12
Creativity

12.1 Introduction

Suppose that you are writing a book on imagination. Your next-door neighbor engages you on the topic: "So…it's about creativity?" If you're a philosopher of mind in the analytic tradition, you may be tempted to say no—to explain that being creative is one thing (a personality trait, say), while imagining is the use of a type of mental state or process that may, or may not, be closely linked to being creative. This is not an uncommon move.[1] But you can tell how it will sound before you say it: like a cop-out. People want a theory of imagination to shed some light on the nature of creativity. Even if there is a distinction to be made between "being creative," on the one hand, and the mental state of imagining, on the other, there is no margin in denying an important link between the two. If a theory of imagination leaves creativity a mystery, so much the worse for the theory. In this closing chapter, I will do my best to face up to the difficulties that creativity presents to a theory of imagination, even if much will remain unresolved.

Even on its face, creativity has an air mystery. As with imagination, it's tricky to specify the nature of the phenomenon we want explained when we ask for an explanation of creativity. It's easy to end up following the wrong scent, diverted by questions that don't get to the heart of the matter. It's also easy, as philosophers, to suffocate or oversimplify the topic in our efforts to say something precise. So, in the first half of this chapter, I want to focus on distinguishing the most interesting and difficult questions about creativity from others that might catch our eye. To that end I will discuss some recent work by philosophers of mind that, I think, misunderstands the nature of the explanatory task that creativity presents to us. Then I will move toward articulating and answering some of the most important questions about creativity, connecting those answers back to this book's broader project of explaining imagination.

---

[1] Currie & Ravenscroft (2002, pp. 8–10) make essentially this claim in explaining why they will not be discussing "creative imagination" in their book devoted to "recreative imagination." While they grant that some instances of creativity may rely upon recreative imagining, they do not see it as a limitation on their theory of recreative imagining that it shed no light on creativity. In their view, "getting things done by using recreative imagination is not definitive or even criterial for displaying creative imagination" (p. 11). While I agree that (in my terms) neither A-imagining nor I-imagining is sufficient for creativity, it is still, I think, a strong desideratum on any theory of A-imagining (or "recreative imagining") that it meaningfully connect imagination to our capacity for creativity.

12.2 Creativity and A-imagining

Creativity, as a feature of persons, is nothing other than the ability to make creative products. To understand our own creativity, then, we need to know what it is that makes a product—be it a work of art, a marketing plan, or a car design—creative. Happily, there is a general consensus on two features being necessary and jointly sufficient for a product's being creative: novelty and value (Boden, 2004; Carruthers, 2018; Gaut, 2003; Paul & Kaufman, 2014).[2] The novelty condition, intuitively, is that there be something original, uncopied, or distinctive about the work.[3] The value condition holds that there is something *good* or *useful* about the product's novelty. Here the notion of value is to be taken as entirely general, extending from monetary, to practical, to aesthetic value; to be valuable is simply to be valued by people. Thus, *merely* novel creations—extemporaneous scribbles and babbles—do not clear the bar as creative; nor do hundred dollar bills, sapphire rings, and other merely valuable objects. (Whether there could be an *aesthetically* valuable product that was not novel is less clear; supposing that Warhol painted 100,000 Campbell's Soup cans, would each one be aesthetically valuable? Each one creative?)

Bones can be picked with this characterization of a creative work, certainly. There may be some creative works with little value and some with little novelty. But value and novelty at least serve as characteristic features of creative works. If we can explain how it is that people are able to produce works that are highly valuable in their novelty, we will have gone a good way toward explaining creativity.

Imagination enters the picture when we consider the grounds for a person's creativity. To be creative you need a good imagination—you need to be *imaginative*. This might be a controversial psychological claim if it assumed a specific theory about the kinds of mental states, processes, or representations involved in being "imaginative." But here I only mean to voice a platitude. There are no creative geniuses lacking in imagination; and there are no creative acts in which the creator's imaginativeness played no role. This is just because being creative and being imaginative come to the same thing, as a matter of platitudes. What about the different *senses* of imagining—A-imagining and I-imagining—distinguished

[2] "Creativity is the ability to come up with ideas or artefacts that are *new, surprising, and valuable*" (Boden, 2004, p. 1).

[3] Following Boden (2004), we can distinguish products that are novel with relation to the creator's personal history—in that they are not a matter of the person simply repeating something she has seen someone else do, or that she has done before—from products that are *historically* novel, in that they are significantly unlike anything the culture has previously witnessed. Boden calls the former "Psychological Novelty," and notes that psychologically novel products can be valuable and praiseworthy even when they are not historically novel. As historically novel products are a sub-set of the psychologically novel products, we can confine our investigation to psychologically novel products without loss. (Boden follows the same procedure.)

in earlier chapters? How do they relate to "being imaginative" in the sense (trivially) equivalent to being creative?

There appears no very tight connection between I-imagining—i.e., the having of (seemingly) image-like mental states—and being imaginative. Mental images likely feature in the thoughts of imaginative people, sure. But they're also present in the thoughts of the least imaginative and least creative among us. This must be true if, as is generally thought, imagery is an essential feature of much episodic memory (Schacter, Addis, & Buckner, 2007; Michaelian, 2016b). Some creative products will certainly rely, for their conception, on the use of mental imagery. But some will also rely, for their conception, on non-imagistic mental representations as well, supposing there are such. There is no obvious reason to think that having thoughts in a particular (imagistic) *format* is closely linked to the general capacity for creativity.

On the other hand, imaginativeness does seem closely tied to the notion of A-imagining. Imagining, in the A-imagining sense, is the having of rich, elaborated sequences of thought about the possible, fantastical, or unreal, in an epistemically blameless way. The notion of an imaginative, creative person is of someone who is good at thinking of unobvious possibilities and new ways of doing things that result in valuable products. It seems clear that the former would support the latter. How else are we to think of unobvious possibilities than by thinking—in rich, elaborated ways—about the merely possible? And how else are we to arrive at new ways of doing things than by considering many different possible actions and scenarios and selecting from among them? Having a basic capacity for A-imagining may not be sufficient for being imaginative in the sense that is equivalent to being creative. But it appears at least necessary.

This link between A-imagining and being imaginative is still a superficial one, however, as it doesn't commit us to any substantive thesis about the nature of the mental states and processes involved in A-imagining. In earlier chapters, I've argued that the A-imagining that occurs during pretense, conditional reasoning, and fiction-appreciation can be explained in more basic folk psychological terms—as drawing exclusively on beliefs, desires, intentions, and so on. To the extent that creativity requires conditional reasoning—as it surely does, at times—those prior arguments can be applied to the problem of creativity in showing how (some of) the A-imagining required for creativity is explicable in more basic folk psychological terms. Indeed, I think there is much to be said for the view that creativity draws on quite general reasoning and problem-solving skills (Weisberg, 2006) including, especially, abilities for conditional reasoning.

But there is much more to the story than that. There are mental processes relied upon in the generation of the most valuable creative works that, I will argue, are not well characterized as inductive or deductive *inferences* of any sort. This puts pressure on the idea that they could be characterized as bouts of conditional reasoning—or, indeed, as any kind of *reasoning* at all. Creative cognition, and the

A-imagining it involves, is less rule-governed and more associative—less proof-like and more dreamlike—than inferential trains of thought. My aim in what follows is to shed some explanatory light on these cases of A-imagining, some of which are unlike anything else so far discussed in this book.

12.3   The Easy versus the Hard Problems of Creativity

We first need to recognize that creativity comes in degrees. Creativity is in that way like intelligence. No healthy human has literally *none*. We are all capable of being *somewhat* creative. This leaves us with a choice in what to take as our target when trying to explain creativity. One option is to focus on the question of how creativity is at all possible in human beings. Here the goal would be to reveal some cognitive abilities, states, or processes shared by everyone, without which creativity would be impossible—perhaps contrasting humans to some species of animal (goldfish?) that lacks creativity altogether, precisely because it lacks this kind of state, faculty, or cognitive ability. Alternatively, we could ask, more narrowly, what it is that separates the exceptionally creative from the rest of us. My focus will be on this second question. I want to know what it is about Dylan, Dickinson, Bjork, and Nabokov that sets them apart. If this cannot precisely be known, then I at least want to brush against the answer and feel its outlines. This will require comment on the kinds of abilities and mental processes exploited in more modest creative efforts. Keeping one eye on the masters, however, will ensure that we don't undersell the problem.

Philosophers and psychologists often focus instead on the first question, addressing the issue of how creativity is at all possible in creatures like us (Carruthers, 2002, 2011, 2018; Stokes, 2014). This is due in part to the *prima facie* clash between the kinds of computational theories of cognition featured in many naturalistic accounts of mentality, on the one hand, and the *novelty* and *freedom from constraint* associated with creativity, on the other. It seems odd, to many, to suppose that a computer program following fixed algorithms could be capable of generating the sort of novelty required in a creative product.[4] If we are indeed sophisticated computers of a kind, the question extends equally to ourselves. The puzzle grows more acute in a context—such as contemporary philosophy of mind—where inductive and deductive inference are the primary psychological transitions formally modelled and discussed. On the face of it, creativity requires a capacity for a-rational trains of thought—an ability to transcend or sidestep anything so pre-determined as a rational inference.[5]

---

[4] Boden (1998, 2004) has done much to challenge this assumption, however.
[5] See Carruthers (2018) for a recent attempt to deal with this problem by inserting quasi-random associative processes within an otherwise deterministic, mechanistic cognitive architecture to achieve "constrained stochasticity."

But once we take account of the fact that creativity comes in degrees—granting that every child's stick-figure portrait is *a bit* creative—the nature of the puzzle can begin to elude us. For there are already computer programs capable of generating novel musical pieces (Cope, 1991), abstract visual art (Cohen, 2009), poems (Vandegrift, 2016), and even "co-authored" novels (Olewitz, 2016). Visual artworks created by Cohen's AARON algorithm have been displayed in major museums; free-verse written by Scholl's poetry-generating algorithm has been selected (unwittingly!) for publication in literary magazines (Merchant, 2015). Existing programs are capable of generating, from a photograph, an impressionistic image in the style of Monet, and—more impressively, in my view—of flipping the processing so as to create a realistic photograph-like image from an existing Monet (Zhu et al., 2017). The elegance and sophistication of such products will continue to increase. Google recently initiated the Magenta project, aimed at using machine learning to "generate compelling art and music" (https://magenta.tensorflow.org/). Its researchers have already moved well beyond the problem of generating novel musical works to focusing on how to create AI capable of composing artistic products with valuable aesthetic characteristics. For instance, one project aims to create AI that can compose and perform music with expressive *feel* (https://magenta.tensorflow.org/performance-rnn). Thus, it is the creation of works of considerable aesthetic *value*—and not mere novelty—that currently poses the most significant challenge to researchers in artificial intelligence.

A response to consider at this point, however, is that it is the *writer* of such programs who is the creator of the resulting artworks and not the programs themselves. If that were the case, then even the capacity for "mere" artistic novelty might remain beyond the grasp of such programs. This would leave the door open to holding that minimal creativity is, for all we know, something we cannot model on a contemporary computer. But there is little basis for such skepticism. A human being is no more responsible for his or her own creation than is a computer program. And yet we manage to be creative. We would be no less creative were it revealed that, like computer programs, we are the products of an intelligent designer. Unless one is going to take issue with the entire project of explaining human cognition in computational terms, there is no reason to deny that some current programs display *a degree of* creativity in generating the products that they do. After all, we wouldn't dream of calling our children uncreative if they presented us with the same.

My point here is not to suggest that computer scientists already understand and can perfectly model creative cognition. It is to focus attention on what most needs explaining. Notwithstanding current AI, there is much about creativity we do not yet understand. Distinctions from Margaret Boden (2004) help to pinpoint what that is. (NB I will put these distinctions to my own ends and do not mean to hold Boden to my understanding of them—only to credit her where due.) Boden

distinguishes three forms of creativity: combinatorial, explorative, and transformational (pp. 3–7).[6] *Combinatorial* creativity consists in "making unfamiliar combinations of familiar ideas," typically by means of an interesting analogy. "Think of a physicist comparing an atom to the solar system," she writes, "or a journalist comparing a politician with a decidedly non-cuddly animal" (p. 3). Ideas are "combined" in the sense that two things not normally associated are shown to have an interesting, even enlightening connection. The combination need not be by analogy; it can simply involve putting two things together—two ingredients, say—in a new and valuable way.

Explorative creativity, on the other hand, occurs in the context of a discipline or practice with a codified set of rules or principles—such as chess, or landscape art. It is possible to work entirely "within the rules" of chess to arrive at a creative move—a novel means of attack—just as it is possible to work entirely within the rules of deductive logic to arrive at a creative proof, or to use one's understanding of visual perspective to make a creative painting of a new landscape. In each case, Boden proposes, we do so by exploring the possibilities available within a particular "conceptual space." In the arts, the "rules" or "principles" explored can be understood as the basis for a particular *style*—be it the style employed by a caricature artist in the park, by John Coltrane in his rendering of jazz standards, or in the impressionisms of Monet. Once there is a recognizable style in play, we can—at least potentially—extract a set of principles or heuristics that can be used to generate additional works in that style. When a work manages to be creative through the artist's exploration and exploitation of these principles and heuristics, we can say it has explorative creativity. Boden analogizes *explorative* creativity to the driving of new routes on roads that are already created. I will understand it, somewhat less metaphorically, as a kind of creativity grounded in one's ability to draw out inferences from a set of premises, in accord with a fixed set of rules. So defined, it remains an open question whether there are indeed such a set of principles or rules underlying, say, Coltrane's improvisatory stylings, or Monet's impressionism. The point is simply that explorative creativity is the kind of creativity that occurs through the exploration and application of such rules and heuristics, where available.

It is only through *transformational* creativity that the limits of a particular style are transcended, the governing rules rewritten, and something occurs that cannot comfortably be characterized as an inference. This may consist in depicting objects in an entirely new way (think of the *first* cubist painting), deciding to ignore the requirement of rhyme (as in free verse), or using a formerly non-musical element as music (such as Jimi Hendrix's musical use of guitar-amplifier feedback)—something akin to the construction of new roads on which to drive. In Boden's

---

[6] These are in fact my terms for her three notions, which she does not explicitly name as such.

evocative terms, the creator changes the rules or style such that "thoughts are now possible which previously (within the untransformed space) where literally inconceivable" (2004, p. 6). These acts constitute what Boden sees as "the deepest cases of creativity." While I agree with Boden about the relative value of this kind of creativity, I want to suggest an amendment (or perhaps just a clarification) to how it is understood. Transformative creativity does not only occur when some new *style* of work is generated. It can occur *within* a style as well. A particular lyric or melody within a song of a recognizable genre may be so surprising, insightful, and inspired that it cannot be seen as an extrapolation of the style, but, rather, a tiny revolution within it. One doesn't have to be the first abstract expressionist for one's abstract expressionism to be transformative in its specific gestures; and a jazz musician needn't play something no longer recognizable as jazz in order for his performance to be transformative. For creativity to emerge in its transformative guise, the valuable novelty in a product simply needs to result from something other than the following of an explicit formula, or the teasing out of implications. It needs to result from something *non*-inferential. In what follows I will aim to provide a more positive characterization of what that is.

Despite emphasizing these three "kinds" of creativity, I doubt clear distinctions can be drawn in every case. We can expect borderline instances of each and creative products containing elements of all three. I have introduced them for their heuristic value in helping us to separate and more clearly articulate different questions about creativity and its relationship to imagination. We can begin to see how they help in that endeavor by considering a fairly ordinary case of creative cognition: the Build Challenge.

12.4 The Build Challenge

My son's fifth grade science teacher gives the class a Build Challenge at the end of each week. The students are split into groups, with each group receiving the same set of materials and a particular challenge to complete with those materials. For one challenge, each group was given six sheets of notebook paper, 10 inches of tape, and five paper clips with which to create the tallest free-standing structure they could. Completing the task required creativity. The students had never before faced the task of building a tall structure with precisely those materials; and they arrived at products unlike any they had made before, with varying degrees of value. (Groups caught copying the designs of other groups—i.e., *not* being creative—were disqualified.)

The winning group taped the top of a notebook page to its bottom, so as to make a cylinder. Then they stood the cylinder vertically on one end and set a second piece of paper across the top of the cylinder. That piece of paper served as the floor upon which a second cylinder was placed, its base secured with a second piece of tape. The process was continued up to three floors. This was, without

question, a creative solution to the problem. It required combinatorial creativity, insofar as students needed to find new ways of combining objects not normally combined. And it required explorative creativity, as students needed to explore the many possibilities for exploiting the known properties of the objects given. The important realization was that tape could be applied to a piece of paper, so as to make a cylinder, and that such a cylinder would have both height and sufficient stability for the process to be iterated. It did not, however, require transformational creativity—at least, not obviously. There was no new architectural concept or construction paradigm exploited. The paper was folded into a familiar shape; tape was used as tape. The winning design was something many of the students likely would have come to with a bit of prompting. The key was to focus on the right properties of the materials so as to infer their most effective combinations.

Despite the evident need for creativity in winning a Build Challenge, it is not hard to see the exercise as a reasoning task—as a bit of complex problem-solving. The desired end result was clear: the tallest structure in the class. The properties of the materials were well known. The problem was how to combine those materials to achieve the end result. (I faced a similar "build challenge" moments ago, trying to determine what sort of sandwich could be made from the items in the fridge.) There is nothing inherent in such problem-solving that cries out for a *sui generis* mental state that stands apart from one's beliefs, desires, and intentions. In exhibiting explorative creativity, the students needed to reason hypothetically about what would happen if the materials were arranged in a variety of different ways. And we have already seen that hypothetical reasoning does not, by its nature, require *sui generis* imaginative states (Chapters 5 and 6). It may have been that arriving at the best answer required students to generate mental imagery of the paper in different configurations—that mental imagery figured in the hypothetical reasoning itself. But we have also seen (Chapters 3 and 4) that the presence of imagery within a mental state is consistent with the state's being a judgment, such as: JIG (the paper taped end to end and turned on its side would be: **a vertical column**…). The combinatorial creativity displayed involved students' making the paper function like a column; the the notions of a sheet of paper and of a column were in that sense "combined." This can even be viewed as a case of object pretense: pretending that the paper is a column. Such pretenses, I argued in Chapter 8, do not require anything other than ordinary belief and desire.

All of this stepwise reasoning will be facilitated by the depth of one's knowledge of the materials in question and general experience with making structures of different kinds. But none of it calls out for states irreducible to beliefs, desires, and intentions. We can instead appeal to sequences of judgments and decisions that build on each other: the judgment that the paper can be made into a column leading to a decision to see what can be done with the paper in the shape of a column, for instance; this decision may then elicit a judgment that the paper can be secured to the desk with tape; then one may judge that another piece of paper

could rest horizontally on top of the secured column, and so on; these judgments lead to decisions to carry out the plan. If the plan doesn't work—if it turns out that the column is too narrow to support a piece of paper laying horizontally across its top—one has to go back to the drawing board, to act on, and develop, another hunch.

Return now to the platitude that creativity requires A-imagining. It looks like the A-imagining relied upon for *this* bit of creative cognition is explicable in more basic folk psychological terms. This will be possible, as a rule, whenever the relevant cognition can be characterized as a series of inferences—provided, again, that neither conditional reasoning nor pretense require *sui generis* imaginative states, as argued in previous chapters.

12.5 Losing the Scent—Recent Missteps in Linking (*Sui Generis*) Imaginings to Creativity

Nevertheless, most contemporary philosophers who have considered the matter seem confident that creativity draws on imagination as a central resource and that such imaginings are irreducible to other, more basic kinds of folk psychological states. One might think they were driven to this view by cases of creativity *unlike* the Build Challenge—cases not comfortably seen as problem-solving. But, in fact, to my knowledge, none of the considerations they offer introduce phenomena not already seen in the Build Challenge. I will survey a few of these views now, as a means to clarifying the real challenge that creativity presents to theories of imagination.

Some put forward *a priori* arguments of the kind dismissed in Chapter 1, which overlook the possibility of imagining being a complex folk psychological state. Here is Berys Gaut:

> Imagination is peculiarly suited…to be the vehicle of active creativity—to be that faculty we employ in being actively creative. For one can imagine various states of affairs without being committed to their truth or to carrying them out, so one can try out various options. In contrast, to believe some proposition is to be committed to its truth, and to intend something is to be committed to carrying it out if one can.   (2010, p. 1043)

As evidence for this analysis, Gaut—in earlier work—offers the following Moorean contrasts. There is awkwardness on the order of incoherence in saying: "*p*, but I don't believe that *p*," or "I will *x* if given the opportunity, but I don't intend to *x*." Yet there is no awkwardness in saying, "I am imagining that *p*, but I don't believe that *p*," or "I am imagining that I am *x*-ing, but I don't intend to *x*" (Gaut, 2003, pp. 279–80). This, he concludes, is because imaginings "lack the intrinsic ends of

belief and intention," and, indeed, "seem to lack any intrinsic end at all…imagination thus exhibits a kind of freedom" needed for creative cognition (pp. 279–80).

Now, the question of interest here is not whether imagining is in some sense "free" and used in acts of creativity—both are platitudes we should accept. What matters is whether such imaginings can be understood in more basic folk psychological terms. Gaut offers the Moorean contrasts as evidence that they cannot, as they show imaginings to have different (or non-existent) "intrinsic ends." Yet, like other superficial arguments for imagination's irreducibility considered in Chapter 1, the Moorean contrasts only serve to establish that imagining that p is not the same thing as believing or intending that p. It leaves open the possibility that imagining that p—like suspecting that p—is a complex mental state that can be broken into simpler, more general parts of some other (folk psychological) kind. After all, we can quite sensibly say: "If p, then q, r, s, t, u, and v…but I don't believe that p." And I have made the case that at least some episodes of imagining that p amount to making judgments in favor of rich conditionals with p as their antecedent. In that case, imagining that p is indeed consistent with one's not believing that p *and* with the imagining's consisting in the use of *other* beliefs. Gaut, and a chorus of others, may object that not all cases of imagining that p can receive such treatment. There I agree. The question, however, is not whether there is one specific kind of belief, desire, or decision that can be equated with each case of imagining that p, but whether there is always *some* more basic state or other that any arbitrary instance of imagining that p might be. To establish that there is not requires subtler arguments than an appeal to general Moorean contrasts. For there is no *prima facie* reason to think that all the mental activities that satisfy the commonsense criteria for being A-imaginings must receive the same analysis; as earlier discussed, imagination appears heterogeneous to many *even on its face*.

Peter Carruthers, in a series of papers, links creativity to imagination by very different means, describing a developmental pathway from pretend play to creativity, with imagination at the root of each. Both creativity and pretense, he tells us, "can be seen as sharing essentially the same cognitive basis, in so far as both involve exercises of imagination" (2002, pp. 228–9). Carruthers does not explicitly argue that "exercises of imagination" cannot be viewed as the exploitation of more basic propositional attitude states. However, in elaborating his position, he approvingly cites theorists who explicitly endorse such a view—including Nichols & Stich (2000). And, in other work, Carruthers finds it "very plausible" that imagining, conceived of as a distinct cognitive attitude by Nichols & Stich (2000), "can't be reduced to believing, or to desiring, or to any combination thereof (nor can it be reduced to any sort of planning or intending)" (Carruthers, 2006, p. 89).

In explaining the connection he sees between pretense and creativity, Carruthers notes:

Each involves essentially the same cognitive underpinnings—namely, a capacity to generate, and to reason with, novel suppositions or *imaginary scenarios.* When pretending, what a child has to do is suppose that something is the case…and then think and act within the scope of that supposition…Similarly, when adults are engaged in the construction of a new theory, or are seeking a novel solution to a practical problem, or are composing a tune, they have to think: 'Suppose it were the case that P', or 'Suppose I did it like *this*' or 'Suppose it sounded like *so*'.   (2002, pp. 229–30)

In later, co-authored work, Carruthers places imagination within the context of the "GENPLORE" model of creativity developed by Finke and colleagues (1992), according to which creative action involves both a generative phase, where a new idea emerges, and an exploratory phase, where the idea is developed and expanded. Carruthers and Picciuto propose that imagination, *qua* supposition, is involved in both phases:

Perhaps what pretense does, then, is provide practice in making suppositions and reasoning within their scope, thus supporting both the 'generate' and 'explore' components of GENEPLORE creativity.

(Picciuto & Carruthers, 2014, p. 215)

On the one hand, it can certainly be granted, on a platitudinous level, that both pretense and creative action require us to generate suppositions and, further, that childhood pretense allows us valuable practice in doing so. This leaves open the question of whether such "supposings" or "imaginings" are basic cognitive states or, instead, reducible to others. As Carruthers (2006, p. 89) appears to maintain that the relevant supposings make use of a *sui generis* type of state or cognitive module, we are brought back to the earlier debates (of Chapters 5–8) concerning whether the imaginings at work in pretense and conditional reasoning must be viewed as *sui generis*, irreducible states. Like Nichols & Stich (2003), Picciuto & Carruthers cite the difficulties people with autism show with both pretense and (certain forms of) suppositional reasoning as evidence that there is a single type of state or module relied upon by both (2014, pp. 211–12). But, as noted earlier (Chapter 8), little can be inferred from the co-presence of any two deficits in autism; ASD is a multi-faceted developmental disorder affecting many aspects of cognitive, emotional, perceptual, and sensorimotor functioning. The fact that people with autism show a characteristically abnormal gait (Rineheart et al., 2006; Calhoun et al., 2011), for instance, does not warrant the inference that a single cognitive module underlies both pretense and ambulation.

In any case, the debate to be had with Carruthers over the nature of the imaginative states exploited in pretense, conditional reasoning, and—by

extension—creativity, has in large part already occurred in earlier chapters. I will not revisit it here. The important point for present purposes is that, in emphasizing the importance to creativity of thinking "within the scope of a supposition," Carruthers does not introduce or aim to explain any phenomenon that is not itself a kind of drawing-out-of-inferences (albeit "offline," through the use of *sui generis* imaginative/suppositional states). Nor has he introduced or explained any aspect of creativity that we didn't already face in addressing the Build Challenge, which also plausibly required hypothetical reasoning. If transformative creativity really is non-inferential in nature, as suggested earlier, then it is no easier explained by appeal to inferences drawn "in imagination" than it is by appeal to inferences drawn within one's beliefs.[7]

Dustin Stokes—also the author of several articles on imagination and creativity—takes an approach closer in spirit to Gaut's in arguing that creativity relies upon *sui generis* imaginative states. Stokes (2014) develops, to my knowledge, the most detailed philosophical argument for *sui generis* imaginings' being necessary to creativity. Stokes begins by distinguishing "truth-bound" mental states, such as beliefs, whose function it is to "faithfully represent the information of some conceptual space," from non-truth-bound states, which lack that function. He then proposes that creativity makes essential use of non-truth-bound states. This is because:

> Any cognitive state that functions to faithfully represent the information of some conceptual space—be that cognitive state a true belief, propositional or procedural knowledge, or a memory—can at best play a necessary but insufficient role in the thinking required for an accomplishment like [Bach's] *The Well-Tempered Clavier*. *Truth-bound cognitive states*...are rarely sufficient for creative thought.   (Stokes, 2014, p. 160, emphasis in original)

Stokes goes on to conclude that "imaginings" are the best candidates for being the relevant non-truth-bound states at work in creativity. Why, exactly, does creativity require the use of non-truth-bound states? In explanation, Stokes offers a thought experiment reminiscent of Frank Jackson's (1982) Mary, the brilliant neuroscientist who knows all the physical facts about visual perception but still doesn't know what it's like to see red. Stokes asks us to consider "Super-Bach," who:

---

[7] It is worth again mentioning Carruthers' more recent work (2018), in which he takes a very different (and I think promising) approach to explaining creativity, positing quasi-random associative processes that generate thoughts or ideas semantically related to a task at hand. I do not have space to discuss this newer view here. It bears noting, however, that he is silent, in that work, on the relation of imagination to these associative processes.

> Knew *everything* there was to know—both in terms of all the facts and all of the relevant skills—about the clavier, tempered tuning, and the 12-tone scale. This would not have been sufficient for the creation of *The Well-Tempered Clavier*…Super-Bach's knowledge (just like actual Bach's knowledge) of the space would indeed constrain his composition, but this knowledge alone would not amount to, afford, or even imply the musical work in question. This is for the simple reason that there is nothing in this conceptual domain, or cluster of domains, that includes or entails (by itself) *The Well-Tempered Clavier*. (2014, p. 157)

Certainly, it would be odd to propose that a piece like *The Well-Tempered Clavier* arrived to Bach as a kind of deductive or inductive inference from his standing knowledge—irrespective of any goals, desires, values, or intentions he may have had. But as soon as we give Super-Bach some desires and intentions—to write a new piece using the clavier, to incorporate the 12-tone scale in an innovative way—it is not *obvious* that any more exotic non-truth-bound state is required. As we saw above, Carruthers would envision Bach as reasoning: "Suppose it sounded like *so*." Granted, Super-Bach needs to have a reason for exploiting his existing knowledge and skills as he does. He needs goals, interests, and aesthetic values—all of which are non-truth bound states—that lead him to draw particular connections and try out new combinations. In that sense, non-truth-bound *motivational* states are indeed necessary, as they are for the creation of any artifact. But we don't yet have a clear role for *sui generis* imaginings.

What Stokes still finds missing is an ability for Super-Bach to engage in "cognitive manipulation." Unlike Super-Bach, actual Bach:

> Presumably imagined, working from within the constraints that he imposed for himself, how certain musical combinations and structures would achieve certain goals. He did not, as it were, simply read off or abstract from the relevant music-theoretic information. He had to manipulate, by use of the imagination, that information (or perhaps add to it) in ways unbound to accurately representing it. (2014, pp. 162–3)

Stokes's argument here is still difficult to follow. We can agree that no new music will arrive as an inductive or deductive inference from what one knows, without relevant goals and desires intervening. But what is it to use information "in ways unbound to accurately representing it"? When we *use* information we do not, in general, *represent* that very information in any way at all. What Stokes seems to have in mind is that, in cognitive manipulation, we often start with representations whose contents we believe and then move on to make use of related representations whose contents we do not believe—at which point the information ceases to be "truth-bound." But once we've seen that hypothetical reasoning, and "considering new possibilities" does not in general demand *sui generis* imaginative states

(Chapters 5 and 6), we need some other reason to think that the kind of cognitive manipulation at work in creativity does.

The issues get a bit clearer when we move away from a complex piece like *The Well-Tempered Clavier* to consider less exalted creative works. For according to Stokes, "even minimally creative thought and behavior require cognitive manipulation" (2014, p. 165). He offers the example of children who are asked, as part of experiments on creativity and cognitive flexibility, to draw "nonexistent houses, people, and animals." A typical work produced in response is a simple figure that combines elements of two categories—a house with eyes as windows, for example. On the one hand, such drawings were in general "quite predictable—largely generated in line with the relevant conceptual schemes" (p. 165). However, Stokes asks:

> Are any of these drawings possibly enabled *merely* by the relevant conceptual knowledge? We know that the subjects consistently deployed their concepts of HOUSE, PERSON, etc., to make their drawings; was this knowledge sufficient? No. The concepts of HOUSE and PERSON, no matter how rich, will not (by themselves) enable a child to draw a house with eyes for windows, a mouth for a front door, and arms and legs. These cross-category changes require the child to cognitively manipulate, rather than faithfully mirror, the conceptual space in particular, albeit minimal, ways. These drawings require non-truth-bound cognitive states.   (p. 165)

The key claim here is that drawing a house with eyes requires the child to "cognitively manipulate, rather than faithfully mirror, the conceptual space." What does this mean? Suppose that I am simply *told* to draw a house with eyes for windows. Does this require me to cognitively manipulate my conceptual space? I don't see why it would. I know how to draw a house. I know how to draw eyes. I follow the task instructions by drawing a house with eyes. This requires no adjustment to my concept of a house, or to my concept of eyes. Does doing this require that I enter into states whose contents I disbelieve? I don't think so. There's a perfect parallel here to the most simple cases of pretense discussed earlier (Chapters 7 and 8). To pretend to be a lion, I need to recall what lions are like and use that knowledge to make myself somewhat lion-like. But I do not need to represent that *I am a lion*; there is no need, either, to readjust or modify my concepts of lions or of myself. There is no need, then, to enter into any cognitive state that is in some sense "non-truth-bound." The same goes for making a drawing of a lion-like house: I draw a house and, recalling a lion's salient features, aim to mirror some of those features in the drawing. As Boden notes, this sort of novel category-combination is:

> easy to model on a computer. For nothing is simpler than picking out two ideas (two data structures) and putting them alongside each other…a computer could merrily produce novel combinations till kingdom come.   (2004, p. 7)

Would Stokes hold that, in achieving such combinations, the computer must exploit its own functional analog of non-truth-bound states? Presumably he must. But there is no clear reason why we should.

12.6 Back to the Deep Waters

Despite these criticisms, I think that Stokes is on the right track in observing that there is nothing in Bach's knowledge that "includes or entails" the various pieces he composes. The problem is that we don't get from Stokes a clear picture of why his knowledge, *together with* his goals and interests, is not a sufficient mental resource for the piece's creation. To move in that direction, let's consider a more complex creative activity: writing a story. Just as the science teacher assigns a "build challenge" with limited materials, we can imagine a creative writing teacher assigning the following *story challenge*: confining yourself to thus and such characters (a wife, husband, identical twin boys), thus and such setting (Los Angeles in the mid-1970s), and thus and such problems (a contaminated swimming pool; a heart attack), write the best story that you can. Writing the story will require, *inter alia*, reasoning hypothetically about how such individuals might interact, and what would happen if they did this, that, and the other. So far we have a close analogy to the build challenge. A difference is that the "materials" provided by the fiction writing assignment are more open-ended. Decisions will need to be made about their finer features. Do the twins get along, or are they rivals? We can decide that they are rivals and see how that affects the equation. If the story progresses well, we can stick with that decision about what is true in the story; if it does not, we can revise it. Was the pool contaminated by accident, or by design? Again a decision needs to be made about what is true in the story. Once it is made, the story can be developed with that element as one of the "materials." The process of making a decision about what is true in the story and extrapolating about what might happen in such a situation builds on itself cyclically until one arrives at a satisfactory product—a finished story. Writing the story, then, requires an ability to make decisions about the materials to use, to reason hypothetically about likely results of combining them, and to make aesthetic judgments about the value of those results. We can see plenty of *combinatorial* and *explorative* creativity in such work. But none of this creates any obvious need to draw upon states other than beliefs, judgements, decisions, and desires—not lest we re-litigate the debates of earlier chapters. The same points apply to many daydreams, which can be seen as a set of decisions about what is happening in a story we are telling ourselves. (See Chapter 4 for more on daydreams.)

I do think that a story could be written in the way just described—as a kind of interplay between one's goals, values, inferential capacities, and the constraints that have been set. Simply writing a story, then, does not require any *sui generis*

imaginative states. However, I don't think that a very *good* story could be written in this way! Or, to put a finer point on it: such an explanation of the psychology of story-writing does not give us any insight into how one should go about writing a good story. We will do well to focus on this gap in the explanation. For as we saw at the outset, the notion of a creative product is inherently value-laden—there has got to be something valuable about the novelty in the product. That value, like creativity itself, comes in degrees. And if we confine ourselves to explaining creative products at the low end of the value spectrum, we are likely to miss the most interesting puzzles creativity presents to us. We may come to see creativity as little more than the dogged application of reasoning heuristics and aesthetic rules of thumb in the service of one's goals and desires. This is nothing that is not already well simulated by current artificial intelligence. To move forward, we need to think more about the difference between a generic fiction and a great fiction, a passable song and a classic, a stick figure and a masterpiece. That is the project I turn to now.

12.7 Songwriters on Songwriting

Paul Zollo's (1997) interviews in *Songwriters on Songwriting* are a valuable resource for anyone interested in the nature of songwriting, and in artistic creation more generally. Zollo is able to draw accomplished songwriters into exceptionally deep, careful, and often hilarious reflection on their craft. I will quote a number of their responses at length to reveal a common theme. They all gesture at something less accountable, less conscious, less reasonable, and more dreamlike in the process of artistic creation than we have so far considered.

Paul Simon: "I don't consciously think about what a song should say. In fact, I consciously try *not* to think about what a song should say…As soon as your mind knows that it's on and it's supposed to produce some lines, either it *doesn't* or it produces things that are very predictable. And that's why I say I'm not interested in writing something that I thought about. I'm interested in *discovering* where my mind wants to go, or what object it wants to pick up…If it doesn't come to me in that surprising way, I don't tend to believe it or get excited about it…I mean, it may be that that's what is slowing me down to such a slow pace, you know, that I keep waiting for this stuff instead of just writing. But just writing what? How do I know what I'm going to write if I don't discover it? If I make up what I'm going to write, all I'm going to write is what I saw on television or what I read in the paper or what I saw…it's not going to be from the underground river of your subconscious. Because that just comes to the surface occasionally and you have to capture it when it happens."

(Zollo, 1997, pp. 95, 98, 120)

Bob Dylan: "It's nice to be able to put yourself in an environment where you can completely accept all the unconscious stuff that comes to you from your inner workings of your mind. And block yourself off to where you can control it all, take it down…There's no rhyme or reason to it, there's no rule…Still staying in the unconscious frame of mind, you can pull yourself out and throw up two rhymes first and work it back. You get the rhymes first and work it back and then see if you can make it make sense in another kind of way. You can stay in the unconscious frame of mind to pull it off, which is the state of mind you have to be in anyway…It's a magical thing, popular song. Trying to press it down into everyday numbers doesn't quite work. It's not a puzzle. There aren't pieces that fit. It doesn't make a complete picture that's ever been seen."   (Zollo, 1997, pp. 72, 81–2)

Leonard Cohen: "If I knew where the good songs came from, I'd go there more often. It's a mysterious condition. It's much like the life of a Catholic nun. You're married to a mystery…Things come so damn slow. Things come and they come and it's a tollgate, and they're particularly asking for something that you can't manage. They say, 'We got the goods here. What do you got to pay?' Well, I've got my intelligence, I've got a mind. 'No, we don't want that.' I've got my whole training as a poet. 'No, we don't want that.' I've got some licks, I've got some skills with my fingers on the guitar. 'No, we don't want that either.' Well, I've got a broken heart. 'No, we don't want that.' I've got a pretty girlfriend. 'No, we don't want that.' I've got sexual desire. 'No, we don't want that.' I've got a whole lot of things and the tollgate keeper says, 'That's not going to get it. We want you in a condition that you are not accustomed to. And that you yourself cannot name. We want you in a condition of receptivity that you cannot produce by yourself.' How are you going to come up with that?"   (Zollo, 1997, p. 335)

With respect to writing the song "Anthem," Cohen remarks:

"I didn't start with a philosophical position that human activity is not perfectible. And that all human activity is flawed. And it is by intimacy with the flaw that we discern our real humanity and our real connection with divine inspiration. I didn't come up with it that way. I *saw* something broken. It's a different form of cognition."

Neil Young: "Usually I sit down and I go until I'm trying to think. As soon as I start thinking, I quit…then when I have an idea out of nowhere, I start up again. When that idea stops, I stop. I don't force it. If it's not there, it's not there, and there's nothing you can do about it…There's the conscious mind and the subconscious mind and the spirit. And I can only guess as to what is really going on there."   (Zollo, 1997, pp. 354–5)

On the one hand, each of these songwriters has a mass of articulable knowledge about songwriting—theories, strategies, influences, and goals they *can* put into words and consciously exploit. To give a small sampling from elsewhere in the interviews: Dylan recommends changing the key of a song when stuck on where to go with it, or working backwards to a lyric with a particular rhyme in mind. Simon bounces a rubber ball to achieve the right focus and emphasizes finding a first line for a song that "has a lot of options"; Cohen works in the morning but recommends trying everything: "thought, meditation, drinking, disillusion, insomnia, vacations…" We can include among these the strategy of *trying not to consciously think* about what the song should be like. In addition to these little "tricks," each has mastered the (fungible) "rules" of different musical genres and knows a great deal of music theory—concerning keys, song structures, harmony, and so on. They know what roads there are to drive on and have paved many of them themselves. Without question, prior practice and expertise are essential components to the kind of creativity we most value.

And yet, for all their practice and expertise, and all their tricks and heuristics for facilitating creativity in themselves, these songwriters are still *waiting*…they are waiting on something else. Their practical skill, theoretical knowledge, goals, values, and desires only take them so far—far enough, I think, to write a passable song, a creative song. But not far enough to write a *great* song. For the latter, something else is required. This "something else" is not an active examination of logical space through the use of non-truth-bound mental states. In trying to explain what it is, they each appeal instead to "the subconscious"—or, for Cohen, to "a condition of receptivity that you cannot produce by yourself." It is from within this condition that *transformative* creativity may emerge—creativity that they are at a loss to explain in terms of *reasons* and that, therefore, is not happily analyzed as kind of means-to-ends *reasoning*. Insights from across the arts and sciences no doubt often spring from the same obscure source. In drawing on it, one appears no longer simply to be combining elements in predictable ways, or exploring the possibilities within a fixed logical space. I want to investigate this condition in the balance of this chapter, to see what more can be said of it and its relation to imagination.

It must first be granted, however, that these songwriter reflections don't constitute an explicit argument that transformative creativity draws on some mental resource "not happily analyzed as a kind of hypothetical reasoning." I think their comments are, instead, *strongly suggestive* of the need for some other account—one that doesn't see artistic creation as a stepwise inferential process. Should that hunch prove incorrect and transformative creativity simply a rarefied form of deductive, inductive, or abductive reasoning, this would only render our work in explaining creativity easier, as I've already shown how a good deal of A-imagining can be understood as the making of inductive and deductive inferences of different kinds (including in favor of conditionals). The problem I saw with Stokes' and

Carruthers' examples was that they didn't highlight the need for anything *other than* ordinary hypothetical reasoning in creativity. The songwriter reflections give us a better sense of where we need to go if we are to face up to the full challenge that creativity presents.

12.8 Creativity and the Subconscious

Let's suppose that transformative creativity very often, or even always, draws upon the subconscious, as these songwriters propose. What would this tell us about the role of imagination in transformative creativity? On the one hand, it is natural to say, as a kind of platitude, that songwriters are relying upon their imaginations in their moments of transformative creativity. And it seems reasonable to add that they are (at least often) engaging in rich thought about the possible and the fantastical, in an epistemically blameless way. To that extent, they are engaged in A-imagining. On the other hand, two platitudinous features of imagination appear at odds with its being involved in such cases. First, imagining—in both the I-imagining and A-imagining senses—is a paradigmatically *conscious* phenomenon, familiar to our mental lives. Yet the songwriters emphasize the need to draw upon the subconscious, and the importance of quieting the conscious mind. This is one reason to think that the imagining they undergo—subconsciously—is not the same sort of process we are familiar with from everyday imaginative acts. Second, one of the key characteristics that, for many, distinguishes A-imagining from believing or judging is its being subject to the will. We can imagine whatever we *choose* to imagine, it is said, and this is the source of its freedom. Yet being subject to the will in this way cuts directly against everything the songwriters have to say about the resource needed for transformative creativity. Whatever it is that lies in the subconscious, it cannot be forced or willfully controlled. Songwriting is instead "a waiting game," a "mystery," with no "rhyme or reason" to it. Try as one might to purposefully bring it about, "nothing works," Cohen laments. To the contrary, when he feels himself *trying* to think of the next line, Neil Young knows he needs to stop.

In earlier chapters I proposed reductive accounts of A-imaginings, such as occur during ordinary pretense or hypothetical reasoning. It is possible—*just* possible—that the kind of (putatively) subconscious cognitions relied upon for transformative creativity are subconscious versions of these same processes. In that case, were we to shine a light on the subconscious during transformative creativity, we'd simply see more of what we are already familiar with from our conscious episodes of action planning and conditional reasoning (whatever that might be!). But I doubt that is the case. Why, after all, would it be so important for the process to occur subconsciously, and without being subject to the will, if the very same resource were available consciously? We should instead consider other

possibilities for the kind of mental process that might realize A-imaginings in this context. While *some* A-imaginings relied upon for creative cognition may be much the same as those underlying pretense, conditional reasoning, and the appreciation of fictions, others—including especially those required for trans-formational creativity—may be quite different. To understand them, I will argue we need to consider a role for *associative* thought processes—ways of transition-ing from one mental state to the next that are not determined by principles of reasoning (be they inductive, deductive, or probabilistic), but that are also not at all *random*. (The difference between an associative transition in thought and a deductive or inductive one will be clarified momentarily.)

Before doing so, a word on the relation of transformative creativity to some competing accounts of A-imagining. As we have seen in earlier chapters, the most fully-developed theories of A-imagining focus on its role in guiding pretense, facilitating conditional reasoning, and explaining our engagement with fiction. A key feature of many such accounts is their proposal that the dynamics holding among *sui generis* imaginative states mirror those that hold among beliefs with the same contents (Currie & Ravenscroft, 2002; Doggett & Egan, 2012; Nichols & Stich, 2000; Weinberg & Meskin, 2006b; Williamson, 2016). Such imaginative states are said to be "belief-like" in this respect. The fact that they develop in ways that mirror belief-like inductive and deductive inferences is supposed to explain why they can be relied upon when reasoning our way to conditionals, recovering inexplicit fictional content, and generating pretenses that follow realistic scripts. Yet this cuts against their suitability for explaining transformative creativity, which, it appears, requires something other than drawing out implications from a set of premises. These theorists may respond that it is only a norm or regularity of the poisted *sui generis* states that they have belief-like inferential characteristics— that this "default mode" can be suspended in acts of transformative creativity. Would that open the door to such states serving in explanations of transformative creativity as well?

Perhaps, but only at the cost of turning one's theory into mush. These *sui generis* imaginings would then, at times, be paradigmatically conscious mental states, under volitional control, with tight belief-like inferential characteristics and, at other times, unconscious states unbounded by any inferential constraints and beyond control of the will. It is not flatly impossible for there to be a mental Jekyll-and-Hyde of this sort, with one set of causes and effects in one case, and quite the opposite in another. But we might then wonder: is there anything imagination *can't* do? Could it then be that some of what we *thought* were ordin-ary beliefs and desires are in fact *sui generis* imaginings that have ditched their default mode and are now behaving *exactly* like beliefs and desires? It's an unhappy possibility to consider, bordering on incoherent. At a minimum, it would require a way of identifying and typing states other than by their causes and effects. If there is any default position in this area, it should be that the states

exploited in transformative creativity (alluded to by the songwriters, above) and hypothetical reasoning are different in kind. This is compatible with their still being A-imaginings, however, so long as we allow A-imaginings themselves to be heterogeneous.

In any case, I don't think this criticism would come as an unwelcome surprise to the philosophers in question. Currie & Ravenscroft are explicit in distinguishing "creative imagination" from "recreative imagination," holding they are only theorizing about the latter (2002, pp. 9–11), while Nichols & Stich (2000) and Williamson (2016) never make any effort to extend their views to creativity, nor to other subconscious processes beyond control of the will. The point is worth emphasizing, however, that *whether or not* one posits *sui generis* imaginative states to explain pretense and hypothetical reasoning, there remain *further* challenges in characterizing transformative creativity. We cannot "buy" explanations of both by positing a single *sui generis* type of imaginative state without, in effect, positing a state with quite distinct functional roles in different contexts. And that just raises the question of whether we indeed have just a single type of state at work in each context. In short: we all need to say something *new* about transformative creativity.

12.9 Creativity and Associationism

The songwriter reflections are intriguing in the way they highlight the need for mental transitions beyond conscious control—transitions that are difficult to characterize as *inferences*. As mysterious as this may seem in the context of creative cognition, it is not so unusual that our thoughts unfold in non-inferential patterns beyond conscious control. Suppose that you are asked to name as many animals as you can in thirty seconds. Will you be able to explain why you thought of the specific animals you did, and not others? Perhaps you will think of a farm and use that setting to generate ideas. This will still leave open the question of why some farm animals came to mind, and not others—and why a farm itself came to mind, and not some other place where animals are found, such as a zoo, a forest, or an African savannah. In answer to such questions, we can of course *speculate*. ("It's probably because I saw a copy of *Animal Farm* on the table this morning.") But, in most cases, we will just be making an educated guess as to why a specific setting or animal sprang to mind. It will not be that we made a choice to think of animals in that order, or desired to think of a farm more than a zoo. Folk psychological explanation comes up empty-handed in such cases, insofar as there are no beliefs, desires, or intentions available to explain why we first settled on the farm as a prompt, or why a chicken was the third, instead of the fourth, animal that came to mind.

In this sense, there are quite unremarkable aspects of ordinary cognition that are no more *rationalizable* than Paul Simon's spontaneous arrival at the lyric,

"There's a girl in New York City who calls herself the human trampoline."[8] This is not to say there is no explanation *at all* for these things. It is instead to say that the explanation will not appeal to a sequence of folk psychological states deployed in an act of practical reasoning. And where there are no steps of reasoning to be adduced, the transition in thought cannot be considered an inference. For inferences, traditionally understood, are simply transitions in thought that follow a reason scheme (such as *modus ponens*, or statistical induction), whereby one thought provides the reason for the next.[9]

Yet, despite our inability to explain the order of animals uttered as the result of an inference, there may be associative explanations available, whether or not we are in a position to articulate them. The same may be true for Simon's lyric, even if the associations are more obscure. So we will do well to consider the nature of associative explanation in more depth, as a means for broadening the tools we have for thinking about creative cognition in cases where it outruns anything we can comfortably characterize as an inferential processes.

We can begin with some background on associationism as a general approach to understanding transitions in thought. Associationist theories of thought have had adherents from the British empiricists (most notably Hume), through Skinnerian behaviorists, to present-day connectionists. Associationist principles are also alive and well within contemporary empirical psychology—particularly evident in work on so-called "semantic networks" (Collins & Loftus, 1975; Goñi et al., 2011; Sowa, 1991). While these approaches all differ in important ways, each is committed to what Eric Mandelbaum (2017) calls, somewhat pejoratively, "a certain arationality of thought." On any associationist approach, Mandelbaum explains:

> A creature's mental states are associated because of some facts about its causal history, and having these mental states associated entails that bringing one of a pair of associates to mind will, *ceteris paribus*, ensure that the other also becomes activated.

What leads one thought to follow another, on an associationist picture, is not adherence to a deductive or inductive reason scheme, but rather "facts about [one's] causal history." Unlike the inductive or deductively warranted thought

[8] In Simon's telling: "That line came to me when I was walking past the Museum of Natural History. *For no reason* I can think of. It's not related to anybody or anything. It just struck me as funny" (Zollo, 1997, p. 110).

[9] John Pollock connects the dots between inferences, reasons, and reasoning as follows: "Reasoning proceeds by constructing arguments for conclusions, and the individual inferences making up the arguments are licensed by what we might call reason schemes" (Pollock, 2008, p. 451). Reason schemes, for Pollock, are equivalent to inference rules—such as *modus ponens*, in deductive reasoning, or the statistical syllogism of probabilistic reasoning (viz., "This is an A and the probability of an A being a B is high. Therefore, defeasibly, this is a B."). The individual inferences of an argument collectively make up the reasons for the conclusion.

transitions that are the centerpiece of heavy-duty conceptions of folk psychology (discussed in Chapter 2), associative thought transitions "are not predicated on a prior logical relationship between the elements of the thoughts," and are "not based on the logico-syntactic properties of thoughts" (Mandelbaum, 2017). They are based instead on associative relations among the thoughts, which consist in two thoughts having been causally contiguous in the past (e.g., "The bell is ringing. Food is being dispensed."), or bearing some other non-logical, causal relation (such as contiguity in a neural semantic network (Collins & Loftus, 1975)). The causal facts responsible for two thoughts being associated may at times contrive to elicit thought transitions mirroring those of someone following a particular reason scheme. But they need not by nature. Associative transitions in thought, Mandelbaum notes, can be:

> Just a stream of ideas that needn't have any formal, or even rational, relation between them, such as the transition from THIS COFFEE SHOP IS COLD to RUSSIA SHOULD ANNEX IDAHO, without there being any intervening thoughts. This transition could be subserved merely by one's association of IDAHO and COLD, or it could happen because the two thoughts have tended to co-occur in the past, and their close temporal proximity caused an association between the two thoughts to arise (or for many other reasons).
>
> (Mandelbaum, 2017)

Taken as a claim about the nature of *all* thought processes, associationism is controversial and at direct odds with the "heavy-duty" computational theory of mind defended by Fodor and others (see Chapter 2). Less contentious is the claim that at least some thought processes are associationist in nature. Consider, for instance, *free association*. In naming things we associate with, say, Isaac Newton (apples, gravity, wigs, Fig Newtons...), different ideas or concepts trigger each other, due to their being associated with the initial concept. We are not drawing inferences but simply associating one thing with another. Associative transitions can also plausibly occur among propositional thoughts, and not merely concepts.[10] Many ordinary cases of propositional memory—remembering that p, because one has just judged that q—are plausible examples. Judging that you will need to drive to work today (= q), may, for instance, remind you that the car could use an oil change (= p). It is not surprising that thinking of driving would prompt you to

---

[10] Associationism comes in both "pure" and "impure" variants, where impurity amounts to thoughts having structure (such as predicative structure) over and above anything that can be accounted for in purely associative terms. A common challenge to pure associationist theories is that they cannot account—in purely associationist terms—for the difference between thinking about x and then about F (because being F is associated with x), from thinking that x is F (Fodor, 2003; Mandelbaum, 2017). I consider an impure associationist account here—where not all thought must be explained in associationist terms—as it is likely to have wider appeal.

remember that you should change the oil. But neither is this case of remembering well viewed as an *inference*, made in accordance with an abstract reasoning scheme. Thinking about driving to work simply reminded you that the car could use an oil change.

Less obvious, and more controversial, is whether we acquire *new* beliefs through associative processes. Mandelbaum's example of moving from the thought that "this coffee shop is cold," to "Russia should annex Idaho," would be an example of this, supposing the second thought were a newly acquired belief. Let us suppose, for the sake of argument, that we do at times arrive at new beliefs through associative processes. As Mandelbaum notes, calling such a process an *inference* would be something of an oxymoron, insofar as the notion of an inference carries with it the implication of a rule or principle relating the relevant propositions, in virtue of which one provides rational warrant for the other. A dog's associative thought transition from "The bell is ringing" to "Food is being dispensed" may have the air of an inference, but the principle that links the two is brute causal contiguity. Any thought at all, on any topic, that comes to follow "The bell is ringing" due to constant causal contiguity would qualify as the same kind of "inference."

Where a proper inference occurs, there is the possibility of providing a folk psychological description of the states it involved. For to give a folk psychological explanation for why one has inferred what one has is simply to attribute a collection of thoughts that *rationally warrant* those that follow them. And a thought rationally warrants another insofar as the transition from one to the next is explained by the following of a reason scheme. Accordingly, when we find ourselves at a loss to explain—even to ourselves—a certain sequence of ideas or thoughts we have had, it is often because there is no way of seeing the sequence as a set of reasonable *inferences*, and hence no available folk psychological explanation for why we should have thought what we did. In such cases, it will be tempting to attribute their genesis to something beyond conscious control—to something, perhaps, "in the subconscious." In some cases, the appeal to the subconscious may be quite correct. But another possibility is that the entire process occurred consciously, and that we are unable to *explain* it just because it was not an inference—and, hence, not something to be rationalized via the attribution of beliefs, desires and a reason scheme governing their relations.

Consider again the case of generating animal names as quickly as you can: it could be that there is unconscious reasoning taking place that explains why we come up with the specific animals we do, and that this is why we are unable to give any clear account of why we named just those animals in that order. Alternatively, it could be that the animals that were chosen each followed each other in the order they did due to associative factors we are simply unable to articulate. In that case, the relevant processing was only unconscious in the sense that we do not have conscious awareness of the associative principles that move us from one thought to the next. But the thoughts themselves were, perhaps, conscious. There need not

have been some further acts of picking and choosing from among possible names that occurred subconsciously, and which we would apprehend were we to shine a light on the subconscious.

The point I am after here is that the mental processing that songwriters and other artists often attribute to the subconscious may be associative in nature, and, *ipso facto*, not the sort of thing to be viewed as an inference. In that case, we should expect the relevant thought transitions to be inexplicable to the people having them—insofar as they are not subsumable under patterns of practical reasoning involving beliefs and desires—even if they were in fact entirely conscious. With this idea in mind, I will look in some detail at an actual associative process carried out in a recently developed artificial neural network. My claim will not be that it gives an accurate model of what in fact goes on during creative cognition—only that it *might*. Importantly, its mere possibility throws light on a whole range of strategies for explaining transformative creativity, and for relating such creativity to A-imagining.

12.10 Generative Adversarial Networks

Connectionist networks—otherwise known as "artificial neural networks" or "neural nets"—are schematic models of neural activity, representing sets of neurons connected to each other by varying degrees of strength. Different sets (or "layers") of "nodes" in a neural network serve as analogs for different sets of neurons, with input nodes standing for sensory-perceptual neurons, output notes corresponding to motor output neurons, and hidden layer nodes corresponding to all the relevant neural activation in between.[11] During a lengthy training process, during which the strength of the connections among the many different nodes in each layer are adjusted in response to its errors and successes, networks can be trained to accomplish complex tasks mirroring human perceptual discriminative capacities, such as face recognition, object identification, and phonology to text translation. Adjusting the strength of a connection within an artificial neural network is meant to serve as an analog to the strengthening and weakening of synaptic connections between neurons that occur during learning and development. As earlier remarked (Chapter 2), much of connectionism's appeal lies in the clear picture it provides of how specific cognitive abilities could be realized in a physical object with the structure and causal dynamics of the human brain.

While the general ideas behind connectionism have been widely discussed in philosophy and psychology since the 1980s, the last ten years have seen an exponential increase in the complexity of the tasks such networks are able to carry out.

---

[11] See Garson (2018) for a helpful primer on the basics of connectionist networks.

This is due in large part to substantial new investments by companies like Google, Facebook, Apple, and Amazon in developing the necessary theoretical and technological tools. As these models, and the algorithms that implement them, have grown in complexity, adding substantially more "hidden layers" of nodes between input and output, they have become known as "deep" neural networks. In 2020, Google Translate, Amazon's voice-to-text assistant Alexa, and Facebook's face-recognition software are some of the most visible applications of this technology.

In a "supervised" deep neural network, the network gains its discriminative abilities by being "trained" on a preexisting set of stimuli that have already been appropriately labelled. For instance, to create a network that can discriminate turtle images from images other animals, we might start with a large collection of photos of animals serving as input. The person training the network already knows which pictures are of turtles and which are not. The activation of each individual input node in the network, we can suppose, is determined by a single pixel's level of illumination in each image. Yet there will be just two output nodes corresponding to the answers: turtle/not-turtle. At the beginning of training, the weights among nodes in the hidden layers between inputs and outputs are randomly distributed. When the network's input nodes register the first picture, *some* of the connections among nodes in the hidden layer will have been "correct" in the sense that, due to their weighing, they favored activation of the output node corresponding to the correct response ("turtle" if indeed it was a picture of a turtle). The network's learning is then *supervised* in the sense that, after each new input is registered and a verdict given, the weights of the connections that favored the correct response are slightly strengthened, while those of the connections favoring the wrong response are decreased (using the same fixed "backpropagation" algorithm). Over the course of many, many trials and subsequent re-weightings, the network gradually becomes more reliable at activating the "turtle" output node when and only when the input image is of a turtle.

Importantly, this learning process does not involve giving the network a set of explicit rules and reason schemes by which to identify turtles. For instance, the network is never given a rule to follow such as: "If X has a round shell, and four legs, and a reptilian head, THEN X is a turtle." Its discriminations may well *conform* to such a rule, insofar as it only identifies things with round shells, four legs, and reptilian heads as turtles. The point, however, is that its training never involved its receiving such a rule but, rather, the gradual adjustment of the strength of connections between hundreds of nodes in hidden layers. Partly because of this, a well-trained network will be less "brittle" than a rule-governed classification system, with the system less likely to rule out a stimulus on the basis of its missing a particular salient trait. Moreover, programmers themselves do not need to know any such a rule in order to set up and supervise the network. Someone just needs to tell the trainer of the algorithm which pictures are of turtles and which are not, so they can adjust weights accordingly. (And, as we will see in a moment, even

this degree of "supervision" is not always necessary.) Because the network is not using any explicit rules or reason schemes in making its judgments, the sequence of states within the network leading up to its judgment are not *inferences* in our earlier-defined sense. The algorithm is not manipulating symbols in a stepwise, serial manner to compute anything akin to *modus ponens*, or the statistical syllogism. Its computations are instead distributed and parallel. The weightings of many different nodes simultaneously influence the network's output. And these weightings themselves are simply established through trial and error. If the network is trained on a sufficiently broad set of stimuli, including turtles of exotic appearance, it can come to classify turtles correctly without anyone's being able to articulate the rule that it is following to do so. Indeed, researchers typically have to conduct a separate experiment, with carefully constructed stimuli, to determine which features a network is in fact favoring when making its discriminations.

Now, when we think about creativity, we are interested in how creative products are *generated*, not in how they are discriminated from each other. But a neural network initially set up to discriminate a type of thing can, in effect, reverse its processing to generate plausible new instances of that type. These are called *generative neural networks*. Taking input from what would be considered the output node of a discriminative network, a generative network exploits the same set of weightings in the opposite direction to generate an instance of something likely to fall within the class of things the network was initially set up to discriminate. An ingenious method was recently developed for fine-tuning these generative networks, by pitting two networks—a generator and a discriminator—against each other (Goodfellow et al., 2014). Taken as a pair, these are known as *generative adversarial networks* (GANs) (Goodfellow et al., 2014; Radford, Metz, & Chintala, 2015). The generator has the role of producing outputs likely to fall within a certain class of things (turtle images, say), while the discriminator has the role of detecting whether the output of the generative network is indeed a member of the class of objects the discriminator has been trained to detect. When the discriminator is able to detect that the output of the generator does not fall within the class of things the discriminator was set up to detect (e.g., images of turtles), the hidden nodes within the generator are adjusted to decrease weightings that favored the "not a turtle" response in the discriminator. Likewise, whenever the discriminator is "fooled" by the generator, classifying something as a turtle image that was really just a creation of the generative network, the discrimator's weights are adjusted accordingly to deemphasize weightings that led it to give a false positive. This "double feedback loop" allows each network to continually fine tune itself against the other. The generator gets successively better at generating convincing members of the class to be discriminated, while the discriminator grows increasingly adept at discriminating real members from phonies.

In one striking deployment of GANs, researchers at Nvidia created a generative network capable of producing photorealistic images of (non-existent) celebrities

(Karras, Aila, Laine, & Lehtinen, 2017). The network was originally trained on thousands of photographs of celebrities gleaned from the internet. This was an instance of "unsupervised" training, however, in the sense that the images were not contrasted to another set of, say, non-celebrity images. Rather, taking only the celebrity images as input, the network was able to extract regularities within the set to calibrate itself for later discriminating celebrity images from images of other things (e.g., non-celebrities). The network's fine-grained discriminative capacity was subsequently used in reverse to generate images of (nonexistent) people who sort of look like celebrities—i.e., possible members of the set of things it was trained to discriminate. (It is an oversimplification to say the discriminative network is simply "run in reverse" when in generative mode; but the idealization is harmless for our purposes). This generative network was fine-tuned by being placed in a double feedback loop with a mirror-image discriminator network that received the same unsupervised training. After thousands of iterations, and resulting adjustments to each network, a generative network emerged capable of creating photorealistic images people who have a "celebrityish" look to them; human observers cannot reliably discriminate them from photos of actual celebrities they do not know (Metz & Collins, 2018).

Now to the philosophical relevance of GANs. As noted, my claim will not be that human creative cognition in fact relies upon processes akin to GANs—only that it *might*, and that the very possibility of its doing so opens up a range of promising strategies for answering the puzzles we have confronted about transformative creativity. With this in mind, suppose that we could ask a system deploying such a GAN how it came up with the image it generated of a possible celebrity. What would it say? There would be no set of deductive or inductive rules it could mention—not if it was speaking truthfully—and, thus, no sequence of states that constituted the following of such rules. The image was not arrived at by drawing inferences, or by following an explicit procedure whose inference schema could be stated in English. And yet we can understand its generation as resulting from a complex, distributed computation arrived at though an unsupervised learning process, during which it latched on to regularities it could not itself name.

Suppose now that the thought-transitions that occur during transformative creativity are rather like the GAN processing just described. We should not, in that case, be surprised to find that songwriters, novelists, painters, and the like are not able to fully explain why or how they arrived at their products, or to characterize their insights in inferential terms.[12] For while there are no doubt *some* processes involved in creativity that can be seen as inferential, our most finely

---

[12] Boden (2004, pp. 134–8) also discusses connectionism as a means to explaining the associative processes apparently at work in creative cognition, highlighting "the ability of neural networks to learn to associate (combine) patterns without being explicitly programmed in respect of *those* patterns" (p. 137).

tuned generative capacities may in fact be associative in nature. Artists receive "unsupervised training" through years of exposure to works they find aesthetically interesting and inspiring. They are, in effect, extracting regularities from the raw data to which they are exposed, in a way that cannot be linguistically expressed or formulated into a maxim. Eventually, they become sufficiently adept as critics to reverse their discriminative capacities in generating novel works of their own—possible members of the set of works on which they were trained. Being skilled discriminators of aesthetic value (and the features on which it depends), they are able to serve as their own adversarial networks, setting off alarms and changing their approach when their works do not meet the standards of those they admire.[13] Through an iterative process of revision and criticism, they eventually arrive at their goal. They are unable to distinguish their own work from a "fake"—i.e., from something lacking in the kind aesthetic features they have trained themselves to discriminate. Yet their abilities will in a sense remain a mystery to themselves, as there will never be clear *reasons* they can provide for how or why they made what they did—for the process was not inferential. They may as well chalk it up to "the subconscious."

It may, however, seem that the analogy to GANs leaves no room for genuine originality to emerge—at least, not orginality of the sort we expect of *transformative* creativity, which was our explanandum. The GAN that generates celebrity*ish* photos is, after all, highly imitative in nature. It succeeds precisely by making images indistinguishable from those in a preexisting set. How would something like a GAN create room for a new *style* to emerge? How do we get from fine-grained imitation to creative generation?

Two points in response. First, most art—even great art—*is* highly imitative in nature. Consider Dylan's debt to Woody Guthrie, or Amy Winehouse's to Billie Holliday. An artist's work is understood, in the first instance, in relation to her influences. Second, and more important, we have to consider the diversity within each person's "training set." A GAN trained only on images of celebrities is like a poet who only ever read the work of William Carlos Williams. Such a poet would likely write poems too highly derivative of Williams to be considered creative in any transformative sense. But now add to her training set thousands of poems from the middle ages to present day—an idiosyncratic cluster weighted by her tastes and interests. In understanding and enjoying the works within that set, she is implicitly tuning her generative capacities to a singular frequency—one that distills her individual training and that emerges with an identity of its own. The same capacities—the same network—may one day be exploited in the generation of new poems in an original, transformative style.

---

[13]   Here again is Paul Simon: "So I let the songs go this way and that way and this way and whatever way it is and basically what I do is be the *editor*: 'Oh, that's interesting. Never mind *that*, that's not so interesting. That's good, *that's a good line*'" (Zollo, 1997, p. 95).

Where could *imagination* fit in to all this? Because GANs make use of connectionist "deep" neural networks that rely upon parallel distributed processing, it is not proper to describe them as exploiting language-like representations in a stepwise, serial fashion. This means we cannot accurately describe them in heavy-duty folk psychological terms (see Chapter 2); there is no fit to be made between notion of a GAN and, say, a Belief Box or Imagination Box. Yet we can still give light-duty folk psychological descriptions of the GAN processing, insofar as the processing gives rise to certain dispositions in the system that can be described and tracked in folk psychological terms. Supposing that the celebrity-photo-generating GAN were embedded within a system we were otherwise comfortable describing in (light-duty) folk psychological terms, it seems reasonable to describe its generative processing as "imagining possible celebrities," insofar as it results in images of celebrity-like faces. Likewise, the poet who was just described as using GAN-like processing can be seen as imagining possible poems and possible next lines. Yet we could alternatively, and equally accurately, describe the celebrity photo GAN-system as *judging* that certain images look like celebrities, and the poet as *judging* that thus and such would be a good poem, or as *deciding* that thus and such would be a good next line. In short, the dispositions we ascribe—invoking imagining in one case, and judging and deciding in the other—are essentially the same. It is, then, possible to explain this sort of (transformative) imagining in more basic (light-duty) folk psychological terms.

12.11 The Importance of Being Earnest

The account I have sketched of transformative creativity, and of the imaginings it involves, aims to show how creativity can rely upon years of training, even if experts are unable to articulate the principles behind their work, or to generate new works by following an explicit procedure. We can see why trying *intentionally* to come up with a creative new song might backfire—why Neil Young is right to stop when he senses himself trying to think up the next line of a song, or why Paul Simon tries not to think about what a song will be about. For in trying to think our way to an answer, we will be tempted to reason our way there through a set of inferences. ("If I have lyrics x, y, and z, then I ought to follow them with…what??") This is not the way works are created on the model of a GAN. Trying to arrive at the next line as an inference will leave us spinning our wheels. A better analogy is that we must learn to play our minds like an instrument, letting the strings ring out in their natural tuning.

But this last analogy—and indeed the entire appeal to GANs above—may seem to paint too simple a picture. To generate great art, could it really be that we just need to reverse our discriminative capacities and let 'er rip? No. We have to recall Leonard Cohen's dismal assessment that, when it comes to eliciting the state of

mind needed for writing a good song, "nothing works." This much is evident from the fact that many a refined critic, with carefully tuned discriminative capacities, cannot *create* compelling work. What are they missing?

We first must take account of the basic teachable skills necessary to any particular field. It doesn't matter how refined your taste in piano concertos may be. If you haven't learned to play the piano, and don't grasp the basics of music theory, you won't be able to write a concerto yourself. The same goes for painting, fiction-writing, philosophizing, songwriting, and any other creative endeavor: there are essential skills and techniques—the sort of things imparted in MFA and PhD programs—whose mastery requires practice. These are necessary to convert one's discriminative capacities into productive ones. The interesting question is what *more* transformative creativity still requires. To move closer to an answer—to really understand the bait on Bob Dylan's line[14]—we need to pan out from neural networks and teachable skills to consider matters of character and motivation. Here Matthew Kieran's writings on creativity are particularly illuminating, in ways I'll describe in the next section. We will not arrive at a guide for how we ourselves can take our place among the great artists, of course. But, with luck, we will gain a better understanding of their process.

12.12 Character, Creativity, and Conscious Dreams

Matthew Kieran (2014) analyzes creativity as a virtue of character, prioritizing traits like courage, patience, and honesty over intellectual capacity. Whether one is working in art, philosophy, or science, he notes:

> It takes honesty to evaluate the nature and value of what one is doing properly; it takes courage to be prepared to fail; it takes humility and open-mindedness to recognize when one has gone wrong; and it takes perseverance and fortitude to continue to work at something for its own sake or in seeking to do justice to an idea.   (2014, pp. 132–3)

Kieren's claim isn't that you need to be a *saint* to generate transformative creative works. It's rather that there are specific virtues—diligence, open-mindedness, courage, fortitude—that tend to promote one's ability in that direction. Such character

---

[14] From Zollo's (1997) interview with Dylan:

ZOLLO: "Arlo Guthrie recently said 'Songwriting is like fishing in a stream; you put in your line and hope you catch something. And I don't think anyone downstream from Bob Dylan ever caught anything.'" [Much laughter].

ZOLLO: Any idea how you've been able to catch so many?"

BOB DYLAN: "It's probably the bait."

virtues are as much a requirement for transformative creativity as are training in the basic skills of an art, or the development of an ability to tell the aesthetically good from the bad. They affect one's ability to gain the required skills in the first place, of course; but, equally important, they raise the odds that one will overcome the drives that "tend to pull us toward compromise, self-deception, and over-inflated estimations of the nature and value of what we do" (p. 133).

Kieran rightly draws links between these virtues and the having of "intrinsic motivations," where these are motivations grounded in "values internal to the relevant domain." Compared to the person with extrinsic motivations, someone with intrinsic motivations will be "more sensitive to and motivated by reasons bound up with the goods internal to the activity in question" (p. 129). For instance, if one's motivation is intrinsic,

> a subject is more likely to take risks, more likely to attend in an open-minded way to what she's done, envisage different possibilities, and be directed by thought in action toward realizing the inherent values in a given domain.   (p. 129)

By contrast, when the motivation is extrinsic, we'll be inclined to take shortcuts and apply formulaic strategies, using the product as means to some other end. Work done in the service of extrinsic goals, Kieren adds, "often only coincidentally tracks and typically pulls away from the intrinsic values of the given domain" (p. 131). "The extrinsically motivated agent is keying into certain social dynamics and goods," he adds, "where the aesthetic quality and value is coincidental to such" (p. 130).

There are two kinds of social courage Kieran identifies that facilitate transformative creativity. First, there is a kind of courage-through-modesty required in honestly assessing one's own efforts and abilities; second, and perhaps more importantly, there is the courage required to maintain an intrinsic motivation—for this requires dismissing, as irrelevant, the approval of others; and, more troublingly, it may expose oneself to ridicule. I am reminded again of a comment from Zollo's interview with Paul Simon, where Simon describes the self-imposed barriers that lead to "writer's block":

> When I have writer's block (though I don't think of it as writer's block anymore), what it is is that you have something to say but you don't want to say it. So your mind says, 'I have nothing to say. I've just nothing more to say. I can't write anything. I have no thoughts.' Closer to the truth is that you have a thought that you really would prefer not to have. And you're not going to say that thought. Your mind is protected. Once you discover what that thought is, if you can find another way of approaching it that isn't negative to you, then you can deal with that subject matter.   (Zollo, 1997, p. 98)

Simon is pointing to a kind of courage—in front of oneself and others—needed to overcome the natural self-censors that shield us from our more unpleasant, offensive, vulnerable, or embarrassing thoughts. The word "courage" doesn't quite capture this virtue, as it seems equal parts humility and patience. But we at least have an inkling for why the virtue is not easily acquired—and why "freely" generating thoughts worth recording is not as easy as it might seem. Each of us has an armada of self-protective rationalizations aimed at stopping it.

Earlier, in connection to GANs, I spoke, metaphorically, of transformative creativity resulting from learning to play one's own mind like an instrument, letting its strings ring out in their natural tuning. To be more precise about what that might involve, the idea is this: self-conscious, self-protective, reason-concerned thought may perhaps interfere with the production of the kinds of new thoughts and connections that would result from running our discriminative capacities in reverse (as on the GAN model). For the latter sort of associative thoughts will reflect the actual shape of our minds—as weathered by experience—as opposed to what we might wish that shape might be. The artist's mantra of "try not to try" can be heard as a directive not to interfere, through self-editing and self-censorship, with such associative thought production. This is much easier said than done, of course, because obeying the self-censoring, self-rationalizing inner voice is so fundamental to being a socially engaged human being. It is not only a matter of skill, but also of character that one is able to obey the artist's directive to ignore it. It takes a certain kind of courage, perhaps to the point of recklessness, to lower your defenses in order to, in effect, dream while awake. The comparison to dreaming is intended literally: for in dreams, too, our self-supervisory, reason-giving systems fall away to let our more interesting fears, wishes, desires, and quirks reveal themselves. Dreams are indeed evidence that we all *have* the basic ability to generate rich, creative scenarios without *feeling like* we are doing so. The person skilled in transformative creativity has gained the special ability to summon, shape, and explore those states while awake. That skill is inseparable from his or her character.

We can, then, explain a person's creativity *in part* by appeal to their interest in the intrinsic features of the form itself, together with their more general character virtues: their patience in developing and revising their work, their humility in accepting critique, their courage to risk looking like a fool, their confidence in their own abilities. If we want to understand why one person with a highly developed aesthetic sense and finely tuned discriminative skills shows transformative creativity, while another does not, their respective characters and motivations are good places to look. These traits will feed back into the development of their aesthetic skills and discriminative abilities: the deeper a person's concern for the intrinsic values of a discipline, and the greater their diligence, the more time and effort will be expended in sharpening their aesthetic sense and productive skills.

Through it all, a certain foolhardy courageousness will be essential to quieting the self-conscious, self-rationalizing mind, so as to let its proper tuning ring out.

12.13 Concluding Thoughts

There are many things we might wish to explain about creativity. I have tried to focus on those most pressing from the perspective of contemporary philosophy of mind, with an eye toward shedding some light on the nature of the "imaginings" required for creative cognition. I drew distinctions among kinds of creativity—combinatorial, explorative, and transformative—in order to highlight a more basic distinction between *inferential* and *non-inferential* (associative) thought processes. While I suggested that transformative creativity most obviously requires the latter, that link is itself a bit artificial. Some acts of transformative creativity may well result from complex inferential processes, while some explorative and combinatorial creativity may draw upon non-inferential associative processes. Ultimately, the distinction between an inferential and a non-inferential thought process cuts deeper than the heuristic distinctions I have made among "kinds" of creativity. It also cuts differently (if not deeper) than the distinction between imagining and judging. Just as imaginings can be seen as in some sense inferential in nature (as it is by those who see it as a kind of "offline" inference (e.g., Nichols & Stich, 2000; Currie & Ravenscroft, 2002; Williams, 2017), judgments can be seen as associative—as they were by early associationists, such as Hume, and as they are within theories of judgment modelled on contemporary neural networks. So if the reflections of songwriters and other artists convince us that not all creative cognition can be inferential in nature, they don't yet tell us whether such cognition requires *sui generis* imaginative states, or judgments instead.

In any case, the fact that artists lack any sense of control over their most important generative cognitions suggests that such states are not well analyzed as the kind of (conscious, controllable) cognitions that *sui generis* imaginings are normally thought to be. Instead I offered recently developed generative adversarial neural networks as possible models for how we should think about creative cognition that is not inferential in nature. Such computational structures—like all artificial neural networks—have the advantage of being modelled on the structure of the brain, and of showing how both pattern-recognition *and* the generation of novel products that fit a certain pattern can occur in an "unsupervised" manner, without need for specific inferential principles to be programmed in at the start. While we do not yet know the extent to which human creativity draws on similar processes, there are promising and rapidly developing research programs in place.

Finally, there are questions about creativity that aren't well answered by appeal to a cognitive architecture shared (presumably) by all humans. After all, the ability for transformative creativity is not evenly dispersed. This is why I concluded with some thoughts on character and motivation, which contribute essentially to the development and refinement of our basic capacities for creative cognition. There is a sense in which we all have the cognitive tools needed for transformative creativity; whether we are able to cultivate, refine, and exploit them is a matter of character.