Creativity as a Higher Agency

Forthcoming in Philosophy and Phenomenological Research

Kenneth Walden

Abstract. Can human agency produce things that are genuinely creative and original? Some philosophers are skeptical. Here I argue that the case of creative activity should lead us to reexamine and ultimately expand our conception of agency. When we do this, we see that rather than being incompatible with agency, creativity offers an especially robust form of agency: a form in which agents are responsible not just for token events but for the general patterns that characterize those events as forms of human activity.

1.

Many believe that creativity is not something I do but something that befalls me. Not a task I undertake in the light of self-aware deliberation, but something visited upon me in the shadows of the mind by muses and possessing angels. Plato seems to be a champion of this view. "A poet is an airy thing, winged and holy," he says, "and he is not able to make poetry until he becomes inspired and goes out of his mind and his intellect is no longer in him." No mortal poets can claim responsibility for their poetry, he continues, because "god himself is the one who speaks, and he gives voice through them to us."[1]

If Plato is right about this, then we face a real puzzle, since, as many others have pointed out, creativity is not something that happens, but something that a person does. "We may attribute beauty or other aesthetic properties, but we do not (properly) attribute creativity to an unusual array of cracks in a rock wall or to the image of a mythical creature in the clouds," Dustin Stokes observes. "If, however, we come upon an abandoned artefact of some sort, say a painting, we might attribute all of the same properties plus creativity." This is because "we see the cracks and clouds as lacking in any marks of agency; while the

[1] *Ion* 534b-c in (Cooper 1997).

painting betrays the fact that it depends upon agency. We are willing to call the second, but not the first, creative."[2] Thus an event does not so much deserve the name "creativity" if it is not somehow dependent on the agency of a creator.

For all that, there are obvious reasons to agree with Plato. We associate agency with certain powers: setting and pursuing ends, forming and following plans, adhering to rules, and weighing reasons. But it is hard to see how these abilities could produce anything truly original—anything more than the next iteration of a tried-and-true formula. Creativity is too spontaneous, too indeterminate, too much the result of inspiration, we might insist, to be extruded through these stiff holes. If I told you that I had a comprehensive plan to become a creative genius in just ninety days, that I had discovered a set of rules that guaranteed to make me a visionary, or that my new work of surrealist fiction was the result of a meticulous balancing of aesthetic reasons, you would call me a dope.

This kind of argument requires elaboration. We might wonder, for example, why there is this alleged tension between these tools of agency and the demands of creativity. Let's take planning as a representative of the class. The purpose of plans is to allow the agent to interact with the world systematically. After I am done at the office, I plan to drop off a book at the library before proceeding to the grocery store for a pint of ice cream. I make this plan because I would prefer the ice cream be out of the freezer for as little time as possible, and I believe my stop at the library will be brief enough to enable me to make it to the grocery store before it closes. In making this plan I am orienting my behavior and goals within a network of patterns that make up my little corner of the world. Patterns about how long it takes one to return a library book, about how long it takes to go from library to grocery store, about the hours of the grocery store, about my walking speed on a drizzly afternoon, about my preferences for different states of ice cream, about the thermodynamic properties of that ice cream. The plan's function, very broadly speaking, is to orient me within these patterns so I can interact with them comprehensively and intelligently, so I can take on the various problems my ends present to me systematically rather than one by one. It's because of this orientation that I am more likely to achieve my ends if I plan than if I wing it.[3]

We can think of planning as a kind of technology, as a mental tool that has allowed us to expand our capacity for making a difference in the world. It is one of several technologies made possible by our self-consciousness. We can *reflect* on ourselves and so

---

[2] (Stokes 2011, 659–60). See also (Gaut 2010, 1041), (Kieran 2014), and (Paul and Stokes 2018).
[3] (Bratman 1987, chap. 1).

come to represent ourselves as an integrated part of a larger world, and we can regulate our actions in sophisticated ways in virtue of that representation. The capacity for this kind of reflection represents one of the most important joints in nature. "Everything in nature works in accordance with laws," Kant says, but only certain creatures have "the capacity to act in accordance with the representation of laws."[4] Marx observes, even though "a spider conducts operations that resemble those of a weaver, and a bee puts to shame many an architect in the construction of her cells … what distinguishes the worst architect from the best of bees is this, that the architect raises his structure in imagination before he erects it in reality."[5] This power of reflection gives us a tremendous advantage in the pursuit of our ends. Because of it we needn't consign that pursuit to blind instinct and pray we can cope with whatever comes our way. We can design and perfect that pursuit by imagining different prospective courses of behavior unfolding across the wide network of patterns that make up the world, predicting which of these offers the best path to our goal, and steering the ship of the self accordingly. Reflection makes a more systematic way of interacting with the world possible.

Reflection is also the origin of the apparent conflict between creativity and agency. For whatever else creativity involves, it involves novelty. To say that something is creative is to say that it is new. (New for us or new for its creator or produced in a new way or demanding of a new kind of understanding or all of these things.) But insofar as the technologies of reflective agency that I listed depend for their efficacy on orienting our prospective behavior within *existing* patterns—within the patterns that make up the world as it presently is—they will be poorly suited to bringing about things that deviate from those patterns and so poorly suited to bringing about anything of more than token originality.[6]

I am hatching a plan to write a poem. If this is to be an effective plan, I must be able to rely on certain connections between my poetic acts and artistic ends. The best way to ensure that reliability is to look at what the world already contains—at existing examples of poetical production and the causal relationships they depend upon. And if my plan is to be a good one, it should allow me to insert myself into these patterns by, for example, exploiting the connection between a certain kind of poetical activity (writing in iambic pentameter) and a certain kind of result (an impression of formal refinement). But this is

[4] (Kant 2002, 4:412).
[5] (Marx 2000, 493).
[6] Compare (Tomas 1958) and (Peacocke Forthcoming), the latter of whom understands the problem as one about the possibility of creative action being regulated by intention.

just a long-winded way of saying that a good plan for writing poetry will be a plan to imitate other poets.

If I followed this kind of plan in hopes of crafting an original poem, it would prove either useless or self-defeating. On the one hand, the plan may be very vague. Maybe it says to write original poetry by finding inspiration on a tour of Italy. But that's not much of a plan at all. Even if novel poetry does follow close on the heels of my *Italian Journey*, we would not say the plan produced that poetry because the connection between the two is too feeble. At best, the plan put me in a favorable position to exercise some independent faculty of creativity. On the other hand, we can imagine a plan that is effective—a plan where the connection between each component and a particular poetic output is reliable. We might plan thus: first, write a long dash and then the letter "A" and then a space and then the word "simple" and then the word "child" and then insert a line break and then write "That lightly draws its breath", and then, well you get the idea. This is an effective plan for writing a poem because it reliably connects the desired outcome to the planned behavior. Unfortunately, it is a plan for writing Wordsworth's poetry, and it is effective precisely because it is a form of imitation. So it goes with plans generally.

Now, this does not mean that the presence of a plan completely excludes the possibility of creativity. My suggestion is that the point of planning, and thus what makes something a plan good or bad, depends on a particular regularity in the world, and that makes planning ill-suited to producing irregular things. It is true that artists and inventors can plan out large-scale projects that turn out to be creative. But it is not quite right to say these people are planning their creativity. Instead, they are erecting a kind of scaffolding, a kind of support for creative endeavors that makes the overall project more likely to succeed. This scaffolding is an auxiliary to creativity, not its source.

We can also, of course, be creative in the making of our plans. Napoleon's battle plans, Henry Ford's factory plans, and Peter Scott's burglary plans were all creative. But this creativity is decidedly limited by their status as plans. Napoleon's battle plans consisted of a collection of standard sorts of subplans: marching troops over there, arranging cannon in that direction, surrounding this copse or that thicket. If they weren't composed of these familiar maneuvers no one could follow them. The plans' originality lies in the assembly of these parts, in their higher-level structure. And even this limited originality would be a burden on the plan if not for a very special feature of war: that it's useful to catch the enemy by surprise. Plans can be creative, and sometimes we will have

special reasons to seek a creative plan, but generally, the more creative a plan, the less reliably it orients us to the world, and so the worse it is at doing its job *as a plan*.

Attempts to fit creativity within the standard framework of planning agency often stumble on this fact. Some say that creativity involves guidance by a plan or intention, just like other forms of action, but that this plan or intention is somehow indeterminate. This, I believe, is Antonia Peacocke's strategy.

> A creator first has in mind an early proto-work—i.e. a draft or component of the final idea to be conceived. This proto-work might be conceived fully inactively. But once there, it sets up its own standards for success. The creator might grasp these standards only via their access to the particular that is the proto-work. But once grasped via the proto-work, these standards can figure into a creative intention that also enfolds a form of value: the creator can intend to give that what it needs, or do justice to that, where "that" refers to the proto-work itself.[7]

As a description of the creative process, I find Peacocke's account congenial. But I don't think it solves our problem. To the extent that Peacocke's "proto-work" can determine how a creative endeavor will unfold—to the extent that the work created has certain definite characteristics because of the proto-work—that proto-work functions just like a plan. And to that extent, the endeavor is not, after all, creative. Meanwhile, to the extent that the proto-work is not determinative in this way, we are left with little sense of what actually does this work, of what makes the product the way it is. This is the gap in understanding from which fantasies about muses emerge.

The more indeterminate a plan is, the less useful it is as a plan, and so the less work it does in explaining how a given act is brought off. We can and do adopt indeterminate plans, schematic plans, and skeletal plans when something more refined would be a mistake. But the more a plan assumes these qualities, the more we should expect to look beyond that plan, to other aspects of their moral psychology, to understand how a person does this something. This "beyond" is where we should expect to discover whatever it is that is distinctive of creativity. Trying to explain creativity by pointing to the indeterminacy of one's plans is like trying to explain a puddle by pointing to a hole in the roof. True enough, that's part of the story. But just as there must be rain to pass through the hole and

[7] (Peacocke Forthcoming). See also (Gaut 2009).

create the puddle, so must there be something that passes through the holes left by our indeterminate plans.

If we generalize from the case of plans, we get the following hypothesis. Agential technologies made possible by our capacity for reflection are poorly suited to creativity—though not strictly incompatible with it—because the characteristic function of those technologies is to orient our prospective behavior to existing patterns in the world: to the fact that the library closes at five o'clock, that it takes me ten minutes to walk eight blocks, that ice cream has a specific heat capacity of 2100 J/kgC. These technologies are useful precisely because they put us in reliable contact with the causal levers we must manipulate to produce our ends. But they are poorly suited to creativity for the same reason. Creativity involves substantial deviation from these patterns, and that is something plans are not designed for.

2.

This hypothesis explains our Platonic suspicion that agency and creativity are not a happy couple. But it is only a weak vindication of that suspicion. What I have argued is that the clash between creativity and agency only arises for certain "technologies" of agency, not for agency as such. It is very tempting to identify agency with these technologies and suppose that agency *just is* the ability to form intentions, make plans, and follow policies since so much of what we do is done through these devices. Many philosophers do this. But there is a more liberal and arguably more natural conception of agency that does not make this assumption. It is the conception that Maria Alvarez and John Hyman suggest when they note that, "our pre-theoretical talk of agency extends to animals and plants, and also to inanimate things."

> We say, for example, that beavers build dams, or that oxygen rusts iron. Indeed, inanimate substances are often described as agents of one sort or another: cleaning agents, analgesic agents, etc. The pre-theoretical view of what it is to be an agent that this sort of talk implies acknowledges that inanimate substances and living creatures incapable of acting voluntarily or intentionally can be agents.[8]

[8] (Alvarez and Hyman 1998, 221). Compare (Frankfurt 1988).

This kind of talk suggests a more basic and wider conception of agency. Because it is a form of agency available not only to reflective creatures, but to beavers and bears, it is sometimes called "animal agency".[9] But I think it is a mistake to associate it too strongly with non-human animals, so I will call it the *liberal* conception of agency.

The heart of the liberal conception of agency is difference-making. It understands agency as an ability to be why things came to be as they are—to be their "ground" if you like that way of talking. The analgesic is an agent because it is why the patient fell asleep. The beaver is an agent because it is why the river slackened. I am an agent because I am why there are three hundred popsicles on the roof. To be an agent in the liberal sense, then, is to be the answer to a question about how something came to be. If we think that these answers will invariably invoke causal powers, then we have something like Aristotle's theory of agency: my finger's being wiggled counts as an exercise of my agency—as my wiggling my finger—just insofar as it is the exercise of a causal power that belongs to me.[10]

This conception binds agency very tightly to explanation, and this connection introduces an important complication. A phenomenon is seldom explained by one or two lonely propositions, but by vast "corporate bodies" of theory and circumstance.[11] We ask, "why did the river's flow stop?" and are answered, "because a beaver dammed it." In most instances this pithy answer is satisfactory, but it is incomplete. The river's slack is not wholly explained by a single fact or a single beaver. It is explained by a whole network of facts about beavers, logs, the topography of the forest, the gravitational constant, Bernoulli's principle, the reproductive cycle of local insects, the phases of the moon. When we explain an event, we are not just offering its proximate cause. We are showing how it fits into the larger picture of the world suggested by a systematic theory.

This feature of explanation has a tricky consequence for our conception of agency. I said that the beaver's agency, at least at one moment, rested with the fact it was the reason why the flow had slackened, but this claim seems to shortchange other parts of our vast explanation. Why just the beaver? Why not the termites who weakened the tree? Or the coquettish squirrel who lured the beaver to the tree? Or the tectonic plate that smashed into its neighbor a thousand years ago to produce the mountain from whence the stream flows?

I don't think these questions have good answers, and instead of trying to silence them, we should abandon the notion that the beaver is the exclusive explanation of the

[9] (Steward 2012).
[10] (Aristotle 1983, III.i). See also (Coope 2007).
[11] (Quine 1953, 41), (Duhem 1954).

river's slackening. Fortunately, there is an obvious amendment to the liberal conception of agency that accommodates this change. We should accept that agency comes in degrees. Even if we cannot say that this thing is *the* reason why a state of affairs came to be, we can say that it made a larger or smaller contribution to it. This may then allow us to assess degrees of agency—to say *how much* of an agent an individual is at a given moment with respect to a particular feature of the world.[12]

It is relatively easy to introduce this idea into our usual ways of talking about agency. Imagine Elroy pruning his hedge. In one moment, he snips off exactly the branch he planned to, but in the next a great gust of wind blows him over and he removes the wrong branch. Elroy will certainly play a role in the explanation of both branch snippings, but there is an obvious sense in which the part he plays in the explanation of the first is greater. For in the second snipping, he must share the explanatory stage with the gust of wind, since it is part of the explanation for why this branch, and not the other, was snipped. This means Elroy's contribution to the difference-making will be proportionally smaller, and so, consequently, will his agency.

Here one might object that according to this standard, all the agency in the world belongs to some primeval first cause, since the ultimate explanation for Elroy's pruning, gust of wind or no, will pass through a long series of causal antecedents stretching back to the beginning of the universe. But this worry depends on a bad view of explanation. Our best explanation of Elroy's pruning will not consist of facts about the initial arrangement of the universe plus a set of physical laws that somehow guarantee a bush pruning occurs at a certain time. That kind of explanation ignores all the higher-order patterns—the regularities sought by the meteorologist, the biologist, and the sociologist—that make the world legible as more than atoms tumbling in the void. Our best explanation of many events, especially human events, will depend on these patterns and the emergent properties they trade in—on the properties of persons, bushes, and wind gusts. It is at this level where a contrast between a difference made by a person and one made by the wind will emerge.

If we follow the example of Elroy, we get the following standard. Ascriptions of liberal agency will reflect a ratio in our explanations between influences that *belong* to a given individual and those that arise from elsewhere—from a gust of wind or a rapacious termite or a tectonic plate. The more an explanation ultimately turns to the individual and

[12] On the idea of agency admitting of degrees and the cases that might leads us to brook such a possibility, see (Mele 2008, 271–74), (Coates and Swenson 2013), (Nelkin 2016), (Timpe 2016), and (Timpe 2022). For a conception of this gradient closer to the one I am working with see (Korsgaard 2009, 174) and (Kierkegaard 1983, 30).

their powers, the more the event is attributable to that individual. The more it is an expression of their agency. To be sure, this is a rough standard—one that relies on mushy notions about how "central" a thing is to an explanation and what factors "belong" to an agent—but that does not make it uninformative.

3.

What can we say about the gradient made up of different degrees of agency on the liberal conception? For one, it appears quite long: there are many degrees of agency. It will be helpful to have a brief tour of this slope, so let's start near the bottom. The principal way to enhance your agency is to integrate difference-making powers within your self—to make them your own. Organisms are the classic example. Organisms have greater agency than most non-organisms because they are organized in such a way that the causal powers of their parts belong to the whole. A bacterium is made of various constituent parts, and these parts have powers like locomotion, metabolism, and protein production. But because these causal powers complement each other and ensure the overall well-functioning of the bacterium, we should also attribute them to the bacterium as a whole. We can say that *E. coli*, not just the flagellum affixed to it, is capable of locomotion. In effect, organismic form binds the causal powers of the organism's components and locates them in the agent that emerges in that form.[13]

Another enhancement to agency accompanies the introduction of self-propulsion. The structure of simple organisms like plants is maintained through the pre-established harmony of their component parts. The flower's roots and petals "work together" in the sense that they are programmed to perform complementary functions. But animals possess something in addition to this natural harmony. They possess what Aristotle calls an "inborn spirit", an executive capacity that enables one part of their body to exercise control over others.[14] This capacity allows animals to play a more active role in maintaining their well-functioning. The animal can initiate motions and resist external motions in ways that the more passive organism cannot: beavers can shoo flies that flowers must grin and bear.

These two base levels of the *scala naturae*—simple organisms and animals—correspond to two levels in what we might call a *scala agentia.*[15] But just as there are higher

[13] (Goldstein 1995) and (Merleau-Ponty 1963).
[14] (Aristotle 1986, 703a10).
[15] Compare the image in (Thompson 2008).

rungs on the *scala naturae*, rungs beyond the beaver, so there are there higher forms agency. Many of these forms of agency come about from the seminal technology we have already considered, from self-conscious reflection. Reflection makes higher forms of agency possible by enabling agents to, in effect, inflate their self—to constitute themselves as an entity occupying a relatively greater share of the causal order.

For example, these technologies allow us to spread our agency across time. I am, naturally enough, the same agent now as I was two years ago. The causal powers I exercised then belong to the same being exercising similar powers now. But what makes this fact true? It is not, I want to insist, a brute metaphysical fact. It is an accomplishment. I have behaved in ways that make it appropriate to explain past-me and present-me as moments of a selfsame agent, and this is part of *why* I am the selfsame agent.

When I say that this is an accomplishment, I mean that it is conceivable that I could have failed to enjoy this kind of identity. Had my behavior and attitudes become radically different over this interval, then it may not be appropriate to credit me with this continuity. This would diminish the degree to which we can say that the difference-making powers exercised by past-me belong to present-me, and that, in turn, would diminish my diachronic agency. Plans are especially useful for securing this kind of diachronic continuity. If I have a history of plan-making and conscientious plan-following, and one day I make a plan and subsequently carry out each of its steps, then the leading explanation for this course of events will posit a single self, a single me, that exists across different times and is responsible for both the adoption and implementation of the plan. Planning helps me behave in ways that constitute my behavior as the behavior of a single diachronically unified being. It earns me a form of diachronic agency.[16]

This is what I mean when I say that our reflective powers help us constitute ourselves as beings that occupy a greater proportion of the causal order of the world. Planning enables me to behave in such a way that I am thereby constituted as a diachronic agent. It enables behave in such a way that a temporally extended portion of the causal order of the world *is me*.

This is not the end of the agency enhancements made possible by reflection, though. (Small wonder that so many philosophers have been tempted to identify agency with this power.) Our reflective capacities also underwrite certain forms of self-regulation that have a buttressing effect on the self. I can reflect on my impulse to throw a pie at the

[16] (Bratman 2007) and (Ferrero 2009).

speaker and reject it as a momentary hysteria. I can do the same for my burning desire to listen to Cyndi Lauper, only this time endorsing it as an expression of my aesthetic identity. I can scrutinize my suspicion that the man with the twirly mustache is a spy by asking whether it's based on obsolete stereotypes. When repeated, these acts of scrutiny, rejection, and endorsement help me create and sustain a stable evaluative point of view on the world. And this, in turn, can secure my constitution as a unified agent, a single entity to whom diverse kinds of activity are properly attributed.[17]

4.

The previous section charted a *scala agentia*, a sequence of capacities that can enhance the agency (in the liberal sense) of the individuals possessing them. Where does this leave us with Plato's skepticism about creativity and agency? In a slightly awkward position, it seems. The best argument for Plato's suspicion rested on the conflation of agency in general with the kind of agency made possible by our reflective capacities. This is a mistake because there is a more liberal conception of agency that does not depend on these reflective capacities. I have been trying to spell out this conception in hopes of showing how much of our folk theory of agency can be reinterpreted as claims about greater and lesser degrees of this more liberal brand of agency. So in one sense Plato's suspicion is unfounded. But in another sense it has been vindicated. If the "higher" forms of agency, the forms most dear to us and distinctive of our status as human beings, really do depend on our powers of reflection, then it seems very important that creativity is incompatible with those powers. Who cares if creativity is compatible with some very liberal conception of agency that we associate with beavers and cold medicine if it is indeed at odds with the distinctively *human* powers of agency?

No one should care. Or rather, they shouldn't care unless creativity itself makes possible *another* kind of "higher" agency. Not unless creativity offered its own, unique path up *scala agentia*. But that, I shall now argue, is exactly the case.

5.

Our reflective capacities give us access to importantly "higher" forms of liberal agency that outstrip what is available to other organisms. But so, I now want to insist, does creativity. I have two arguments for this thesis, the first of which occupies this section.

[17] (Korsgaard 2009, 18ff).

"Men make their own history," Marx says, "but they do not make it as they please; they do not make it under self-selected circumstances, but under circumstances existing already, given and transmitted from the past." These circumstances, we might suppose, are the material conditions we face when acting—the fact that we have these tools and face these hurdles. But this is not the whole story. As Marx continues: "the tradition of all dead generations weighs like a nightmare on the brains of the living."[18] From this metaphor it is clear that Marx is also talking about our intellectual inheritance, about the discrete, historically conditioned forms our thinking can take.

In saying all this, Marx is on to an important fact. Our mental attitudes, including most of those that guide our actions—our plans, policies, and so on—are *artifacts*, sometimes quite intricate ones. Think of the cunning plan that guides the cat burglars, the carefully cultivated equanimity of the country priest, the web of policies that governs the dutiful bureaucrat, the folksy persona invented by the ambitious politician, the verbal and sartorial preferences of the punk, the complex narratives that we construct to motivate us in our long-term endeavors. These items are all the result of a kind of craftsmanship. And *how* these things are crafted—or more precisely *by whom*—matters to questions of agency in a straightforward way. If I create one of these intellectual artifacts myself, then it is more fully mine than if I take it from somewhere else. (Mill calls this having a *character*.[19]) If a mental artifact is my own creation in this sense, then the activities based on it, the things I do because of it, are more fully attributable to me than things based on an attitude that is not mine. Other things being equal, this little bit of creativity embiggens my agency.

Consider the following examples. Felicia undertakes an extensive quest in the role-playing game *Mazes & Monsters*. She is playing with her level twenty-eighty wizard, a character whose basic features she designed months ago and whose evolution she has guided through innumerable quests. She and her party have devised an intricate plan for today's adventures. They have discussed how to avoid getting separated, how to use magic and range weapons to their best effect, where they should try to steer the battle, and what snares to beware of. Felicia and her party play through the quest and win when Felicia casts ice thunderbolt on Cephalopokus the Land Squid.

Now, suppose everything is as I described it in the previous paragraph, except Felicia is called away to perform brain surgery and can't play. Her friends insist on playing anyway and ask Felicia's little brother Gerald to step in and be her wizard. As it happens

[18] (Marx 2000, 329).
[19] (Mill 2003, 125)

Gerald is an adroit *Mazes & Monsters* player himself, and he has observed Felicia playing this character many times before. He also thinks it's an excellent character: flexible, resourceful, with an interesting backstory. While he wouldn't have made all the same design choices as Felicia, he appreciates their wisdom, and indeed he has been interested in playing it for a while. He understands and approves of the plan that Felicia and her party had previously concocted. Gerald and the party play through the quest effortlessly. The party wins when he casts ice thunderbolt on Cephalopokus the Land Squid.

Finally, imagine Horace is playing *Mazes & Monsters* with a group of people he has never played with before. He can't use his regular character, but they give him a menu of several, all created by Felicia. He thinks that the level twenty-eight wizard is a very good character and is happy to adopt it. They also say that the quest they're going on is one where a couple of different plans might work. They don't have time to develop one with Horace, but once more they give him a menu of choices developed with Felicia. He makes a choice, they follow his suggestion and play through the quest. Horace plays fluently. The group wins when he casts ice thunderbolt on Cephalopokus the Land Squid.

In these cases, Felicia's play is more self-devised and more self-directed than Gerald and Horace's. Her gameplay manifests causal powers that *belong* to her more fully than Gerald and Horace's play belongs to them because the plan and character design that animate that play were created by her. To explain Gerald and Horace's gameplay, we must cite facts beyond them; we must cite Felicia's creative activities. We cite these not because they have been imposed on Gerald and Horace, but because Felicia made the mental artifacts that they are borrowing. On the other hand, an explanation of Felicia's play need not venture beyond Felicia herself. All she does is the result of her own handiwork. And so according to our rough standard about explanation and the liberal conception of agency from before, Felicia's agency is greater than Gerald and Horace's, at least when it comes to playing *Mazes & Monsters*.

The general point is this: other things being equal, coming by your action-guiding states secondhand—borrowing or imitating or choosing from a menu—diminishes your agency because it corresponds to a part of the explanation of your activity that will necessarily advert to things outside of you. This is true even if you tick all the usual boxes contrived by philosophers of action: if you, for example, wholly endorse a perfectly good plan for perfectly adequate reasons. In this case, you may meet all the familiar requirements imposed by reflective agency, but your overall agency, the agency corresponding to the

liberal idea of difference-making, would be greater if your activity were more self-initiated—if you had made your own plan.

We can reinforce this idea with more examples. An oracle tells Irving that the objectively best life will be had by living just as the inventor of Vaseline lived. This oracle has a long and spotless track record, so Irving is justified in believing her. Through painstaking research Irving comes to learn a great deal about the inventor of Vaseline, the legendary Robert Chesebrough. On the strength of the oracle's testimony, Irving sets about to imitate Chesebrough (or "Old Cheesy" as he calls him) in as many ways as possible. He internalizes as many of Old Cheesy's habits, dispositions, preferences, and values as possible. Of course, Irving can't adopt all of Old Cheesy's plans and projects, since they live in different times and places, but he can adopt reasonable analogues. With enough time, this imitation becomes seamless, something Irving automatically endorses. What should we make of Irving's way of life? I think we would be hard-pressed to call it an autonomous life. True, Irving *chose* a life of wholesale imitation, and he chose it for good enough reasons. It was not foisted on him, and he was not tricked into it. But that lone decision is a very thin reed to support a person's entire agency. Besides that seminal choice, nothing about Irving's behavior is ultimately his.

The rote adherence to rules presents similar problems. Norm is an older man who long ago internalized a voluminous system of rules that cover nearly every decision he has to make. And when he comes across a situation that is not treated explicitly by any of his rules, he can be sure that his bounty of higher-order rules will yield a new rule automatically. In this sense, Norm's activity is completely exhausted by the set of rules he adopted many years ago. He lives on automatic pilot: he can make all his choices with as little invention, improvisation, or practical wisdom as it takes for most of us to microwave a burrito. We can stipulate that these rules cohere with Norm's "deep self" and are, objectively speaking, perfectly good. We can also stipulate that Norm assembled this network of rules himself. And yet, despite all this, Norm's agency seems seriously stunted: he seems like a windup monkey left to totter around the nursery for too long, detached from its own actions, a witness to the passive application of a rule. To be clear, I don't think the problem with Norm is the rule-following *per se*. It is that his rules are so comprehensive that they leave no room for present-day Norm to intervene and contribute anything to the conduct of his life and, moreover, that his adherence to them is mechanical and unintelligent. Had Norm adopted some high-level, abstract rule like the Doctrine of the Mean or the Golden Rule or those suggested by *The 7 Habits of Highly Effective People*,

it may not threaten his autonomy, since that kind of rule leaves plenty of room for improvisation. But Norm has done something more radical: he has, to borrow Jonathan Gingerich's word, surrendered his freedom of spontaneity.[20]

In all these examples, a person's agency is affected by the provenance of the states that produce their actions. It is stunted when those states do not originate with them and strengthened when they do. This question of provenance is also one of creativity. Creativity, as we observed, is a capacity to produce things that are original for the person creating them. Thus the contrast in each of the examples I laid out above is between someone who comes by their plan, policy, or other action-guiding state creatively and someone who comes by it uncreatively. (Perhaps it's better to see the difference as a matter of degree: between someone more creative in coming by this state versus someone less so.) And this contrast corresponds to a difference in agency in the liberal sense. Gerald and Horace manifest less agency in their gameplay than Felicia does because that play is borrowed—it is not original to them. Irving manifests less agency than Old Cheesy because the approach to life that they share reflects Cheesy's efforts, not Irving's. And Norm is less of an agent than his more spontaneous counterpart because there is less exercise of his powers of creativity, less opportunity for present-day Norm to make something original in the way he behaves. Taken together these examples provide an opening case for my main thesis:

> *Creativity*. Other things being equal, a person's activity exhibits greater agency in the liberal sense when it is creative.

It's important to mind the hedge here: plenty of very creative activities may exhibit rather little agency for reasons unrelated to that creativity. The claim simply says, in the lingo of contemporary normative philosophy, that creativity is a pro tanto agency enhancer.

6.

Before coming to the second and more ambitious argument for *Creativity*, it will be helpful to answer some preliminary challenges to the claim. At the center of my argument is the idea that patterns of activity that spring from plans or designs that I crafted myself are more fully mine than those acquired from elsewhere. One consequence of this position is that two people can have different degrees of agency even when acting in nearly identical ways. By itself this is not strange: there are many cases where what matters to

[20] (Gingerich 2022).

questions of agency is not just the contents of a person's mind but the origin of those contents. The originality of a pattern of behavior is just one more way that origin matters.

We might think it strange, however, that the difference between greater and lesser agency might lie quite directly in a person's ignorance. A scholar and an ignoramus dress exactly as Oscar Wilde did. If my argument is sound, then the former's agency may be rather less than the latter's simply because he knows so much more about the history of their common style. The scholar can be no more than an imitator, but the dunce can be a second originator—someone who comes to Wilde's style themselves. I don't think this is so strange, though. It is a special case of the more general point that knowing too much can make it harder to act. The person who knows too much has a more complete and so more cramped image of the universe, and this makes it comparably harder for them to find room for their own agency. Everything they consider doing they suspect of being just another historical pattern playing itself out. This is a deeply dismaying thought. It is also why omniscience seems incompatible with agency.[21] And it is why Nietzsche says that "forgetting is essential to action of any kind, just as not only light but darkness too is essential for the life of everything organic."[22] Of course, as Nietzsche also says, we need history to give us some sense of what we are doing, so ignorance is not the preferred way to overcome the past: much better "to break up and dissolve a part of the past … by bringing it before the tribunal, scrupulously examining it and finally condemning it."[23] This is exactly what creativity offers.

Still, we might wonder, is there nothing that the uncreative person can do to make an activity *their own* in the way that the activity of a creative person is? Let's take another example. Steve becomes a stevedore because his father was a stevedore before him and his father before him. Steve is initiated into his stevedoring by imitating others, and so this life is unoriginal and uncreative. But he loves it deeply and authentically: pictures of cargo containers decorate his walls, banners from the ILA Local 1814 hang from his porch, stolen merchandise fills his garage. Is his agency forever limited because the object of his sincere attachment is acquired through imitation? Can't a sufficiently thoroughgoing identification with this life overcome these limitations?

It cannot. There is an obvious sense in which Steve's activities belong to him, but it is too close to the sense in which the analgesic's effects belong to it. In both cases this

[21] (Kapitan 1991).
[22] (Nietzsche 1997, 61).
[23] (Nietzsche 1997, 75–76).

sense involves agent and action being bound together by a nomic regularity—by a pattern—that exists independently of the agent. And this, I have suggested, is where their agency is less than it could be. Indeed, I think identifying very strongly with a certain preformed way of life can be a form of Sartrean "bad faith". It can represent a denial of the possibility of abandoning or remaking one's way of life, and that is tantamount to a denial of freedom. A person like Steve achieves a very robust sort of lower agency, not unlike the agency of a powerful analgesic. But he does so at the cost of ever achieving the higher forms of difference-making made possible by creativity, the kind of difference-making involved in crafting one's own patterns of activity.

If that conclusion sounds harsh, I may mention some things that Steve can do without abandoning the life of the stevedore. He can, in a word, make this kind of life his own. This does not mean identifying with it more strongly, for the problem is not the strength of his bond. It means exercising small amounts of creativity within the bounds of that life. Steve can realize some measure of creativity in his chosen life by fashioning himself as a distinctive sort of stevedore with a distinctive approach to stevedoring. Steve can, in Nietzsche's words, *give style* to his stevedoring. "Here a great mass of second nature has been added; there a piece of first nature removed—both times through long practice and daily work at it. Here the ugly that could not be removed is concealed; there it is reinterpreted into sublimity."[24] Not being a member of the noble order of stevedores myself, I cannot say much about how he might do this, but perhaps Steve can develop his own special rhythm to unloading ships or find ways to integrate popular music into the task or be a leader in making the union hall more welcoming to untraditional stevedores or cultivate an aesthetic sensitivity to hustle and bustle of the docks. These efforts may prove trifling or superficial, but it's Steve's job to make them otherwise. In "giving style" to his chosen form of life, Steve may be able to self-fashion without leaving his preferred form of life.

7.

My first argument for *Creativity* relied on the idea that when an action-guiding state reflects our creativity, it is more fully and ultimately attributable to us than if we take it from somewhere else. My second argument relies on a different feature of creativity—on surprise.

[24] (Nietzsche 2001, sec. 290).

"Creativity is the ability to come up with ideas or artefacts," Boden says, "that are *new*, *surprising*, and *valuable*." Forget about the first and third of these and focus on the surprise part. The relevant kind of surprise, Boden continues, comes in degrees:

> An idea may be surprising because it's unfamiliar, or even unlikely—like a hundred-to-one outsider winning the Derby. This sort of surprise goes against statistics. The second sort of surprise is more interesting. An unexpected idea may 'fit' into a style of thinking that you already had—but you're surprised because you hadn't realized that this particular idea was part of it. Maybe you're even intrigued to find that an idea of this general type fits into the familiar style. And the third sort of surprise is more interesting still: this is the astonishment you feel on encountering an apparently impossible idea. It just couldn't have entered anyone's head, you feel—and yet it did. It may even engender other ideas which, yesterday, you'd have thought equally impossible. What on earth can be going on?[25]

These final examples of creativity are the important ones. They occasion not just the shallow kind of surprise that we experience when a fat horse wins a race, but the durable bemusement that comes from not being able to figure out how or why someone came to an idea, or even, perhaps, how to understand the idea. This is the kind of surprise, or astonishment, that we experience when we find something inexplicable, when its creation is not just statistically unexpected or incompatible with the prevailing *gestalt* but nearly impossible to fit into our reigning understanding of the world.

Of course, we can think that surprise is a feature of creativity without doubting that creative acts can *ultimately* be fit into any particular picture of the world (for example, a "naturalistic" one).[26] Our questions about explanation are scalar, not all-or-nothing, and I am suggesting that the things we judge especially creative will be ones that are difficult to explain using whatever our existing tools for explanation are, and that the explanations we do produce will likewise be worse than our explanations of less creative activity. This means that what we find especially creative may be different from what others do. If scientists ever discovered a convincing biological story about how Shakespeare and no one else was able to write *King Lear*, then those scientists might find Shakespeare less creative than we do. (That we find such a story far-fetched is a testament to our admiration.)

[25] (Boden 2003, 1–3).

[26] Thus I don't necessarily disagree with arguments like that of (Kronfeldner 2009).

Creativity is therefore indexed not to the world itself, but to our best explanatory resources, and when I say that creativity is, to some extent, inexplicable, I am saying that the explanation we can fashion with those resources is comparatively poor.

But what do surprise and inexplicability have to do with agency? To answer this question, we must say a little more about the goals of explanation. According to the lore of our forefathers—the logical empiricists—an explanation is a derivation of the explanandum from two sets of premises: a set of laws of nature and a specification of various worldly conditions. If you want to explain why the kettle boiled at four o'clock, you derive that fact from the kettle's initial condition using general laws about thermodynamics and the boiling point of water. Likewise, if we wanted to explain a certain action, we would need a set of initial conditions—perhaps the agent's beliefs, desires, and capacities—and suitable laws of human behavior. We would then need to demonstrate that these things jointly entail the performance of the act.[27]

This account of explanation has many problems and essentially no living partisans. But one of its insights has endured. The principal goal of explanation is to show how a mysterious phenomenon fits into a larger regularity. When we explain a thing, we are trying to unify it and other, seemingly different phenomena under common patterns. Our explanations improve as the patterns get more general and more inclusive, and as we are forced to rely on fewer and fewer independent epicycles and caveats to capture our target phenomena. Explanation aims for the ideal that Kant calls "systematic unity": the condition that "makes ordinary cognition into a science, i.e. makes a system out of a mere aggregate of it."[28]

The kinetic theory of gases, for example, is an effective explanation because, as Michael Friedman explains, it offers "a significant unification in what we have to accept. Where we once had three independent brute facts—that gases approximately obey the Boyle-Charles gas law, that they obey Graham's law, and that they have the specific heat capacities they do have—we now have only one—that molecules obey the laws of mechanics."[29] Newtonian mechanics also succeeds because of its ability to unify different phenomena under more general patterns. We could be forgiven for thinking that the celestial and terrestrial realms were radically different worlds governed by fundamentally different laws. But Newtonian mechanics manages to explain the phenomena of both using

[27] (Hempel and Oppenheim 1948).
[28] (Kant 1999, A832/B860).
[29] (Friedman 1974, 14–15).

a small store of common physical principles, and that's why they are such good explanations. Likewise, Darwinian explanations succeed because they allow us to see how a great diversity of phenomena—camel humps, bacterial flagella, the silly dances that birds of paradise do, the silly dances that human beings do—are manifestations of simple principles about variation, survival, and reproduction.[30]

It is here that explanation's conflict with agency emerges. These explanatory aims and the demands of agency we have already noted pull us in opposite directions. A creature's agency depends on its being the ultimate explanation for some state of affairs. It depends on localizing some bit of difference-making *within the agent*. But explanation's own constitutive demands urge us to find, as far as possible, the same difference-making potential in the impersonal forces of the wider world.

I think this tension lies at the heart of our concept of agency. It is first cousins with a famous worry from Thomas Nagel. For him the worry is about a certain "perspective" we can take on ourselves.

> From an external perspective … the agent and everything about him seems to be swallowed up by the circumstances of action; nothing of him is left to intervene in those circumstances. This happens whether or not the relation between action and its antecedent conditions is conceived as deterministic. In either case we cease to face the world and instead become parts of it; we and our lives are seen as products and manifestations of the world as a whole. Everything I do or that anyone else does is part of a larger course of events that no one "does," but that happens, with or without explanation. Everything I do is part of something I don't do, because I am a part of the world.[31]

If we borrow Nagel's terms, we can put the worry like this. The demands of explanation push us to see our actions from this "external perspective"—not as fundamentally and irreducibly *our* doing, but as an undistinguished part of a global pattern. And this threatens to rob us of our status as difference-makers.

There is a natural worry about my presentation of this "tension". I seem to be suggesting a rivalry between the explanatory work done by individual entities and that done by general principles reflecting the overall structure of the world: a local explanation, which secures the greater agency, versus a global one, which is urged on us by the demands

[30] (Kitcher 1989, 512–15).
[31] (Nagel 1986, 114).

of explanation itself. But are these two approaches really opposed to each other? When I explain an explosion by citing the combustibility of its fuel, I am implicitly relying on a general principle about the conditions under which that substance explodes. And when I explain the same explosion with the general principle, I am implicitly attributing a certain causal power—combustibility—to that substance. Thus, we might suggest, the global and local are not competing forms of explanation, but different sides of the same coin.

This concern relies on a substantive assumption. If we are going to explain something by attributing a power, then our explanation relies on a general principle articulating the dynamics of that power. But this assumption is false. If my headache disappears after I take an aspirin, we may attribute my relief to the analgesic powers of the aspirin. This is a serviceable explanation even though aspirin doesn't always relieve headaches, and even if we cannot even begin to formulate a general principle listing the conditions under which it does relieve headaches. The local explanation that attributed analgesic powers to the aspirin does not rely on anything like a general principle because there probably is no such principle and certainly none we could articulate.

Nancy Cartwright has shown this to be a situation we find ourselves in constantly. We can explain an event by attributing certain capacities to a thing, but we do not understand these capacities well enough to codify their operation in general principles, even ones embroidered with *ceteris paribus* clauses. If we tried to spell out these principles, we would probably end up saying something false. In these moments we may kid ourselves by thinking that the general principle is closer to hand than it really is, or that it can aid our explanation despite being unavailable to us. But this is an illusion. And it is an especially ridiculous illusion in explanations of human action, where we typically have nothing remotely like a universal principle in mind, and yet manage to offer decent explanations.[32]

So the tension remains. Our workaday explanations will tend to be local. They will locate explanatory ground in the agent. These explanations would be better if they could continue to capture the relevant phenomena while being more general and more unified with other explanations. But this improvement will also tend to efface the agent by taking the difference-making out of their hands.

What should we do about this? We might propose metaphysical remedies. We might put forward claims about the essential nature of action designed to secure the agent's

[32] (Cartwright 1989, chap. 4).

place in action explanations amidst even the strongest explanatory headwinds.[33] I am not necessarily opposed to this, but I think there is a more direct response. We can resist the effacement of our agency *through that agency*, through the behavior we introduce into the world. We have a say in how closely our explanations can approximate the ideal associated with the "external perspective" because we have a say in producing the phenomena they are trying to explain. We can, in effect, act in a way designed to thwart the threatening explanatory tendency.

A sense of this possibility comes out of our daydreams about counterfactual versions of ourselves. We can easily imagine vaguely Huxleyan worlds in which human behavior is radically easier to explain because it is more homogeneous. Maybe evolution never fostered the parts of us that care about art and religion. Maybe a relentless authoritarianism has pruned away all but the most prosaic interests. Maybe our imaginations have been smoothed over by our telephones. The explanations that best fit the behavior of our cousins in these worlds will be better in a certain abstract sense: they will be simpler, more parsimonious, and, most importantly, more unified. But these explanations would fail as explanations of *us*—of people as varied and unruly as actual human beings. It follows, I want to suggest, that our agency is greater than that of these imaginary counterparts. Because it is harder to subsume our behavior under simple, general principles than it is theirs, it is that much harder to take up Nagel's external perspective on us and see our behavior as part of the impersonal machinery of the universe. This means that when we come to explain the behavior of people like our actual selves, we end up assigning a greater share of the explanatory ground for this behavior to the individual powers of agents. We will still be able to offer explanations of course, but these explanations will rely on stubbornly particular facts. In other words, because our behavior is so unpredictable and occasionally so inexplicable compared to our ovine cousins, it will be that much harder to see our behavior "as products and manifestations of the world as a whole".

But we can also imagine less tractable versions of ourselves, individuals who think *we* are the sheep. In fact, we can do better. We can try to become this version of ourselves by deliberately frustrating explanation's demands. We can behave in ways that make it harder to assimilate our behavior to universal patterns and therefore make it more appropriate to understand that behavior as an expression of our particularities. Doing this

[33] Agent-causation theories do something like this: (Chisholm 1976), (Clarke 1993), (O'Connor 2000).

obviously requires an awareness of our own behavior compared to others', and so it requires the reflective capacities I have been expounding on here. But it also requires something more.

8.

This brings us to the connection between agency, inexplicability, and creativity. In some cases and in some respects, a creature will exhibit greater agency for reasons that make the best explanation of her behavior comparatively poor: because her behavior is so unusual or novel. Insofar as creativity is a talent for these qualities, it offers a distinctive opportunity for expanding and preserving our agency. In a word, creativity can make us greater agents by stymying explanation's drive to abstract from and efface our individual powers. Something like this was Frank Zappa's ambition when playing the guitar: "I have a basic mechanical knowledge of the operation of the instrument, and I've got an imagination, and when the time comes up in the song for me to play a solo, it's me against the laws of nature." In doing this, we can, in principle, make it so that the best account of our activities—the best explanation given the phenomena as we find them—locates more difference-making *within* us than the explanation of an uncreative act would.

An example may make this suggestion clearer. The music that traditionally accompanies the arrival of circus clowns is a march called "Entrance of the Gladiators". It was written by Julius Fučík, a prodigious composer who produced more than four hundred works for military band. If we want to explain the event that was Fučík's composition of this famous but aesthetically unremarkable march, we would probably be satisfied with an account that reflected his workmanlike approach: something detailing the specific demands of military music, Fučík's musical training, the prevailing fashions in Czech popular music, requests from his patron, and Fučík's baseline musical preferences. This isn't going to be an exhaustive explanation, of course—there is an element of serendipity in his coming up with a catchy tune—but it will be pretty good.

We would certainly feel that this kind of explanation is more adequate to Fučík's achievement than an analogous explanation would be of a more creative act—of, for example, Beethoven's creation of the late string quartets. Part of what makes those quartets extraordinary is our astonishment at their very existence. They are a miracle. We can scarcely fathom how they could have followed from what came before. They seem to leapfrog fifty years of musical development. They come from the future. Any explanation citing such mundane factors as the demands of a string quartet, the requests of a patron, or

Beethoven's preferences will seem monstrously inadequate to the inventiveness of these works. It will feel inadequate because it cannot answer a pivotal question. How did Beethoven come to compose something extraordinary instead of something ordinary?

There are a few ways we can respond to this inadequacy. We may give up and say Beethoven's achievement was dumb luck—the musical equivalent of a hole-in-one. But we should be loath to do that for someone who makes a habit of creativity. And so we are likely to turn to something more *sui generis*, to claims about Beethoven's distinctive compositional powers. This is what critics do when they talk of "Beethoven's genius" or his "expressive resource" or his "strength and imaginative force" or his "spiritual power" or his "limitless musical resource".[34] These are phrases we turn to at moments of impasse: when the material provided by Beethoven's desires, techniques, and interests plainly does not explain the creation of something exceptional, but it would be a mistake to fill the explanatory gap with references to serendipity, luck, or divine intervention. These phrases denote Beethoven's personal powers, capacities that we want to mark as distinctively his.

These phrases are a common recourse for overawed critics. Wesley Morris says of Sidney Poitier: "The acting he did required every inch of his long body—for exuberance, rapture, caution, solemnity and rage. … Some of that is training; he was our most famous Black Method actor. The rest is simply *him*."[35] Or criticism that resorts to eponymous adjectives for otherwise ineffable qualities: Wagnerian bombast, Beckettian exhaustion, Pinteresque foreboding, Lynchian strangeness. These expressions are not wholly unexplanatory, but they are certainly *less* explanatory than characterizations that suggest how an artist fits under general principles. They are explanations that include a residue of inexplicability, sometimes quite a large one. And this, I think, is something characteristic of attempts to explain creative activity.

The comparison between Beethoven and Fučík is meant to bring out the connection between agency, creativity, and inexplicability. Because it is harder to subsume Beethoven's activity under general laws in the way explanation aims to do—harder than Fučík's—but we nonetheless want to attribute that composition to Beethoven, rather than to luck or the muses, we end up turning to *Beethoven himself* and his distinctive compositional powers. Such an explanation is worse as an explanation than the one we can offer of Fučík, but it's the best we can do for Beethoven. One consequence of this, I believe, is that these explanations present Beethoven as a greater agent than Fučík, as more fully

---

[34] These quotes are from two studies of the quartets: (Marliave 1928) and (Kerman 1979).
[35] (Morris 2022).

responsible for his compositions. It is harder to see Beethoven's composing as simply *how things go* given certain antecedents (compositional strategy, musical fashion, patron demands). It is harder to see it as an instance of a system of general laws. We must instead see *Beethoven himself* as an irreducible and ineliminable difference-maker in the course of events. And to do this is to see Beethoven as that much more of an agent.

9.

I have made two arguments for *Creativity*. The first is that when our activity is creative it is more fully attributable to us because the mental artifact that produced that activity—a plan, a policy, a comportment—originates with us rather than somewhere else. And this, according to the liberal idea at the heart of the liberal conception of agency increases our agency relative to a hypothetical foil who comes by their activity second-hand. The second argument claims that when we are creative, it is harder to subsume our behavior within the patterns that describe the dynamics of the world and that means that the explanation of our behavior will assign a greater share of its ground to our personal powers. These arguments rely on slightly different features of creativity and suggest slightly different paths to my thesis. Nonetheless, I think they point toward much the same idea about the *kind* of agency that creativity makes possible.

To get a better handle on that idea, I turn to what Korsgaard calls "the ideal of agency". This ideal, Korsgaard says, is one "of inserting yourself into the causal order, in such a way as to make a genuine difference in the world."[36] I think this is a very good way of summarizing what is at stake in the liberal conception of agency, but I want to add one clarification that I think Korsgaard overlooks. What she calls the "causal order" is not just a series of events lined up like toy soldiers. It is, as the name suggests, an order: a network of objects and events organized into *patterns*. Our agency is bound up with these patterns because it is through explanations involving them that attributions of difference-making are made. I have argued that our reflective powers aid our pursuit of Korsgaard's ideal by allowing us to orient ourselves within these patterns in such a way that our behavior constitutes us as the bearers of corresponding causal powers. On first inspection, then, reflective agency is the agency of effectively "inserting" oneself into the causal order.

Creativity does something else. It makes it possible for us not only to insert ourselves into Korsgaard's "causal order" and constitute ourselves as nodes in that order.

[36] (Korsgaard 2009, 89).

It allows us to be responsible for certain parts of the form of that order, for the *patterns* that make up the causal order's structure. The kind of agency made possible by creativity is one on which the causal patterns that Korsgaard treats as more or less given may be attributed to *the agent*. When Felicia plays *Mazes & Monsters*, she is instantiating a certain pattern corresponding to her game plan. But her relationship to this pattern is very different from Gerald and Horace's relationship to it. Because Felicia is the author of this plan, the pattern that she introduces into the world in following it is, to a much greater extent, *attributable to her*.

This is especially obvious in Beethoven's case. Beethoven's music suggests, as Dmitri Tymoczko puts it, that "music is fundamentally under our control, that we can change the rules when we want to—that we can, for want of a better phrase, *take ownership of style itself*."[37] This ownership of style is the essence of the kind of agency creativity affords us. The late quartets exemplify new ways that music can be, and the musical patterns that describe those quartets and their production can be attributed to Beethoven in a way that the patterns describing the activity of more conventional composers cannot.[38]

This is exactly what we say about Beethoven when we talk about his influence. When we say that Schumann and Mendelssohn, Berlioz and Bizet, couldn't escape Beethoven's shadow, that they lived in a world he created, that he was the silent co-author of all their music, we mean, of course, that his musical ideas are found in their work, but we also mean that the patterns of musical activity that these composers found themselves instantiating—their *style*—are, to a surprising degree, Beethoven's handiwork. In saying this we are acknowledging an agential power that Beethoven possesses but Fučík does not: the power to be ultimately responsible for patterns of human behavior.

Some care is needed on this point. This example may lead us to conflate two things that must kept apart: the ability to introduce new patterns of activity into the world and the fortune of having them adopted by others. In Beethoven we find both originality and influence, but it's the first thing that matters. It would still be appropriate to credit Beethoven with the introduction of a new pattern of music-making even if Schumann and Mendelssohn never existed. Even then there would be at least one individual who instantiated that novel pattern—there would still be Beethoven himself.

[37] (Tymoczko 2023, 456).
[38] Here I bracket questions about style is, but see (Eck, McAllister, and Vall 1995), (Riggle 2015), and (Hopkins and Riggle 2021).

When I say that Felicia and Beethoven introduce new forms into the "causal order" of the universe, I don't mean that they change the laws of nature, of course. The patterns that we call on to explain a person's behavior may sometimes correspond to nature's majestic laws, but usually they won't. They will be the patterns of social life—"institutional" facts some call them. These are genuine regularities that explain some target phenomena, but which are nonetheless the product of human agency. Intellectual property law, the biology of Pokémon, alternate side parking, the nation-state, emoji syntax, mathematics preprint servers, tenant farming, Christmas tree decoration, rickrolling, the nuclear family, cockney rhyming slang, subway etiquette. All these institutions are the site of distinctive patterns of activity, patterns we can and should rely on in explaining how people behave, but they are themselves the products of human agency. When I say that the creative agent can shape the causal order, I am talking about these things, not Maxwell's Equations.

This is also why *Creativity* is orthogonal to debates about freedom and determinism. If we agree that determinism is, as a metaphysical matter, compatible with the existence of agency, then it will hardly disturb that comity to say that we sometimes exercise our agency by creating new higher-level principles about string quartets. Nor should we think the claim that creative activity is less explicable than uncreative activity is any threat to determinism. If the truth of determinism depended on every event succumbing to *total* explanation, then it would certainly be false.

10.

I have been trying to explain what is distinctive about the agency of creativity with reference to Korsgaard's ideal. But we might challenge my explanation from a different angle. Is there anything particularly remarkable about the kind of agency I have pointed to? After all, there are dozens of things we can do to expand our agency. Lifting weights, building a cyborg suit, becoming a millionaire, getting elected Bundeskanzler, founding a cult of personality: all these things would seem to augment my difference-making powers. So if creativity does too, that need not be surprising or particularly impressive.

Consider a foil, someone who amassed many of the standard agential powers but was not particularly creative—a hulking muscleman, a political giant, a captain of industry. Let's use Cornelius Vanderbilt. From the point of view of our liberal conception of agency, Vanderbilt looks mighty indeed. He is responsible for colossal changes in the world:

railroad lines, steamships, harbors, shipping depots, boomtowns, banks, train stations, mansions, universities, and endless Fifth Avenue social intrigue. He was a stunning difference-maker. And yet from a different point of view, a more general one, Vanderbilt himself seems less important. The ultimate explanation for the steamships, railroad lines, and even the debutante balls, lies not in a single man but in very general forces of history: the development of early industrial capitalism in America, the consolidation of commercial power, and the transition from one stage of social development to another. Vanderbilt's achievement lay in inserting himself into this causal order at a particularly robust point, in constituting himself as the proximate cause of the railroads and steamships and enjoying the benefits of that position. This is nothing to sneeze at, but Vanderbilt himself is hardly responsible for the network of forces. Indeed, in the final analysis he is mostly superfluous to it: if he hadn't existed, some other robber baron would have taken his place, and the world would have gone on much as it did. From this more synoptic point of view, then, Vanderbilt's difference-making seems to dissolve into the background in precisely the way we worried would happen when we take up the "external point of view".

When we first encounter the liberal conception of agency it looks to be all about might. It is the agency of making differences, and to make a big difference you need muscle, money, or clout. But Vanderbilt's case shows this is misleading: these forces operate within the context of larger systems, like early industrial capitalism. If we drop the system from view, then the person exerting these forces will look like they can move mountains. But attention to the wider context reveals that much of what we attributed to the agent is really due, at least in part, to something tucked away in the background. It is not Vanderbilt alone performing these wonders but a much larger segment of the world.

But we cannot take this kind of perspective on Beethoven. We cannot say that he was the foam on the wave of the Viennese School. We cannot say that music would have developed in much the same way without him. We cannot say that if Beethoven had not existed Hummel and Cherubini would have filled the void because they plainly wouldn't have. We cannot say these things because they are not true.

The point, of course, is not about Beethoven's actual influence. We can imagine Hummel or Cherubini exerting substantially more influence on subsequent music in a Beethovenless world. What we cannot imagine is the musical world proceeding in much the same way without Beethoven as it did with him. We cannot imagine this because the Beethoven-dependent path was one that broke with the very patterns that we must turn to in trying to project history in his absence. This makes Beethoven indispensable to our

understanding of the world in a way that Vanderbilt is not. "What you are, you are by accident of birth; what I am, I am by myself," Beethoven wrote to one of his former patrons. "There are and will be a thousand princes; there is only one Beethoven."[39]

I said before that most enhancements to agency in the liberal sense involved the internalization of difference-making capacity within the agent. Organismic form, inborn spirit, and planning all have this character. The idea that creativity represents a higher form of agency is simply a development of this thought. Creativity of the sort exemplified by Beethoven represents the possibility of certain patterns of activity, over and above the token events of that pattern, being attributed to an agent. It represents the possibility of internalizing a capacity for making the high-level regularities that govern the world.

This suggests one final way of explaining why this represents an *elevation* of agency. You and I are apprentices in a master sculptor's workshop. I am tasked with assembling constituent sculpture parts into one of a dozen possible configurations that those parts can fit into. You, by contrast, are told to create a totally new sculpture from whatever materials you prefer; indeed, you are warned against copying. We are both constituting sculptures, but this means very different things for each of us. My sculpture-constitution is a kind of sculpture-assembly. I am arranging the sculpture's constituent parts in such a way that it instantiates a sculptural order that already exists. What you are doing, however, is more like sculpture-determination. You are not just constituting the sculpture according to an existing form. You are creating that form through your sculpting. Now imagine that we are not apprentice sculptors but apprentice human beings, and we are tasked not merely with sculpture-constitution but self-constitution. The same distinction seems possible here. There are some agents whose self-constitution is essentially self-assembly, the merger of constituent parts according to an existing pattern. (Choosing the self from a series of menus, we might say.) And then there are those whose self-constitution is genuine self-determination, the invention, through one's own constitution, of a new pattern of agency.

11.

One fringe benefit of this conclusion is a new way of approaching an old question, the question of creativity's value. According to a standard view, creativity is a disposition to produce ideas that are both original and *valuable*. But some dissent from this connection;

[39] Qtd. (Thayer 1992, 403).

they say that creativity can, in fact, produce perfectly useless things. This has produced a debate about the engineering of a concept: does the best characterization of our paradigms of creativity "bake in" the value of its products?[40]

Both sides of this debate share an assumption about the form that the value of creativity, insofar as it exists, will take. They assume that creativity's value will be inherited from the value of its products. This thought is natural enough, but it is not the only way to think about the value of a capacity. There is, for example, the way that Kant thinks about the value of our rational nature, of our "humanity". For him the value of our rational nature is related not to its status "as the source of *good things* (i.e., of their *existence*), but instead as the source of the fact of their *goodness*—indeed of the fact that anything at all is objectively good."[41] In other words, the value of the capacity Kant calls "rational nature" lies not in its causal relationship with valuable things, but in its *conditioning* relationship with their value.

In this light, my claim about creativity and agency suggests an unexplored possibility. It suggests that creativity may be valuable for the same reason rational nature is. It may be valuable not because of its products but because it enables a kind of agency that *conditions* certain forms of value—because it is, in effect, part of the ground of at least some forms of value in approximately the same way that Kant says our rational nature is. For this to be the case, the "higher" kind of agency I have described here would have to be a requirement for the activities by which at least some forms of value are constituted. There would have to be things whose value is conditioned on valuing activities that are themselves reliant on the form of agency made possible by creativity. Now, I cannot begin to show this here, and so I can do no more than record such an arrangement as a possibility, but it is, I believe, a promising one.

12.

In closing, I should address, if only briefly, a just and important objection. My principal thesis—that creativity is associated with a "higher" form of agency—does not address the problem I put to other approaches to creative agency. It does not explain what it takes to come by this form of agency, or what creative agency consists in. I concede this point. I have not, to return to my metaphor from before, explained from where the rain of

[40] For the affirmative view see (Boden 2003), (Kieran 2014), (Gaut 2018), (Brainard Forthcoming). For dissent see (Hills and Bird 2018) and (Hills and Bird 2019).
[41] (A. W. Wood 1999, 130).

creativity falls. Instead, I have made the negative point that it is a mistake to answer these worries by assimilating creativity to the standard devices of agential control because creativity makes possible a kind of agency above and beyond these techniques. But this still leaves us with a problem. If it is not a constellation of plans and intentions that makes Beethoven's writing the *Eroica* symphony an exercise of his agency, then what *is* it? In virtue of what is this writing something Beethoven *does*?

I can do no more than gesture at an answer to this question, but the answer is something like this. Creativity is one of many behaviors that is unreflective but by no means unintelligent. Peter Railton describes some of them:

> The jazz saxophonist's solo riff, the basketball guard's well-timed jump, the experienced driver's smooth downshift, and the wit's lightning riposte aren't unintentional or mindless behavior, like absent-mindedly tapping one's foot while writing or succumbing without realizing it to the emotional contagion of a crowd. Rather, they are complex, structured, purposeful activities done mindfully but fluently, without deliberation or intention-formation.[42]

It is notable, of course, that at least two items on Railton's list involve forms of improvisation and so embrace a measure of creativity.

Even more notable is the fact that each of these items represents the exercise of a skill. We might therefore be tempted by the idea that creativity is a skill. This is a venerable idea, but there is a familiar problem with it. We generally suppose that skills are directed toward a determinate end: basketball-shooting aims at shots made, cobbling aims at shoes produced. But this directedness seems to rule out the possibility of invention in the same way that planning does. For if, as Berys Gaut explains the problem, "the process of making something is creative, then one cannot know the end: for if one knows the end, one has already created the object. For instance, a poet creating a poem cannot already know what the poem is, for if he knows this, he has already created the poem."[43]

Gaut's response to this problem is in the same spirit as the one from Peacocke we explored earlier. He suggests that a skill's definitive end can be indeterminate, and that creativity can abide in the gaps left by this indeterminacy. But we might wonder about a different approach. Why must skills be teleological in this sense at all? Why say that they are tied to a *particular* end? Presumably it is for reasons of individuation and normativity.

---

[42] (Railton 2009, 97).
[43] (Gaut 2009, 87).

What counts as good cobbling and bad cobbling, and, indeed, cobbling at all, is determined by the activity's constitutive connection to shoes. If we sever this connection between skill and end, it would not be clear what cobbling even *is*. Note that nothing in this explanation requires the characteristic end of a skill to be *given in advance* of its exercise. There is nothing, in other words, that excludes the possibility of what we might call exemplary skills—skills whose characteristic ends and attendant standards are constituted through the exercise of that very skill. Or, more precisely, through what that exercise exemplifies.

This may sound strange at first blush, but we have plenty of examples of activities with this self-constituting character. "This is how you do my new dance," I say while dancing. This is true not because my dance conformed to some existing standard but because it successfully *exemplified* a new standard. I have constituted the dance by dancing.

Taking up this possibility would mean abandoning the idea that the exercise of a skill is always guided by a definite conception of its end. But this need not mean that this exercise is arbitrary, that it is a spasm or fit of caprice. Among other things, this exercise will be guided by the demands of exemplariness itself. Not every odd thing we do is exemplary. To exemplify something means making oneself understood in a particular way—as exemplifying that thing—and when the thing in question is new, that is no trivial task.

This possibility gives us another way of addressing Gaut's problem. We can say that creativity is a skill, or, more precisely, a class of skills that are exemplary in this sense. Not an off-the-shelf skill with an antecedently given end, but a skill that is constituted through its own exemplary performance.

I cannot muster a proper defense of this suggestion, but, if I am right, then it suggests that an answer to our question will lie with a better understanding of a power Kant calls "genius". Some of Kant's readers take "genius" to be little more than a hoity-toity name for what we call creativity, but it is crucial for Kant that genius is not just a capacity for originality. It is a capacity for *exemplary* originality. It is the talent "that gives the rule to art"—a talent for making something not produced according to rules nonetheless conform to rules.[44] The key to understanding what makes Beethoven's creations properly his own will lie, I believe, with an account of how he can "give the rule" to them.[45]

[44] (Kant 2000, 5:306-7). I develop this suggestion in (Walden Forthcoming).

[45] I am grateful to a large group of people for discussions that improved this paper. I am sorry I cannot hope to name all the members of this group, but it includes Sarah Buss, Lindsay Brainard,

Alvarez, Maria, and John Hyman. 1998. “Agents and Their Actions.” *Philosophy* 73 (2): 219–45. https://doi.org/10.1017/S0031819198000199.
Aristotle. 1983. *Physics: Books III and IV*. Edited by Edward Hussey. Oxford University Press.
———. 1986. *De Motu Animalium*. Edited by Martha Nussbaum. Princeton University Press.
Boden, Margaret A. 2003. *The Creative Mind: Myths and Mechanisms*. Routledge.
Brainard, Lindsay. Forthcoming. “What Is Creativity?” *Philosophical Quarterly*. https://doi.org/10.1093/pq/pqae075.
Bratman, Michael. 1987. *Intention, Plans, and Practical Reason*. Harvard University Press.
———. 2007. “Reflection, Planning, and Temporally Extended Agency.” In *Structures of Agency*. Oxford University Press.
Cartwright, Nancy. 1989. *Nature’s Capacities and Their Measurement*. Oxford University Press.
Chisholm, Roderick M. 1976. *Person and Object: A Metaphysical Study*. Open Court.
Clarke, Randolph. 1993. “Toward a Credible Agent-Causal Account of Free Will.” *Noûs* 27 (2): 191–203. https://doi.org/10.2307/2215755.
Coates, D. Justin, and Philip Swenson. 2013. “Reasons-Responsiveness and Degrees of Responsibility.” *Philosophical Studies* 165 (2): 629–45. https://doi.org/10.1007/s11098-012-9969-5.
Coope, Ursula. 2007. “Aristotle on Action.” *Aristotelian Society Supplementary Volume* 81 (1): 109–38. https://doi.org/10.1111/j.1467-8349.2007.00153.x.
Cooper, J. 1997. *Plato: Complete Works*. Hackett.
Duhem, Pierre Maurice Marie. 1954. *The Aim and Structure of Physical Theory*. Princeton University Press.
Eck, Caroline, James McAllister, and Renée van de Vall, eds. 1995. *The Question of Style in Philosophy and the Arts*. New York: Cambridge University Press.
Ferrero, Luca. 2009. “What Good Is a Diachronic Will?” *Philosophical Studies* 144 (3): 403–30. https://doi.org/10.1007/s11098-008-9217-1.
Frankfurt, Harry G. 1988. “The Problem of Action.” In *The Importance of What We Care About: Philosophical Essays*. Cambridge University Press.
Friedman, Michael. 1974. “Explanation and Scientific Understanding.” *Journal of Philosophy* 71 (1): 5–19. https://doi.org/10.2307/2024924.
Gaut, Berys. 2009. “Creativity and Skill.” In *The Idea of Creativity*, edited by Karen Bardsley, Denis Dutton, and Michael Krausz, 83–103. Brill.
———. 2010. “The Philosophy of Creativity.” *Philosophy Compass* 5 (12): 1034–46. https://doi.org/10.1111/j.1747-9991.2010.00351.x.
———. 2018. “The Value of Creativity.” In *Creativity and Philosophy*, edited by Matthew Kieran and Berys Gaut. Routledge.
Gingerich, Jonathan. 2022. “Spontaneous Freedom.” *Ethics* 133 (1): 38–71. https://doi.org/10.1086/720778.
Goldstein, Kurt. 1995. *The Organism: A Holistic Approach to Biology Derived From Pathological Data in Man*. Zone Books.

Keren Gorodeisky, Richard Holton, Samantha Matherne, Antonia Peacocke, Nick Riggle, Karl Schafer, and an anonymous reviewer.

Hempel, Carl Gustav, and Paul Oppenheim. 1948. "Studies in the Logic of Explanation." *Philosophy of Science* 15 (2): 135–75. https://doi.org/10.1086/286983.

Hills, Alison, and Alexander Bird. 2018. "Creativity without Value." In *Creativity and Philosophy*, edited by Berys Gaut and Matthew Kieran. Routledge.

———. 2019. "Against Creativity." *Philosophy and Phenomenological Research* 99 (3): 694–713. https://doi.org/10.1111/phpr.12511.

Hopkins, Robert, and Nick Riggle. 2021. "Artistic Style as the Expression of Ideals." *Philosophers' Imprint* 21 (8): 1–18.

Kant, Immanuel. 1999. *Critique of Pure Reason*. Edited by Paul Guyer and Allen W. Wood. Cambridge University Press.

———. 2000. *Critique of the Power of Judgment*. Translated by Paul Guyer and Eric Matthews. Cambridge University Press.

———. 2002. *Groundwork for the Metaphysics of Morals*. Translated by Allen Wood. Yale University Press.

Kapitan, Tomis. 1991. "Agency and Omniscience." *Religious Studies* 27 (1): 105–20. https://doi.org/10.1017/S0034412500001359.

Kerman, Joseph. 1979. *The Beethoven Quartets*. W. W. Norton.

Kieran, Matthew. 2014. "Creativity as a Virtue of Character." In *The Philosophy of Creativity: New Essays*, edited by Elliot Samuel Paul and Scott Barry Kaufman. Oxford University Press. https://academic.oup.com/book/6463/chapter/150306526.

Kierkegaard, Soren. 1983. *The Sickness Unto Death: A Christian Psychological Exposition For Upbuilding And Awakening*. Edited by Howard V. Hong and Edna H. Hong. Princeton University Press.

Kitcher, Philip. 1989. "Explanatory Unification and the Causal Structure of the World." In *Scientific Explanation*, edited by Philip Kitcher and Wesley Salmon, 410–505. University of Minnesota Press.

Korsgaard, Christine M. 2009. *Self-Constitution: Agency, Identity, and Integrity*. Oxford University Press.

Kronfeldner, Maria. 2009. "Creativity Naturalized." *Philosophical Quarterly* 59 (237): 577–92. https://doi.org/10.1111/j.1467-9213.2009.637.x.

Marliave, Joseph de. 1928. *Beethoven's Quartets*. Oxford University Press.

Marx, Karl. 2000. *Karl Marx: Selected Writings*. Edited by David McLellan. Oxford University Press.

Merleau-Ponty, Maurice. 1963. *The Structure of Behavior*. Beacon Press.

Mill, John Stuart. 2003. *On Liberty*. Edited by David Bromwich and George Kateb. Yale University Press.

Morris, Wesley. 2022. "Sidney Poitier Was the Star We Desperately Needed Him to Be." *The New York Times*, January 7, 2022, sec. Movies. https://www.nytimes.com/2022/01/07/movies/sidney-poitier-legacy.html.

Nagel, Thomas. 1986. *The View From Nowhere*. Oxford University Press.

Nelkin, Dana Kay. 2016. "Difficulty and Degrees of Moral Praiseworthiness and Blameworthiness." *Noûs* 50 (2): 356–78. https://doi.org/10.1111/nous.12079.

Nietzsche, Friedrich. 1997. *Untimely Meditations*. Edited by Daniel Breazeale. Translated by R. J. Hollingdale. Cambridge University Press.

———. 2001. *The Gay Science*. Edited by Bernard Williams. Translated by Josefine Nauckhoff and Adrian Del Caro. Cambridge University Press.

O'Connor, Timothy. 2000. *Persons and Causes: The Metaphysics of Free Will*. Oxford University Press.

Paul, Elliot Samuel, and Dustin Stokes. 2018. "Attributing Creativity." In *Creativity and Philosophy*, edited by Berys Gaut and Matthew Kieran. Routledge.

Peacocke, Antonia. Forthcoming. "The Problem of Creative Intention." In *Philosophy and Art: New Essays at the Intersection*, edited by Alex King. Oxford University Press.
Quine, W. V. 1953. *From a Logical Point of View*. Harvard University Press.
Railton, Peter. 2009. "Practical Competence and Fluent Agency." In *Reasons for Action*, edited by David Sobel and Steven Wall, 81–115. Cambridge University Press.
Riggle, Nick. 2015. "Personal Style and Artistic Style." *Philosophical Quarterly* 65 (261): 711–31. https://doi.org/10.1093/pq/pqv026.
Schafer, Karl. 2023. *Kant's Reason: The Unity of Reason and the Limits of Comprehension in Kant*. Oxford University Press.
Steward, Helen. 2012. *A Metaphysics for Freedom*. Oxford University Press.
Stokes, Dustin. 2011. "Minimally Creative Thought." *Metaphilosophy* 42 (5): 658–81. https://doi.org/10.1111/j.1467-9973.2011.01716.x.
Thayer, Alexander Wheelock. 1992. *Thayer's Life of Beethoven*. Edited by Elliot Forbes. Princeton University Press.
Thompson, Michael. 2008. *Life and Action: Elementary Structures of Practice and Practical Thought*. Harvard University Press.
Timpe, Kevin. 2016. "Executive Function, Disability, and Agency." *Res Philosophica* 93 (4): 767–96. https://doi.org/10.11612/resphil.1451.
———. 2022. "Agency and Disability." In *The Routledge Handbook of Philosophy of Agency*. Routledge.
Tomas, Vincent. 1958. "Creativity in Art." *Philosophical Review* 67 (1): 1–15. https://doi.org/10.2307/2182766.
Tymoczko, Dmitri. 2023. *Tonality: An Owner's Manual*. Oxford University Press.
Walden, Kenneth. Forthcoming. *The Imperative of Genius*. Oxford University Press.
Wood, Allen W. 1999. *Kant's Ethical Thought*. Cambridge University Press.