criticism* 64(1), pp. 19-32.

Willard, Mary Beth. (2019). "Reclaiming the Paradox of Fiction". *Debates in Aesthetics* 14(1), pp. 30-44.

Contents