

The Imaginative Agent

forthcoming in *Knowledge through Imagination*
eds. Amy Kind and Peter Kung, Oxford University Press

Neil Van Leeuwen

Jean Valjean, in Victor Hugo's *Les Misérables*, has just been sent to prison for stealing a loaf of bread to feed his sister's seven children. They are poor, and he'll never help them again. At this point, we find a remarkable passage:

The tears choked his words, and he only managed to say from time to time, "I was a pruner at Faverolles." Then still sobbing, he raised his right hand and lowered it seven times, as if touching seven heads of unequal height, and from this gesture one could guess that whatever he had done had been to feed and clothe seven little children. (pp. 84-85)

Why was Valjean crying? The children were neither in front of him, nor would he see them again. And why do we feel such poignant sentiment on reading this passage? We know Valjean does not exist, yet we cry for him more than for many people who do. He imagines the children; we imagine him.

But now suppose Valjean is a real person with a real human psychology. I want to argue the following. First, the psychological pathway that leads Valjean to cry is *of the same kind* as the psychological pathway that leads readers to cry on reading this passage. Valjean, the agent, and you, the reader of fiction, share something. Second, that pathway from imaginative state to emotional response enables three central agential capacities:

1. *Bodily preparedness* for potential events in the nearby environment.
2. *Evaluation* of potential future actions.
3. *Empathy-based moral appraisal*.

Furthermore, the pathway in question facilitates communication of emotion to other agents in a way that enables them to respond appropriately. In short, this pathway partially constitutes human agency. But it makes us lovers of fiction too.

My argument for this view works like this. First, after making some distinctions that will prove useful (section 0), I posit a pathway from imagistic imagining to emotional activation. My defense of this pathway, which I call I-C-E-C for reasons that will become clear, proceeds *via* consideration of evidence from psychology and cognitive neuroscience (section 1). Second, the evidence I consider shows that this pathway has three important properties: automaticity, reality congruence, and integratability¹ with perception (section 2). Third, I argue that these very properties enable the three agential capacities in question (section 3).

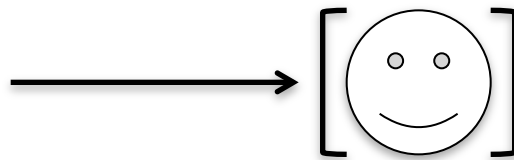
¹ Pardon the neologism. "Integrability" already exists, but it has a specific mathematical sense. So rather than mislead, I prefer a new word to express what I need.

Fourth, I argue that I-C-E-C supports engagement with works of fiction and conclude that this view sheds light both on the so-called paradox of fiction and on imaginative resistance (section 4).

In thinking about agency, philosophers—broadly, not universally—have too much ignored imagination. Imagination not only represents the possibilities from which we choose; it facilitates appraisal of possibilities by sparking relevant emotions. At the same time, in thinking about love of fiction, we have not realized that that love stems from pathways and capacities that also—primarily—enable action in reality. These two oversights—one in action theory and the other in aesthetics—are complementary. I seek to eliminate both at once.

Section 0: What do we mean by “imagine”?

When we say someone imagines something, we imply she has mental representations of things beyond what’s currently available to perception. But despite this loose coherence, there are *orthogonal* uses of “imagine” to distinguish. Suppose you “imagine” that a smiley face is looking at you. Let this diagram schematize the mental states and processes involved in this imagining. We’ll see that the combinatorial structure of the diagram maps to the combinatorial structure of different imaginative mental kinds.



First, you have a *constructive process* of coming up with the representation; the arrow symbolizes this process. Second, since you don’t actually believe there is a smiley face looking at you, you have an *attitude* besides belief toward the content you represent; the brackets represent this attitude, which may be “fictional imagining,” though there are other varieties of attitude imagining as well. Third, your mental representation has a certain *format*, which in this case is *imagistic*—structured like percepts (often also called “depictive”); the smiley face stands for the imagistic format.²

The English word “imagine” can pick out the process, the attitude, the representational format, or *any one or two or three* out of the three. So to be precise about which notion of imagining I’m deploying in a given context, I introduce the terms *constructive imagining*, *attitude imagining*, and *imagistic imagining*.

Importantly, these three things can come together or come apart. You might actually believe a smiley face is looking at you. In that case, you would still have the constructive imagining (arrow) and imagistic imagining (smiley face), without having an imaginative attitude (subtract the brackets); you would have a belief attitude toward the same content instead (one might say, “your imagination is playing tricks on you”). Alternately, you could visually remember something you

² It should be clear from how I set up the terms that not all imagistic imagining (or “imagery”) has to be conscious; in fact, I think much of it isn’t, as Nanay (2013) argues in detail.

saw and believe you saw it; then you'd have just the smiley face, or perhaps the arrow and the smiley face, if remembering involves constructive imagination (I won't judge that issue here). Or you could imagine *discursively*, in an abstract symbolic code (without visual imagery), that a smiley face is looking at you, because imaginative thoughts needn't always involve sensory imagery.³ In this case, we'd put a sentence or formula instead of an image inside the brackets. Or if you look at a picture of a smiley face and close your eyes, you needn't engage in constructive imagining to come up with the imagery you take an imaginative attitude toward; in that case, subtract the arrow but leave the brackets and the smiley face.

Here's why these distinctions are important. Imagistic imagining is incorporated into much mental activity that has nothing to do with fiction, as Williamson (this volume) emphasizes. Much imagistic imagining is of what *has happened* or of what may be *about to happen* in the actual world. But in order for imagery to support human agency usefully, imagistic imagining must yield emotions automatically, whether one believes their contents or merely has an imaginative attitude toward them. If the imagined potential black widow under the bed is to scare us—a fright that is ecologically valuable—the spider from the movie we just saw must do so too. If we conflate the different senses of “imagine” and wrongly think that *all* imagery is fictionally regarded, we'll miss the crucial point that imagery is an area of overlap between fictional and everyday processing.⁴

Section 1: from Imagery to Emotion

Seeing a lion, smelling a perfume an ex-lover used to wear, hearing your grandmother's voice—all these percepts excite emotions. Each causes fear, nostalgia, or affection quickly, before you've thought about why they cause these emotions. For ecological reasons, percepts spark emotions before you have *time* to form beliefs about what you perceive. Nor are such emotional reactions optional. I couldn't decide to feel anything but fear on seeing a lion in the wild, no matter how hard I might try.

So there are pathways from perception to emotion, which we'll explore in more detail. But this is already enough to state the view I endorse about imagery and emotion: imagery has a representational format that is like perception (across all modalities), and imagery is processed in the perceptual cortices of the brain; so imagistic imagining activates emotions along the same pathways as perception. This pathway from imagining to emotion has been posited before by Timothy Schroeder and Carl Matheson (2006) and by me (2011). Here I add detail, acknowledge complexity, and connect the view to imaginative resistance and emotions in response to fiction.

Let me now sketch relevant empirical facts.

³ Discursive imaginings form a proper subset of propositional imaginings, since some propositional imaginings deploy mental imagery as part of their constituent structure, as I argue in 2013.

⁴ Another good reason to recognize these distinctions is that doing so is prerequisite to making sense of the title of this volume: *Knowledge through Imagination*. If we conflated the senses of imagining, we might wrongly think that everything produced by imagination is fictional, in which case the notion of attaining knowledge through imagination would make no sense.

Visual mental imagery activates the brain's visual cortices in the occipital lobe. Auditory imagery—"hearing" in one's "inner ear"—activates the auditory cortex in the temporal lobes. Imagining faces activates the fusiform face area, which processes perceptual face recognition. Motor imagining, like imagining turning a knob with your hand, activates the motor cortex in ways that resemble activation for actual bodily movement. Finally, visual imagining also activates the sensory thalamus, which is active *early* in visual processing.⁵

So our brains have a neat trick: a capacity to use sensory and motor areas not just for processing present, actual perception, but also for representing non-present (often non-actual) motions and objects of perception. This capacity is instantiated just about everywhere it could be in the human brain: a shining example of nature's tendency to give individual structures multiple functions.

Thus, we can understand how the brain relates imagistic imagining to emotion by understanding in more detail how it relates perception to emotion.

There are relatively direct connections between perceptual and emotional areas of the brain. Let's focus on vision and fear for now. Joseph LeDoux (1996) identifies in the mammalian brain two pathways from perceptual areas to the amygdalae. These pathways, the "low road" and the "high road," are largely responsible for preparing fear responses, such as freezing, fighting, or fleeing.

The low road is rapid, but imprecise, casting a broad net that responds to almost anything that resembles a threatening stimulus; the low road is responsible for your jolt of fright on seeing a snake-like stick in the grass. It originates at the sensory thalamus and thus is triggered by perceptual processing that (i) is closer to the perceptual organs themselves and (ii) is likely prior to conscious awareness of a stimulus. Thus, visual input of a snake shape is apt to initiate a fear response *before* we're conscious of it. The low road is especially responsive to fear conditioning, and once an organism acquires a low-road fear association, that association is hard to extinguish and returns easily.

The high road incorporates more sophisticated cortical processing and is slower but more precise: it is far better at distinguishing actually threatening stimuli, like a hawk, from stimuli that merely resemble threats, like a flying squirrel that briefly casts a similar shadow.

To continue along the high road, we should note that visual processing divides into two streams, commonly called "where" and "what."⁶ The "where" (dorsal) stream feeds into the parietal lobe and processes locations of objects. The "what" (ventral) pathway categorizes objects into kinds—a ball, a snake, a tree, etc. The "what" processing, which seems central to conscious visual processing of objects⁷ and ultimately to emotional response, happens largely in the inferior temporal lobe (IT).

The IT, in turn, projects to the amygdalae and other emotional areas. I believe the "what" pathway in the first instance *categorizes* perceptual stimuli, before it conceptualizes them: there is a stage of representation that groups heterogeneous

⁵ Kosslyn, Thompson, and Ganis (2006) summarize this research.

⁶ There are alternate construals of the two pathways, but the differences don't matter for my project.

⁷ See, for example, Leopold and Logothetis (1996).

percepts into one category, where the unity of the category consists in common emotional and behavioral response dispositions. For example, visual percepts of a cat from different angles have very different geometric content, but the “what” pathway still groups those disparate percepts under the same category of CAT. Let’s call this stage of representation *primary categorization*.

Primary categorization is different from *conceptualization*. The latter type of representation deploys concepts that can be constituents in compositional thought and rational inference (Evans 1982, p. 75); the former feeds primarily into emotion and relatively automatic behavioral dispositions. (There are various forms of causal influence between categorizations and conceptualizations, but that doesn’t show they’re not distinct.) The primary categorization of an object you see will typically be the *first* thing the object strikes you as being.⁸ You see dogs first as dogs and only later as particular kinds, like beagles or Great Danes, or as more general types, like carnivore (Rosch *et al.* 1976). In my view, the primary categorization of an object most heavily influences the high road’s impact on emotions, in comparison with more rational conceptualization. You can have a SNAKE primary categorization of an object thrown at you, even if you conceptualize it as a rubber toy. Once an emotion occurs, however, one has what Paul Ekman (2007) calls a “refractory period,” during which cognitive processes, like attention and conceptual thought, favor processing information that coheres with that emotion: fear causes one in the refractory period to think of and attend to other fearful things; joy—joyful things; and so on. So conceptualizations of stimuli are temporally subsequent to primary categorizations and emotional responses, even along the high road. Nevertheless, conceptualizations—what you’re thinking in terms of—prime the perceptual system to categorize incoming ambiguous percepts one way or another.

In sum, the high-road functional neuropsychology from vision to fear works something like this. (1) Early stage perceptual processing leads to a (2) primary categorization, which in turn triggers a (3) fear response (if the category is a feared one). That fear response (4) focuses attention and thought on further fearful stimuli. This continues until the fearful situation is past. I call this pathway **P-C-E-C**: percept-categorization-emotion-conceptualization.

Fear is just one emotion, and vision is just one sense modality; we should be careful when generalizing. Nevertheless, let’s assume for now that *something like* P-C-E-C connects other sense modalities to emotions too: emotions like pity, anger, joy, shame, and disgust, which may also have something like low roads of their own (Zajonc 1984). This assumption, which can be modified as needed, allows us to develop a coherent view of how links between *imaginings* and emotion contribute to agency.⁹ It is an explanatory hypothesis.

⁸ You may be inclined to think that the more sophisticated high-road bypasses primary categorization and goes straight to conceptualization. But the high-road is a structure we share with many simpler creatures, like rats, so it is likely that it employs structures, like primary categorizations, that they have as well.

⁹ In any case, Schroeder and Matheson (2006) give some reason to think the generalization is a good one. They point out that structures that produce multi-modal representations—representations that combine information from more than one sense modality—send signals not only to the amygdala, but

I propose that the pathway from imagistic imaginings to emotions parallels the high road(s) almost entirely. Instead of **P-C-E-C**, we have **I-C-E-C**: imagery-categorization-emotion-conceptualization. Imagistic imaginings in the sensory cortices generate “what” categorizations in the ventral stream, which are largely responsible for activating emotional areas.¹⁰ Subsequent conceptualizations enable inferences and *prime* further primary categorizations.¹¹ An animal shadow in the bush receives a *lion* primary categorization, if one has been thinking about lions, but a *deer* primary categorization, if one has been thinking about deer. There may also be a pathway from imagistic imagining to emotion that parallels the low road. That’s because, as noted, the sensory thalamus is activated during at least visual imagery (Ganis, Thompson, and Kosslyn 2004), and the sensory thalamus is the origin of low road, directly activating the amygdalae.

Section 2: I-C-E-C: Principles and Properties

To grasp the importance of I-C-E-C, try this. First, just discursively imagine—propositionally and without imagery—that *a boy failed his test*. Second, visually imagine him receiving his test back, with despair creeping onto his face, as he hides the grade from other students. For me, the second step, not the first, brings pity. And more imagery yields deeper emotion. Imagine the sound of the boy hurriedly crumpling the test into his bag. Imagine his wet tears, when he lets out his frustration in the bathroom stall. The emotional impact of bare discursive imagining is faint and lacking in nuance compared with that of imagistic imagining.

I-C-E-C operates according to two psychological principles that will be important in what follows.

Percept Similarity: the emotional impact of imagistic imaginings is similar to that of percepts with the same representational structure: if a perceptual episode of the form $P(r_1, \dots, r_n)$ contributes E to one’s overall emotional state, where r_1, \dots, r_n are representational constituents of the percept, then imagistic imagining of the form $I(r_1, \dots, r_n)$ will typically contribute E' , which is similar to E in that it includes the same emotions, though with possibly different intensities.¹²

also to the orbital frontal cortex and the affective division of the striatum, which help distinguish rewards from punishments and produce visceral emotional responses as well.

¹⁰ Olson *et al.* (2007) is relevant here.

¹¹ See Gendler on priming (2003, pp. 134-136).

¹² Humor, especially physical humor, seems to be an exception. Sometimes we laugh at things in imagined fictions that would frighten us or make us sad, were they real. Doesn’t this contradict Percept Similarity? No. The solution here is that humor is very often a release from negative emotions, like fear. So Percept Similarity is upheld, insofar as the negative emotions that get released in humor are still there and not far below the surface. Joëlle Proust (in conversation) has pointed out to me another interesting qualifier on Percept Similarity: emotions consequent on imagery are stronger when that imagery arises from external promptings than when it arises from an internal intention to have the imagery. This, however, won’t make a difference to my argument, since what matters for me is that the valences of the emotions consequent on the imagery are the same in both varieties of imagistic imagining (externally vs. internally prompted).

Conceptual Modulation: discursive imaginings modulate the emotional impact of imaginative episodes by conceptualizing imagistic imaginings:

- i) If mental image *i* is conceptualized as entity *e* and *e* is conceptualized as having property *F*, the imaginative episode will be modulated, relative to mere imagistic imagining with *i*, with emotions similar to those that would typically accompany the occurrent belief *that e is F*.
- ii) Discursive imagining *that p* tends to cause imagistic imaginings that represent constituents of *p*, which cause emotions as specified in Percept Similarity.

To exemplify Percept Similarity, visually imagine/recollect fireworks on a night long ago with a particular lover. This imagistic imagining engenders some of the same romantic excitement as before. Percept Similarity holds. What about non-recollection cases and cases of other emotions? Visually imagining a child with a skinned knee, however fictional, and imagining the sound of his cries engenders some of the same pity as when one actually sees and hears such things, even if the pity is fainter. We can multiply such examples. See/imagine a man robbing a blind man: anger. Hear/imagine acoustically a sinister voice threatening to slit your throat: fear. Feel/imagine touching your favorite dog's fur: affection. Furthermore, an important study by Lang *et al.* (1983) lends support to Percept Similarity. They show that idiosyncratic phobias that are triggered by percepts are also triggered by imagery with the same content, and imagery of objects of a phobia one does not have does not trigger fear. So there is a mapping between the percepts and imagistic imaginings that cause fear, as Percept Similarity predicts.

The natural objection to Percept Similarity is this. Often when one recollects an event imagistically, different emotions arise from those had when experiencing the event. A touch of sad longing accompanies sensory imagining of the fireworks and lover, which wasn't there initially. But this isn't a problem, since Percept Similarity does *not* claim there will be *no* other emotions during imagistic imagining from those that would come (or came) from analogous percepts; it entails only that the imagery contributes the same emotions (with perhaps different intensities) to one's overall emotional state. There may be many background differences. The imagistic elements of the romantic imagining still engender feelings of romantic excitement, but the conceptualization of the recollected events as *gone forever* modulates that feeling with sad longing. Thus, far from undermining Percept Similarity, this is a case of Conceptual Modulation.¹³

¹³ Amy Kind (personal communication) has suggested a particularly strong version of this objection. Suppose the ex-lover is now hated. Would there be any romantic sentiment at all? Percept Similarity seems to require that there at least be some, and it is plausible that in such a case there is none. I'm inclined to say that the emotions here would still be mixed, including some romantic nostalgia, at least for most people. In fact, it is the presence of continued romantic sentiment that makes an ex-lover's sins so hurtful. But in extreme cases, to appeal to the last section, the *primary categorization* that ensues from imagery of the ex-lover may be such as to preclude any romantic feeling. For example, if the primary categorization was no longer LOVER but MURDERER, this would be an exception to Percept Similarity. But this can be seen as an exception to a principle that does in fact govern the normal operation of I-C-E-C.

So let's turn to Conceptual Modulation. Visualize a friend standing face-to-face with a lion. Now discursively imagine *that the lion is wild*. Then imagine *that the lion was raised in captivity*. The same visualization yields different emotions, depending on the discursive imagining accompanying it; fright in the former case and slight relief in the latter. This will be so, I think, regardless of whether one also imagistically imagines the lion's history. Again, cases can be multiplied. Imagine, with as much imagery as you like, a man dying. Now imagine, discursively, that that man is a father. This later imagining adds a tragic pity to the initial sadness.

The neuroscience discussed in the last section, furthermore, coheres with these principles. Percept Similarity coheres with the fact that imagery plays out on the perceptual cortices of the brain. If seeing a certain face yields an emotion, and if visualizing that same face receives much the same processing as seeing it, consequent emotions will be largely the same. The same inference holds, *mutatis mutandis*, for other sense modalities and emotions. Next, Conceptual Modulation coheres with the fact that conceptual thought can prime primary categorizations, and primary categorizations are a major determinant of how percepts and imagery impact emotions. So, in the case of imagining the dying man, thinking of him as a father may yield a FATHER, as opposed to a MAN, primary categorization, which modifies downstream emotional processing. Again, the same holds, *mutatis mutandis*, for other conceptualizations, primary categorizations, and emotions.¹⁴

Now that we've seen the principles according to which I-C-E-C operates, let's consider three of its properties that help explain its roles in agency.

Automaticity. Returning to the pity we had on imagistically imagining the boy who failed the test, we see that the emotion follows the imagery *automatically*. By this I mean two things: first, it is not under direct voluntary control *whether* one has the emotion consequent on the imagery, and second, it is not under voluntary control *which* emotion is consequent on which imagery. It may be that through some kind of training I can learn to have this imagery without having the emotion, but that will be an effortful exception to the default operation of the cognitive-affective system. This automaticity claim coheres with the fact that the way *perception* sparks emotion is automatic in the senses specified; whatever pathways underlie that automaticity underlie automaticity in I-C-E-C too. And it wouldn't make sense for the P-C-E-C connections between perception and emotion to be non-automatic; if *those* were non-automatic, we'd be lunch.¹⁵ Furthermore, it is noteworthy that

¹⁴ You might wonder whether the notion of a primary categorization makes sense in the context of imagining. Don't we start imaginings by *thinking* about stuff, in which case it's conceptualization first? It may be true that we start *many* imaginings conceptually. But the fact that imagery is its own representational format allows it to have processing consequences that weren't anticipated in conceptual thought. For example, visualize an equilateral triangle pointing up (actually do it); now visualize an equilateral triangle of the same size pointing down superimposed? It's only once you've done the visualizing that you arrive at a representation that can receive a RELIGIOUS SYMBOL primary categorization, which wasn't present in the conceptual instructions. Furthermore, this primary categorization has sudden and compelling emotional resonances. In any case, we see that conceptual thought and primary categorization are different processing stages. I thank one anonymous referee for OUP for raising this issue.

¹⁵ Some behaviors may be entirely reflex and bypass emotions altogether. But we'd lose too much flexibility if all behaviors were like that. The automaticity of emotions maintains the priority of

visual imagery at least involves activation of the sensory thalamus, whose projections to the amygdala are along the low road and hence fast.

Reality Congruence. In claiming that I-C-E-C is reality congruent, I mean that the things we imagistically imagine are typically things that could happen in the environment in which we live, given what we believe about that environment.

But isn't imagination, as Hume puts it, free? Not exactly. Imagistic and attitude imaginings are by default constrained by one's environment. And there are at least three reasons to think it takes effort to override the default.¹⁶

First, imagistic imaginings are largely composed of representational constituents perception has already instantiated. Since perception, then, is typically of things in one's environment, the environment constrains imagistic imagining, because it constrains perception. This gives us reason to think that the starting points of imagistic imaginative episodes will often (not always) be of entities that in fact exist (or could exist) in the environment.

Second, factual beliefs, which largely track properties of real entities (Dretske 1983), typically constrain inferences from one imagining to the next. If *S* imagines person *x* is outrageously drunk, *S* will likely further imagine *x* stumbling and having slurred speech. That's because *S* *believes* drunks stumble and have slurred speech. This illustrates how beliefs inferentially govern attitude and imagistic imaginings: they are the informational background that allows inferences from one imagining to the next. Reality constrains beliefs, and beliefs govern imaginative inferences.¹⁷

Third, psychologists Deena Skolnick Weisberg *et al.* (2013) show that children have a strong preference for imagining realistic story continuations, where "realistic" here means *in conformity with familiar patterns of causation*. Their experimental paradigm is to show children an incomplete sequence of images, coupled with a correspondingly incomplete story read aloud, and to ask them to choose among pictures that would continue that storyline. Children mostly choose realistic continuations, *even when* the storyline up to that point had been fantastical (involving magic or other unrealistic elements). Adults, who were tested *via* an online version of the experiment, tended to choose realistic continuations for the realistic stories and fantastical continuations for the fantastical stories. But generally, realistic starting points yield realistic further imaginings.

In sum, much work on imagination suggests that we tend to imagine things that could realistically happen, since imagining defaults to being constrained by cognition of the real. If we imagine in relation to a real environment—what is in that bush?—beliefs, past percepts, and knowledge of causal patterns constrain what we

certain kinds of actions in response to certain percepts, without making the organism lose agential flexibility altogether.

¹⁶ And there are many more than three. Kind (this volume), Langland-Hassan (this volume), Spaulding (this volume), and Williamson (this volume) are all pertinent here.

¹⁷ See my (2013) for more discussion of this point. There are many complications here that don't affect the overall argument of this section.

imagine. As a result, the imaginings at the head of I-C-E-C will tend to be reality congruent.¹⁸

Integratability. Integration is when imagistic imaginings represent objects, properties, and events in the perceived space around the agent. Imagery may be integrated or not, but the possibility of integration (integratability) is important and often overlooked. Valjean's hand comes up and down to different heights *in the space around him*. His imagistic imaginings represent those children as being in that space. Had someone asked, "Which is the youngest?", his hand may have returned to the specific place corresponding to the youngest. And we don't need Victor Hugo for examples of integration. When we plan to move furniture, our imagistic imaginings are typically integrated, allowing us to estimate which furniture goes where.¹⁹ Furthermore, the relevant cognitive neuroscience backs up commonsense on this front. Typical tasks that investigate neural correlates of visual imagery involve having subjects visualize where a recollected item *would* fall on a presently *seen* display; such tasks would make no sense, if integration were not a common phenomenon. What makes Valjean's integration so interesting for present purposes is that his integrated imaginings have the emotional consequences characteristic of I-C-E-C: his emotions were *directed at* the children imagistically imagined around him. We'll see why this is important.²⁰

Section 3: Imagination, Emotion, and Agency

Now that we've seen how imagistic imagining—and, by extension, attitude imagining that incorporates imagery—impacts emotion, we can ask: why might it make sense that we are so constituted? Is I-C-E-C merely a byproduct of P-C-E-C? Or does this pathway do something for us as agents? I argue here that three valuable agential capacities are enabled by I-C-E-C.

Bodily Preparedness

An environment is a space in which a range of potential events could take place. To act responsively to such potential events, you must use your body, which can be more or less prepared for them. I claim that I-C-E-C supports bodily preparedness for actions in relation to potential events in an agent's environment.

Consider an intuitive example. Walking down a wooded path, Stephanie hears a rustle from the nearest bush and leaps back in fright, imagining a mountain lion; her heart racing, she feels ready to grab a stick and fight or just run. A deer darts out. She needs a minute to feel calm again. For the rest of her hike she looks at shadows, continuing to imagine mountain lions, and carries a large branch. The

¹⁸ Does such constraint by belief, perception, and ultimately environment mean the idea that imagination is free is entirely without substance? Not really. Constructive imagination, which generates attitude and imagistic imaginings, is *exploratory constraint satisfaction*, and the existence of default cognitive constraints does not entail that different constraints cannot be chosen. But choice and effort are needed.

¹⁹ See Kind (2013) for discussion.

²⁰ I advocate the integration claim more fully in my 2011. Nanay (2010) also offers arguments pertinent to the integration claim.

intuitive idea here is that Stephanie’s imaginings prepare her body—blood pumping, eyes alert, etc.—for potential encounters with lions.

Psychological research bears out this idea.

Paul Ekman argues that there are *basic emotions*, including “fear, anger, disgust, sadness and contempt” as negative emotions, and “amusement, pride in achievement, satisfaction, relief and contentment” as positive emotions (1999: p. 138).²¹ Ekman essentially has a functionalist theory of emotions, where each basic emotion is characterized by certain universal trigger stimuli, by changes in the central nervous system that process those stimuli, by resultant bodily and facial expressions, and (for some emotions) by changes in the autonomic nervous system (ANS), which regulates bodily functions mostly beneath conscious control (breathing, heart rate, blood flow, etc.). The class of trigger stimuli for a given emotion, in turn, is characterized largely in terms of the evolutionary needs of the organism with respect to a specific ecological problem, though learning can modify the class over time. Facial expressions, as outputs, are characterized by specific, culturally universal face muscle contractions, recognizable to fellow humans, like the facial expression Valjean would have made when weeping.

ANS changes are most relevant to our present claim. Ekman describes ANS changes for fear and anger:

... these ANS patterns evolved because they subserve patterns of motor behavior which were adaptive for each of these emotions, preparing the organism for quite different actions. For example, fighting might well have been the adaptive action in anger, *which is consistent with the finding that blood goes to the hands in anger*. Fleeing from a predator might well have been the adaptive action in fear, *which is consistent with the finding that blood goes to large skeletal muscles* . . .

Freezing in fear might seem to create a problem for this line of reasoning, but not if freezing is interpreted as a fearful state in which the organism is nevertheless still prepared, autonomically, for fast flight if the freezing action does not provide concealment. (1999, p. 50, my emphases)

In short, the autonomic bodily preparedness that partly constitutes emotion has agential utility: it facilitates actions that are adaptively appropriate responses to our immediate environment.

Now we ask: does I-C-E-C, as a source of emotion and hence ANS change, usefully serve bodily preparedness, or does it merely hype our bodies for what doesn’t exist? At this point, excessive focus on fictional contexts would lead us astray, because those contexts are *designed* to trigger I-C-E-C without the presence of urgent ecological challenges that render consequent ANS changes adaptively appropriate. But return to our example of imagined mountain lions. In this case, the ANS response of blood flow to large skeletal muscles is appropriate and could be life saving. Not every threat is *seen* (or heard, etc.), so it benefits us to have imagistic

²¹ There is, of course, controversy about whether Ekman is right about basic emotions. I find his views very persuasive and think that a number of objections to his work are based on uncharitable readings. In any case, let’s take his views as given for now. I suspect that my claims about how I-C-E-C supports agential capacities could work for other conceptions of emotion too, though determining that would be another project.

imagining trigger ANS changes when percepts aren't there. A potential lion attack is a worthy ANS trigger.

Importantly, the three properties of I-C-E-C identified in the last section are crucial to bodily preparedness. If I-C-E-C weren't automatic, its contributions to ANS changes would be too slow and unreliable to be useful, especially in cases of potential lions. If I-C-E-C weren't reality congruent, our bodies would end up being prepared for the wrong sort of thing. Imagining a lover in the bush when a lion is more likely would eventuate in the wrong bodily preparedness. And integration allows us to orient our emotional responses toward or away from specific portions of the surrounding environment. Generalized fear of a potential lion is somewhat useful. But fear of a lion imaginistically imagined *in that bush*, where the imaginistic imaginings are integrated with percepts, is more so.

Evaluation of Future Actions

We need to imagine the future in order to choose what to do in it. One basis for evaluation of future action is how imagined actions and events make us feel. I hold that I-C-E-C supports evaluation of future action by giving us affective responses to imaginistically imagined actions and events.

Again, an example. It's 8:30 AM. David should get to work, but he imagines first stopping to see the charming barista at the local café. He thinks of the pleasant conversation he might have and starts feeling upbeat. But then he realizes this is the café's busiest time. Imagining his attempt at conversation while five caffeine addicts stand behind him makes him feel uneasy. He visualizes a pitying look on the barista's face over his failed attempt. Better to go tomorrow, he decides. Here we see I-C-E-C at work: specific aspects of David's visualizations play a role in how he feels about the content of those visualizations. The valences of those feelings then help him decide on an action.²²

How, more exactly, does this evaluation work? Tamar Gendler and Carson Kovakovich (2005) have clarified how imagining-emotion links contribute to planning, applying Antonio Damasio's neuroscientific work on "somatic markers." I basically agree with their view, which I'll explain in more detail shortly. But first let me highlight the broader context by pointing out contributions that various forms of imagining make to evaluation of future action.

First, an agent must represent options and outcomes to be considered. And she must employ *constructive imagination* to have these representations, since they are not of events in her perceivable environment. So even on simple models of agential choice, there is a role for constructive imagination, which seldom receives comment: one can't choose actions not imagined or attempt outcomes not imagined.²³ That's why "failure of imagination" is often apt; agents may fall short in

²² Just to be clear, I am *not* presenting David as a case of *anticipated emotion*, which would mean he merely anticipates having the uneasy feeling, though he doesn't occurrently. Rather, I am suggesting that he would actually have the feeling in response to the visualization. Thanks to Andrea Scarantino for discussion of this point.

²³ By this, I don't mean to imply that the attitude that one takes to them is that of fictional imagining, though it is *some* form of non-belief cognitive attitude, which we may well call "entertaining." All I am

terms of how many possibilities they build into explicit or implicit decision tables in the first place. And this shortcoming is *prior to* assignment of probabilities: one can't assign probabilities to possibilities one hasn't thought of, so one must constructively imagine any x at all before assigning x a probability. This can be done well or badly.

A splendid example is the Bush administration's "deliberation" over going to war in Iraq. Effectively, their implicit decision table looked like this:

	State of Iraq (as envisioned by Bush Administration)	Expected Value of Action
Action: Go to War	Be greeted as liberators	Good
Action: Not Go to War	Hussein retains WMDs	Bad

It only has one column for the state of nature. They only imagined Iraq one way. And a probability distribution over *only one outcome*, given an action, assigns that outcome a probability of 1. So the administration's main failure in planning was *prior to* the calculation of probabilities²⁴; it was in its failure to imagine alternate possibilities at all. The point generalizes. The Bush administration example is replaceable by any in which agents perform vicious actions, not because of assigning too low probabilities to certain outcomes, but because of failure to imagine those outcomes at all.

Second, *attitude imagining*—take this for the moment to include the broad class of secondary cognitive attitudes²⁵—is required for figuring out likely outcomes. Let's return to the imagined café conversation. David holds fixed the attitude imagining that he goes to the café at 8:30 AM and then rummages around (deliberately or not) in his knowledge base for information that would allow inferences to other things that would be the case, if that were. The impatient customers don't initially occur to him, but by imagining for a time he eventually accesses beliefs that suggest their likely presence. So in discovering outcomes for purposes of planning, agents hold fixed an action in attitude imagination and then use constructive imagination to generate whatever *else* would or might happen.²⁶

Now, third, we're ready to understand the role of emotions in evaluating possible actions. David's felt awkwardness at the visualized presence of the caffeine addicts and his shame at the visualized pity on the barista's face enable him to evaluate the conversation attempt as a poor idea. The important, generalizable features of this example are:

- a) Attitude imagining represents an action to be appraised.

implying at this point is that constructive imagination must be implicated in generating representations of the possibilities under consideration. Thanks to Amy Kind for raising this issue.

²⁴ It's not that much of a simplification, however. See Woodward's (2006) *State of Denial* for a thorough report.

²⁵ *Secondary cognitive attitude* is just my broad term for any cognitive attitude that is not factual belief; this includes hypothesis, assumption for the sake of argument, supposition, etc.

²⁶ I develop this idea of belief governance in my 2013 and 2014.

- b) Constructive imagination, taking the representation in a) as given, generates *imagery* of likely states of the world consequent on that action.
- c) The agent has an emotional response to that imagery because of I-C-E-C.
- d) The emotional response enables the agent to evaluate the action.

Gendler and Kovakovich (2005), as indicated, have a similar view; they discuss Damasio's (1994) neuroscientific research in these terms:

This research seems to show that our ability to engage in practical reasoning rests on the following sort of process: We imaginatively engage with the potential consequences of various courses of action, thereby activating our emotional response mechanisms, and we encode the results of these simulations somatically; the presence of these "somatic markers" then helps to guide our future behavior. (p.248)

Let's flesh this idea out. Humans are so constituted that imagistic imagining accesses internally encoded information that other forms of thought don't, such as likely facial expressions. Such imagery makes important contributions to the agent's overall emotional state because of I-C-E-C. Damasio (1994) argues that affective responses are needed for making rational choices about the future; in particular, those without such responses fail to be properly sensitive to *risk*. For example, people with damage to emotional areas of the brain—especially the medial pre-frontal cortex and the amygdalae—have systematic distortions in decision making, distortions which are correlated with failures of change in the ANS. So if Damasio is right about the necessity of emotional functioning for rational choice, then some imagining-emotional links will be needed for appraisal of future action. I-C-E-C is one such link, and it is the only one that enables use of information that is encoded in a perceptual format.²⁷

In sum, the fact that we imaginistically imagine at all when thinking about the future has at least two functions: (i) accessing of information that might not occur through purely discursive thought and (ii) generating affective states that support evaluation of possible actions.

The three key properties of I-C-E-C enable these functions. Automaticity makes emotional responses to imagining an *informative* source of appraisals of actions and events represented by those imaginings. We *discover* the value for us of those actions and events by discovering the emotional responses that automatically ensue. If the responses were voluntary as opposed to automatic, they would not be informative. Reality congruence is needed too, since otherwise we'd be appraising outcomes that are not likely to occur, which wouldn't serve us. The arrival of impatient caffeine addicts is a reality-congruent feature of David's case. Integratability, finally, supports I-C-E-C in generating appraisals of future action, whenever the action to be appraised relates to one's perceivable environment. In such cases, perception and imagery form a combined manifold to which emotional

²⁷ This paragraph places more emphasis on imagery than Gendler and Kovakovich do. I'm not sure if they would agree with that emphasis. So this paragraph, which represents my view, may be taken as a friendly amendment to theirs or simply as a close but different view.

systems respond: P-C-E-C and I-C-E-C are jointly active. To return to Stephanie's case, integration is important, because the seen distance to the bush in which she visualizes the lion helps her decide whether to fight or flee.

Empathy-based Moral Appraisal

Let's consider one last agential example. "Spare some change?" says the ragged homeless man on the corner. Should you give him the 75 cents in your pocket? You fleetingly imagine the man using the money for cheap booze. But then you recall him digging for scraps in the trashcan yesterday. You then imagine the sandwich your 75 cents would help him buy and his look of relief at having food. So you put your change in his gnarled hand.

Imagery of possible outcomes for *others* sparks emotions *via* I-C-E-C—empathy and pity positively; disgust, loathing, and fear negatively—and those emotions support moral appraisal of action. I mean "moral appraisal" in a broad sense: appraisal of action that is responsive to benefit or harm to others, in addition to oneself, and that treats that benefit or harm as not-just-instrumentally relevant to whether the action should be performed, where benefit counts in favor and harm against.

Three points support this picture of I-C-E-C's contribution to moral appraisal: first, emotions are central to moral appraisal; second, *imagining* in general helps cause moral emotions; third, *imagistic imagining* specifically makes crucial contributions.

The first point has been the subject of so much philosophical and psychological literature in the past decade and a half that I can only gesture at the relevant portions. Jonathan Haidt (2001) and Jesse Prinz (2007) are prominent. Haidt, reviewing three decades of psychology, argues for a *social intuitionist* model of moral appraisal, which contrasts with views like those of Lawrence Kohlberg, which emphasize conscious rule-based reasoning. "Intuitionist" here implies that moral appraisals are largely produced by fast, automatic (*nota bene*), unconscious processes and are triggered by emotional responses to a situation. "Social" implies that feedback from others conditions the emotions and hence the intuitively reached moral judgments. Prinz has a similar view (for my purposes). For him, emotional responses to action situations contribute to moral judgments in a way that resembles how visual percepts in response to surfaces contribute to color judgments; the responses are constitutive of those judgments, but they are also triggered by genuine "concerns" in the world. A congenitally blind person may be able to say "red" and think that things are "red," but she doesn't have robust color judgments in the way normally-sighted people do. Similarly, for Prinz, if the right emotions aren't there, verbal or internal labelings of actions as "right" or "wrong" are not fully moral judgments.

For present purposes, we don't need a view as strong as Prinz'. We must simply hold that moral emotions are typical precursors of moral thoughts, both developmentally (diachronically) and situationally (synchronically). Developmentally: one finds oneself with feelings of disgust, anger, pity, shame, or empathy; one goes on to assume there is something common to everything or most things that elicit each emotion; one then learns to apply internal and external labels

for each of these assumed properties, like “right,” “wrong,” “just,” “virtuous,” and “vicious.” Situationally: each emotion is a trigger for judgments with the respective label as constituent.

If this view of emotion’s influence on moral judgment is right, we should next establish the second point: that imagining (generally speaking) can elicit the relevant social emotions. One can check this oneself: imagining a snickering cashier short-changing a blind person elicits anger automatically; imagining someone physically humiliating a disabled person elicits anger and disgust; imagining a hungry man counting change to see if he can buy a sandwich elicits pity. These imagination-driven social emotions can morally guide action choice. You don’t see the riffling through the garbage or the satisfied look on the homeless man’s face at the time you make the choice; you imagine them.

Daniel Batson *et al.* (2003) support this view. They performed two experiments to test what sort of imagining elicits empathy. In the first, each subject could assign herself or himself to one of two tasks and another person (believed by subjects to be real, but not in fact so) to the other, where one task was relatively pleasant and the other unpleasant. There were three groups: one with no instructions to imagine anything, another with instructions to imagine oneself in the other person’s place (“imagine-self”), and a third with instructions to imagine how the other participant likely feels (“imagine-other”). The authors found that subjects in the imagine-other condition assigned the other person the pleasant task significantly more often than subjects in the no-imagining and imagine-self conditions (the respective portions were .58 versus .25 and .25). Subjects in the imagine-other condition also reported higher levels of empathy using “the six empathy adjectives...*sympathetic, softhearted, warm, compassionate, tender, and moved*” (p. 1195; authors’ emphasis), and the self-reported levels of empathy correlated positively with the assignment of the pleasant task to others.

The third and final point in establishing that I-C-E-C contributes to moral appraisal is that imagistic imagining in particular evokes moral emotions.

Claus Lamm, Daniel Batson, and Jean Decety (2007) provide experimental evidence that visualizing faces generates social emotions. They report a neuroimaging and behavioral study that supports the imagine-self vs. imagine-other difference mentioned above, as well as locating the brain regions implicated in the cognitive and affective sides of empathy. Importantly, imagining the other person’s pain correlates significantly with activity in the fusiform gyrus (FFG)—also known as the “fusiform face area” because of its role in cognizing and imagining faces (O’Craven and Kanwisher 2000).

Experimental work by Elinor Amit and Joshua Greene (2012) and by Eugene Caruso and Francesca Gino (2011) is also relevant to this third point.

Amit and Greene gave subjects two traditional trolley problems. In the first, flipping a switch would steer a moving trolley toward killing one person instead of five. In the second, pushing a person off a bridge and in front of a trolley would kill one and save five. It is well known that subjects tend to respond differently to the two scenarios, even though the costs and benefits from a utilitarian standpoint are essentially the same. But some people are more likely than others to advocate taking

action to kill the one and save the five in *both* cases. Amit and Greene investigated why. We can summarize their main findings as follows:

- (i) people who have a more *visual cognitive style* tend toward the deontological type judgment that killing one to save the five is wrong;
- (ii) having been instructed to visualize the harm one's choice would cause correlates significantly with deontological judgments; and
- (iii) higher levels of emotional response mediate deontological judgments.

By "deontological," Amit and Greene mean morality that appraises actions in terms of exceptionless rules, as opposed to in terms of consequences. In my view, their study shows that employing imagery (visual or otherwise) is one important strategy for moral appraisal, since it is effective at yielding moral emotions.²⁸

Are such moral emotions effective at motivating action? Caruso and Gino (2011) find that *closing one's eyes* before performing an act encourages ethical behavior. For example, subjects who close their eyes while listening to instructions for a one-shot dictator game (where the "initiator" simply chooses how a pile of money is distributed between her and another person) are likely to give more to the other (anonymous) player, as compared to subjects who don't close their eyes. Caruso and Gino hold that the mechanism for this is that closing one's eyes causes one to visualize outcomes, which leads to higher emotional engagement. And higher emotional engagement is morally motivating.

As with evaluation of future potential actions, the automaticity of I-C-E-C is critical to its efficacy in moral appraisal. Affective responses to imagined actions and events serve as orienting guides, as if to say things like, "help him" or "avoid her." If their triggering weren't automatic, they wouldn't be good guides; they'd be like signposts that weren't stuck in the ground. (This is not to say that there is no learning or change in connections between imaginings and emotions over longer time periods; rather, the moment-to-moment connections are automatic.) The reality congruence of I-C-E-C is also especially important for moral appraisal. Without it, we'd have strong moral emotions about events that never would happen, and we'd imagine people with feelings they simply don't have. Imagining that the

²⁸ It's fashionable in some circles nowadays to argue that the more detached, utilitarian style of moral judgment is superior to the empathic style of moral judgment. The general idea is that the utilitarian style is free of the prejudices of the more emotional style and that we do more good for more people by adopting it. One who takes this view, on reading this paper, might think me wrong to write so enthusiastically about I-C-E-C's contribution to empathic moral evaluation, even if I'm right about the psychology. So let me just say that I find the arguments in this vein sophomoric. First, the relevant set of facts isn't always available for purposes of weighing utilitarian costs and benefits, so one must have other strategies. Second, some moral actions are *constituted* at least in part by empathic feelings; kind words to someone suffering are of this sort. Third, we must ask how the basic goods in any utility maximizing scheme are determined in the first place. What is worth maximizing? Utilitarian schemes are empty without some such determination. Here I think our moral emotions in response to actions and events in the world are particularly important. They give us initial purchase on what things are good or not. It may be that sometimes our utilitarian calculi may lead us to act contrary to our moral feelings, but that is because we are choosing to support less salient but greater instances of goods that those feelings identified for us as goods in the first place.

homeless man would thank us for our tough love when we deny him change won't help us behave morally. (Interestingly, integrability doesn't come into play in moral appraisal. If Caruso and Gino are right, we're better off *not* integrating and closing our eyes instead.)

* * *

This has been a whirlwind tour through the agential capacities that I-C-E-C facilitates. But it gives a sense of how much improved our cognitive affective-system is by just one feature: by having a set-up in which imagistic imaginings impact emotions in nearly the same way as percepts do—I-C-E-C works like P-C-E-C—we are much richer as agents.

Researchers often think (implicitly or explicitly) that capacities that impact the phenomenon they study exist *for* that phenomenon: imagination is for fiction; ...for pretend play; ...for appreciating art; ...for geometrical reasoning; ...for counterfactual reasoning; etc. As a result, the research community ends up with as many conceptions of imagination as there are related research areas. But this is not a recipe for progress.

When I first learned a number of years ago that imagery plays out on the perceptual cortices of the brain, I wondered if it might have been better for us if we had had two sets of perceptual cortices, one set for perception and one for imagery. It seemed to me then that we would be capable of more detached imaginative reflection with such a set-up. Perhaps, I thought, the only utility of having imagery on the same set of cortices as perception consists in saving space and metabolic resources. But now, as you see, I think I was wrong to think that. Having imagery play out on perceptual cortices does a world of good for us as agents. One pathway yields three agential capacities—and it drives us to fiction as well. *This* is the kind of unity the research community on imagination should be seeking.

Section 4: I-C-E-C and Fiction

By now one can predict my explanation for why humans find fiction so engaging. In very broad brushstrokes, works of fiction trigger I-C-E-C and are so crafted (in the usual case) that they maintain various forms of emotional activation once we've gotten into the work. This emotional activation prompts further engagement with the work and the cycle can begin anew; the emotional refractory period is crucial here, since it eventuates in further attention to the source of the emotion. The reason that I-C-E-C doesn't shut down just because we know the fictions aren't real is that I-C-E-C *generally* reacts to entities and events that are potential, and fictions, if they are well crafted, sneak into this class. If they're not crafted well (in the relevant sense of "well"), they don't trigger I-C-E-C and don't "do it for us." Since I-C-E-C makes this all possible, the answer to the question of *why we like fiction* simply becomes whatever the answer is to the question of why we have I-C-E-C. And that answer, I have argued, is to be found in examining what the pathway does for us as real agents.

There is much to say about this approach to understanding humans as fiction-loving creatures. For now, I restrict myself to two specific points on two of the most prominent puzzles in the philosophical literature on fiction: imaginative resistance and the paradox of fiction.

Imaginative Resistance

The version of the puzzle of imaginative resistance I'm concerned with is what Tamar Gendler and Shen-yi Liao (forthcoming) call "the phenomenological puzzle." When we attempt to process a work of fiction, it seems easy to incorporate outlandish descriptive propositions into our understanding of the story, but our minds are far less flexible about incorporating outlandish moral propositions. Try imagining that Ebenezer Scrooge was *right* to dock Bob Cratchit's pay for using an extra lump of coal to warm the room where he worked. It's easy to imagine *that* he does this, but our minds falter when it comes to allowing the moral claim into the fictional world. I find myself thinking, "No way was it right for Scrooge to do that!" But by comparison, not for a minute do I think, "Hey, Cratchit doesn't exist!" We take descriptive existential claims on board easily. So the puzzle is to explain why our minds seem to have one form of flexibility and not the other.²⁹

Here's how I (non-deductively) explain that inequality of flexibility.

1. We can imagine mostly³⁰ whatever descriptive fact we like, or whatever an author asks us to imagine descriptively, in a fiction.
2. Whatever we descriptively imagine *automatically* triggers moral emotions *via* I-C-E-C.
3. Those moral emotions generate moral appraisal of actions in the fiction.
4. The moral appraisals of actions in a fiction are thus constrained by descriptive imaginings *via* I-C-E-C.
5. Since descriptive imagining of a fiction constrains moral appraisal of the fiction, the latter is less flexible than the former.

In essence, the automaticity of I-C-E-C makes it such that the descriptive facts we imagine, if we imagine any elements of them imaginistically (which we often do), *settle* our emotional moral responses. An analogy may help. Descriptive imagining of a fiction is like the turning of a key in an ignition; I-C-E-C is like the engine of a car; and moral responses are like the starting of the motor. However much control we have over turning the key in the ignition, we have little control over whether the engine starts, given that we have turned the key and given that we have a functioning engine with gas.

The reality congruence of I-C-E-C is also pertinent here. Though we may be able to vary what we imagine a great deal, our imaginings will default to being

²⁹ As a matter of historic interest, I think the phenomenological puzzle is closest to what Hume (1757) had in mind: "Whatever speculative errors may be found in the polite writings of any age or country, they detract but little from the value of those compositions. There needs but a certain turn of thought or imagination to make us enter into all the opinions, which then prevailed, and relish the sentiments or conclusions derived from them. But a very violent effort is requisite to change our judgment of manners, and excite sentiments of approbation or blame, love or hatred, different from those to which the mind from long custom has been familiarized." He seems to say that we have an easy time "enter[ing] into" what we regard as false descriptions of the world, but we have a terribly hard time doing the same for what we regard as false moral proposition.

³⁰ The exceptions don't make a difference here.

similar to that which could occur in the world in which we live. Thus, we find ourselves, because of reality congruence, imagining things with which it seems to us we could interact. We import much more of our beliefs about the causal structure of the world into cognition of even the most outlandish fictions than we think. So our emotional responses to fictions, including moral emotions, will always occur *as if* those fictions are potential realities for us. (And of course, if they don't strike us that way, we just put the book down.)

The Paradox of Fiction

The Paradox of Fiction is about whether it is rational to experience emotions in response to fiction. I won't attempt an ultimate solution here. But I wish to argue, using the framework I've developed, that an apparent entailment of two of the premises that make up the paradox is false. But that entailment, were it true, would have a great deal of practical significance, so its falsity is important.

The paradox has three (to varying degrees) intuitive premises, which jointly entail a contradiction.

Premise 1: it is irrational to have emotional responses to entities believed not to be real.

Premise 2: imagined fictions are believed not to be real.

Premise 3: it is not irrational to have emotional responses to imagined fictions.

The intuitive force behind Premise 1 comes from the need to explain the following sort of case: I imagine I have cancer (even though I know I don't), and because of this I get extremely emotionally agitated. It seems that we need something to explain the irrationality of this case, and Premise 1 *prima facie* fits the bill. I take Premise 2 to be true; denying it would conflate cognition of fiction with delusion. And Premise 3 will seem sensible to anyone who emotionally enjoys fiction (and does so without embarrassment).

Premises 1 and 2 jointly entail that it is irrational to respond emotionally to fictions. This further seems to entail that *we should train ourselves not to have such responses*. After all, we feel normatively obligated to train ourselves not to fall into other mental patterns we realize are irrational. So the mooted irrationality of emotional responses to fiction seems to entail that we should practice not having them.

I am concerned with this latter entailment. Training ourselves not to emotionally respond to fiction would involve effectively destroying I-C-E-C, or at least making it much weaker. I'm not sure it would be possible to do this. But if it were possible, doing it would be a bad idea, since it would involve destroying something on which three important agential capacities depend. So it is false that we should so train ourselves.

Suppose we were to train ourselves not to respond emotionally to plays by doing away with the *integrability* property of I-C-E-C that lets us turn a sparse set into an imagined richer world. That would be folly, for it would also undercut Stephanie's bodily preparedness for dealing with potential nearby mountain lions and other predators. Or suppose we were to train ourselves not to imagine fictional

worlds in such a reality-congruent way that they are emotionally powerful; we *might* be able to do this by habituating ourselves only to imagine incoherent and outlandish scenarios in response to the promptings of fictional works. But to do away with the *reality congruence* of I-C-E-C would be to make bodily preparedness misdirected, to rest evaluation of future actions on the bizarre, and to divorce moral appraisal from the dispositions of actual people. Or perhaps we could make scary movies no longer elicit fear by training away the *automaticity* of I-C-E-C. But to do this would be to make ourselves no longer properly sensitive to real risk, rendering us—when it comes to choosing for the future—functional equivalents of Damasio’s lesion patients. In sum, anyone who scorns the human propensity to respond emotionally to fictions has just not understood the consequences of doing away with that propensity.

If all parties to the discussion share Premise 2, then the last recourse of the defender of Premise 1 will be to deny the entailment: to deny that the supposed irrationality in question entails that we should train ourselves not to respond emotionally to fiction. Investigating conceptions of rationality according to which one could sensibly deny this is well beyond the scope of this paper. But we can conclude that Premise 1, whether true or false, has no significance for how we should train our cognitive-affective system to operate.

Conclusion: a systematic view

Jean Valjean, as an agent, imagistically imagined his sister’s children standing before him; that imagery helped guide his hand to each of their phantom heads. Those imaginings also, *via* I-C-E-C, triggered feelings of affection in him and ultimately sadness at his inability to help.

You, the reader of the passage, imagistically imagine Valjean and his action of patting the imagined heads. Because your mind too has the I-C-E-C pathway, you first feel affection for Valjean’s kindheartedness, which then turns to sadness over his predicament. The sad thing about the Valjean case—both our perspective on it and his—is that there is nothing we can do to help. But I hope to have shown that these sorts of imaginings, this pathway, and these emotions exist for good reasons: they put us in positions—not always, but often—to act well as agents in our present environments, in relation to the future, and in relation to others.

P-C-E-C, the pathway from perception to emotion, was almost certainly largely in place in our most recent evolutionary precursors. And then, as I see it, Mother Nature—by processes still unknown—discovered an incredible trick. By enabling imagistic imagining in all our brain’s perceptual areas, she created I-C-E-C out of P-C-E-C and thereby in one fell swoop enabled the three capacities of which I write. But as soon as this happened, we became suddenly vulnerable to fictional enchantment, and this enchantment leads to moral entanglement with unreal people, places, relationships, and worlds. Is this entanglement irrational? If we narrow our focus onto the entanglement itself, it may seem so, since it’s not obvious what *this* or *that* fiction does for us. But if we take a systematic view, we see the entanglements are charming byproducts of an extremely well-designed system of cognition and affect. Calling those entanglements “irrational” may not be entirely wrong. But it is pointless, for it would be folly to change the cognitive-affective

system that produces them. Our constructive imagination surrounds us by possibilities that emotionally excite us, *mostly* for good practical purposes. But being so constituted leaves us vulnerable to the excitement of make-believe. Enchantment is the price of human agency.

Acknowledgements

I'd first like to thank the research assistants who helped me explore the empirical and philosophical literatures pertinent to this paper; Scott Danielson, Cameron Hamilton, and Rob Hatcher were extremely helpful. Next, for feedback on earlier drafts, I'd like to thank Leslie Andrews, Ben Freed, Josh Landy, Peter Langland-Hassan, Hyoung Kim, Amy Kind, Peter Kung, Shen-yi Liao, Eddy Nahmias, Andrea Scarantino, Shannon Spaulding, and one anonymous referee for OUP. I'd also like to thank Amy Kind and Peter Kung for having me as an invited discussant at the *KTI* conference in April 2012. I delivered presentations based on this paper at Institut Jean Nicod and at the Centre for Philosophical Psychology at the University of Antwerp in May 2013; I thank Joëlle Proust and Bence Nanay, respectively, for hosting me, and I thank those audiences for their feedback.

References

- Amit, E., and Greene, J.D. (2012) "You see, the ends don't justify the means: Visual imagery and moral judgment," *Psychological Science* 23(8), 861-868
- Batson, C. D., Lishner, D. A., Carpenter, A., Dulin, L., Harjusola-Webb, S., Stocks, E. L., Gale, S., Hassan, O., & Sampat, B. (2003). "As you would have them do unto you': Does imagining yourself in the other's place stimulate moral action?" *Personality and Social Psychology Bulletin* 29, 1190-1201
- Caruso, E., and Gino, F. (2011) "Blind ethics: Closing one's eyes polarizes moral judgments and discourages dishonest behavior," *Cognition* 118 (2), 280-285
- Damasio, A. (1994) *Descartes' Error*, New York: Avon
- Dretske, F. (1983) "The Epistemology of Belief," *Synthese* 55, 3-19
- Ekman, P. (1999) "Basic Emotions," in T. Dalgleish and T. Power (eds.) *The Handbook of Cognition and Emotion*, Sussex, U.K.: John Wiley & Sons, Ltd., 45-60
- Ekman, P. (2007) *Emotions Revealed* (2nd Edition), New York: Holt Paperbacks
- Evans, G., 1982. *The Varieties of Reference*, Oxford: Oxford University Press
- Ganis, G., Thompson, W.L., Kosslyn, S.M. (2004). "Brain Areas Underlying Visual Mental Imagery And Visual Perception: An fMRI Study," *Cognitive Brain Research* 20(2), 226-241
- Gendler, T. S. (2003) "On the Relation between Pretense and Belief," in D. M. Lopes and M. Kieran (eds.) *Imagination, Philosophy and the Arts*, Routledge, London, pp. 125 – 141.
- Gendler, T. S., and K. Kovakovich, 2006, "Genuine Rational Fictional Emotions," in M. Kieran (ed.) *Contemporary Debates in Aesthetics and the Philosophy of Art*, Oxford: Blackwell, pp. 241-53.
- Haidt, J., 2001, "The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment," *Psychological Review* 108(4): 814-34.
- Hugo, V. (1862/1987) *Les Misérables*, trs. Lee Fahnestock and Norman MacAfee, New York: Penguin
- Hume, D. (1757) "Of the standard of taste," in *Essays: Moral, Political and Literary* (1875), London: Longmans, Green, and Co.

- Kind, A. (2013) "The Heterogeneity of Imagination," *Erkenntnis* 78(1), pp. 141-159
- Kind, A. (this volume)
- Kosslyn, S., Thompson, W., and Ganis, G. (2006) *The Case for Mental Imagery*, Oxford: Oxford University Press
- Lamm, C., Batson, C.D., & Decety, J. (2007) "The neural substrate of human empathy: Effects of perspective-taking and cognitive appraisal," *Journal of Cognitive Neuroscience* 19, 42-58
- Lang, P., Levin, D., Miller, G. and Kozak, M. (1983) "Fear behavior, fear imagery, and the psychophysiology of emotion: The problem of affective response integration," *Journal of Abnormal Psychology* 92, 276-306
- Langland-Hassan, P. (this volume)
- LeDoux, J. (1996) *The Emotional Brain*, New York: Touchstone.
- Leopold, D.A. and Logothetis, N.K. (1996) "Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry," *Nature* 379, 549-553
- Milner, D. and Goodale, M. (1995) *The Visual Brain in Action*, Oxford: Oxford University Press.
- Nanay, B. (2010) "Perception and imagination: Amodal perception as mental imagery," *Philosophical Studies* 150(2), 239-254
- Nanay, B. (2013) *Between Perception and Action*, Oxford: Oxford University Press.
- O'Craven, K. and Kanwisher, N. (2000) "Mental Imagery of Faces and Places Activates Corresponding Stimulus-Specific Brain Regions," *Journal of Cognitive Neuroscience* xii(6), 1013-23.
- Olson, I.R., Plotzker, A., and Ezzyat, Y. (2007) "The Enigmatic temporal pole: a review of findings on social and emotional processing," *Brain* 130(7), 1718-31.
- Prinz, J. (2007) *The Emotional Construction of Morals*, Oxford: Oxford University Press
- Rosch, E.H., Mervis, C.B., Gray, W.D., Johnson, D.M., and Boyes-Braem, P. (1976) "Basic objects in natural categories," *Cognitive Psychology* 8(3), 382-439
- Schroeder, T. and Matheson, C. (2006) "Imagination and Emotion," in S. Nichols (ed.) *The Architecture of Imagination*, Oxford University Press, New York.
- Spaulding, S. (this volume)
- Van Leeuwen, N. (2011) "Imagination is where the Action is," *Journal of Philosophy* 108(2), 55-77.
- Van Leeuwen, N. (2013) "The Meanings of 'Imagine' Part I: Constructive Imagination," *Philosophy Compass* 8(3), 220-230.
- Van Leeuwen, N. (2014) "The Meanings of 'Imagine' Part II: Attitude and Action," *Philosophy Compass* 9(11), 791-802.
- Weisberg, D. S., Sobel, D. M., Goodstein, J., and Bloom, P. (2013) "Young children are reality-prone when thinking about stories," *Journal of Cognition and Culture* 13(3-4): 383-407
- Williamson, T. (this volume)
- Woodward, B. (2006) *State of Denial: Bush at War, Part III*, New York: Simon & Schuster.
- Zajonc, R. (1984). "On the Primacy of Affect," *American Psychologist* xxxix(2), 117-23