

This is a section of [doi:10.7551/mitpress/10913.001.0001](https://doi.org/10.7551/mitpress/10913.001.0001)

Beyond the Creative Species

Making Machines That Make Art and Music

By: Oliver Bown

Citation:

Beyond the Creative Species: Making Machines That Make Art and Music

By: Oliver Bown

DOI: 10.7551/mitpress/10913.001.0001

ISBN (electronic): 9780262361750

Publisher: The MIT Press

Published: 2021

The open access edition of this book was made possible by generous funding from University of New South Wales.



The MIT Press

2 Coming to Terms with Creative Machines

We can see no a priori objection to the gradual development of a mechanical life, though that life shall be so different from ours that it is only by a severe discipline that we can think of it as life at all.

—Samuel Butler¹

Words do not express thoughts very well. They always become a little different immediately after they are expressed, a little distorted, a little foolish.

—Herman Hesse²

The formal study of creativity ... [may reach] ... the stage of advancement which botany reached when Linnaeus organized flora into phyla and into classes.

—Mel Rhodes³

How Can Creativity Be Computational?

While artists like Cohen, Lewis, and Cope were reflexively tinkering with their machines, a great body of research and commentary has risen up on the nature of human creativity. This spans myriad academic areas, from developmental and evolutionary psychology to organization science, management, and philosophy. As a subject of study, creativity became a topic of serious focus in the nineteenth century and one of great social importance in the twentieth. In the twenty-first century, it is a veritable juggernaut of debate and research, reshaping our cities, governments, and corporations.

As a concept, creativity has historically been something of a moving target and one that is still very much in motion. According to Raymond Williams,⁴ the terminology of creativity evolved through and in the wake of the Enlightenment. Deriving from the Latin *creō*, meaning “to make,” the

earliest uses of the verb “to create” referred exclusively to the divine creation that was the domain of God, setting the tone for its later designation in the activities of mortal humans: a special form of making that was profound and fundamental rather than mundane. Gradually, forms of artistic practice, beginning with poetry, became designated as sharing this godlike capacity to bring new forms into being, and for a good length of time creativity was exclusively associated with the arts among human endeavors. A persevering association between art and divine creation, through this use of words, suited some of the prevailing conceptions of art during this time, embodying a connection to the sacred.⁵

In the nineteenth and twentieth centuries, along with the rise of atheism, the theory of evolution, and the science of psychology, the concept of creativity came to incorporate a wider set of activities, including scientific discovery, design, and the problem-solving of the everyday bricoleur. The modern sense of the term is most prominently associated with a “power of mind,”⁶ a psychological rather than metaphysical concept, and is less and less limited to any specific domain of activity; *accounting* can be creative. However, as many commentators have observed, the term gets applied inconsistently to different entities. It can denote a process, or a property of the products of such processes, as well as applying to the collective capacity of a group rather than an individual.⁷

Although creativity may have shed its theological origins, it has carried the original contrast with its ontological cousin, “to make,” into the twenty-first century. Making is mundane whereas creativity is extraordinary. Under the microscope these activities may seem hard to distinguish. Yet the latter maintains the higher status of association with a powerful metaphysical trick, not only restructuring existing matter, but, godlike, bringing new essences into the world. Making means making the same thing as everybody else. Creating involves differentiation. It also carries with it the baggage of its historically exclusive association with art. In some circumstances we talk of the “creative industries” as an economic category containing art, music, graphic design, film, and so on, and designate certain workers as creatives. This designation has broadened over time. Recent formal analyses of this sector, such as the work of Richard Florida on the creative class,⁸ correspond to a contemporary sense of creativity as including science, engineering, and research, conflicting with the more specific sense of creative

as referring strictly to art-oriented cultural production. (How exactly we distinguish art-type activities from other activities is a topic for chapter 4.)

Creativity is one of several slippery concepts found in this book. Scientists and engineers can find themselves pining for rigorous, operationalizable definitions that help precisely pin down, measure, categorize, and manage a subject of study. Precision of language, and more formal representations of concepts as in code or mathematics, are powerful tools, but concepts in the wild don't necessarily obey these requirements of precision, as the brief history above highlights. "Ask most people the question 'what is creativity?' and you are more likely to elicit an anecdote, an aphorism, or a metaphor than you are a literal definition," note Cardoso, Veale, and Wiggins.⁹ For Wiggins, "'creativity' is ill-defined, but we do tend to know it when we see it; [but] ... 'creativity' as a word is overloaded, and is usable in distinctly different and confusing ways."¹⁰

In fact, established understanding of how we process concepts such as creativity as mental representations shows that a definitional formulation is generally inadequate.¹¹ Concepts work in practice as effective tools for understanding our world, and communicating that understanding with others, by using one or more prototypes as markers against which we seek a "family resemblance."¹² According to this view of how concepts work in the wild, a concept fans out from a small number of things that definitely fit the concept, toward a wider number of things that lie at the concept's periphery and may or may not fit, as the context demands. As Gregory Murphy illustrates:

You might have an idea of the prototypical dog, for example, that is average-sized, dark in color, has medium-length fur, has a pointed nose and floppy ears, is a family pet, barks at strangers, drools unpleasantly, and has other common features of dogs. Yet, this prototype may be something that you have never specifically seen—it is just your abstraction of what dogs are most often like. Other dogs vary in their similarity to this prototype, and so they differ in their typicality. Miniature hairless dogs are not very similar to the prototype, and they are considered atypical; ... if you saw a thing with four legs but a very elongated body, no hair, and whiskers, you might think that it was somewhat similar to a dog, but not similar enough to actually be a dog. However, this is the point at which different people might disagree, and you might change your mind.¹³

According to this body of theory, typicality is a smoothly varying quantity. At concept boundaries, like the one described above, we are slower to categorize and more prone to change our minds. We also determine category

membership by appealing to nonmembership of other categories. So by extension to creativity, if something exhibits properties of imitation, then it is less likely to be seen as creative. We associate creativity with a number of qualities including novelty, aesthetic value, surprise, non-obviousness, insight, individuality, problem solving, reconceptualizing problems (seeing things differently), and divergent thinking. These are elements that might be used to typify instances of creativity. Similarly, creativity has been described as an “essentially contested concept,”¹⁴ one that by its nature evades being pinned down.

The essentially contested nature of creativity makes it appropriate to appeal to a know-it-when-you-see-it definition of creativity, which is how computational creativity researchers have tended to approach the definition of their core subject of study. Wiggins defines computational creativity as “the study and support, through computational means and methods, of behaviour exhibited by natural and artificial systems, which would be deemed creative if exhibited by humans.”¹⁵ But reassuringly, although the concept of creativity may be ambiguous, studies into people’s judgment of creativity by Teresa Amabile have shown that a robust consensus can be reached in these judgments. That is, within specific domains of creative practice, people tend to agree on what constitutes creativity. This is particularly true of domain experts, who can even agree well on the perceived creativity of an artifact, despite not agreeing on whether they like the artifact or whether it is technically good (a 0.77 correlation among experts in judgments of creativity compared to a 0.52 correlation for liking).¹⁶ Thus despite the necessary subjectivity of creative evaluation, we can, under the right circumstances, reliably use human evaluation as a means to establish when something or someone is creative.¹⁷

Such issues draw attention to the disciplinary differences of approaches that are bound to arise amid any study of creativity. In an article entitled “Soft Sciences Are Often Harder than Hard Sciences,” Jared Diamond describes how some “areas are given the highly flattering name of hard science, because they use the firm evidence that controlled experiments and highly accurate measurements can provide,” whereas “soft sciences, as they’re pejoratively termed, are more difficult to study for obvious reasons. ... You can’t start ... and stop [experiments] whenever you choose. You can’t control all the variables; perhaps you can’t control any variable. You may even find it hard to decide what a variable is.”¹⁸ In such fields, it is less of an issue to bypass the

precise definitions of terms, and the family-resemblance view of concepts supports such a stance. A family-resemblance view transforms the task of dealing with concepts like creativity (and art, which we come to later) from one that is constantly confounded to one that is stable.

The concept of creativity comes with its own set of unmistakable prototypes that readers will share. Our prototypes of creativity might derive from famous individual figures, artifacts, and acts of creation—the *Mona Lisa*, Einstein’s relativity, Madonna, Mozart, the iPhone—or from personal examples of individual creativity, such as those involving the developmental progress of one’s own children. Every parent has an example of recognizing creativity in their children, even if their child is *actually* doing exactly what we would expect them to do at that stage in their development.

As these examples illustrate, creativity is additionally confused by the different objects of focus the term might be applied to. As early as 1961, Mel Rhodes catalogued how overlapping strands of research into creativity had begun to appear and were causing confusion. He termed these strands person, product, process, and press (the meaning of the last term is not that obvious; it refers to the relationship between human beings and their environment). “Each strand has unique identity academically, but only in unity do the four strands operate functionally. It is this very fact of synthesis that causes fog in talk about creativity and this may be the basis for the semblance of a ‘cult.’”¹⁹

Novelty, Value, and the Rest

What are the properties that guide our judgment about whether something conforms to our prototypical expectations of creativity? Two dimensions that are almost universally agreed to be part of the reckoning of creativity are *novelty* and *value*, and a third, more contested, candidate is *surprise*, or *non-obviousness*. Novelty doesn’t warrant much scrutiny at this stage, except to warn against the assumption that there is anything objective about measuring the novelty of something. Any measure of similarity or dissimilarity between two things always assumes a frame of reference, and different frames of reference will produce different measures of similarity between entities. A wine connoisseur will have greater capacity for distinction than a novice and would be more attuned to the discovery of a novel wine. But this aside, it would be safe to argue that identity is strictly mutually exclusive to creativity; novelty is a nonnegotiable criterion for creativity.²⁰

Value is the source of more conceptual trouble. It is, like novelty, also subjective, defined relative to a person or group, situated in a given context. Value, like creativity, may also be seen to represent a cluster of concepts with differences in the contexts in which they are applied. For example, Amabile defines creative results using the cluster of terms “appropriate, useful, correct, or valuable,”²¹ and other creativity researchers have also tended to diffuse the focus on value alone with a similar set of related adjectives.

As I will argue, value makes sense from a psychological perspective where creative individuals work toward creative goals. But it becomes a more elusive quality in the distributed view of creativity that I present in the following chapters, which is irreducible to individual psychological creativity. At a social level, we see a creative process which has more in common with biological evolution, unfolding blindly. While the creations of an inventor or artist have value to that creator or to the people around them, there is nobody or no thing that derives value from the creations made by an evolutionary process. Yes, the elephant’s trunk is certainly useful to the elephant, but it was not created because the elephant’s ancestors needed trunks. Whatever process did create the trunk wasn’t *directly* driven by the value of the trunk.

Surprise or non-obviousness is occasionally also included in discussions of creativity, and this usually depends on the application domain in which the discussion is being held. For example, two theorists I will discuss in the following sections include these terms. Dean Simonton refers to the US patent-office definition of a patentable invention as including the criterion of non-obviousness,²² and Margaret Boden includes surprise in some (but not all) of her influential discussions of creativity.²³ We can consider either of these to be a component in a list of elements that define prototypical creativity. The term “compelling” has also been used in a similar way.²⁴ In some ways, surprise, non-obviousness, and compellingness are linking concepts that overlap with both novelty and value. For example, in science, something can’t be *obvious* and *novel* at the same time, whereas in humor, something may gain *value* from its *non-obviousness*.

Creative Thought

Psychological research on creativity is concerned with the processes taking place in people’s minds when they create. A starting point for such inquiry lies in the reflexive accounts of celebrated creative individuals.

Mathematicians and scientists such as Poincaré, Einstein, and Kekulé have provided well-known autobiographical accounts of their thought processes that offer a starting point for formalizing aspects of creativity. The French mathematician Henri Poincaré, for example, observed:

Often when one works at a hard question, nothing good is accomplished at the first attack. Then one takes a rest, longer or shorter, and sits down anew to the work. During the first half-hour, as before, nothing is found, and then all of a sudden the decisive idea presents itself to the mind. It might be said that the conscious work has been more fruitful because it has been interrupted and the rest has given back to the mind its force and freshness.²⁵

The idea that human creativity might come about through different stages or layers of cognitive processing became formulated in theory as researchers turned their attention to explaining these accounts. One of the earliest such formulations is Graham Wallas's four-stage model of the creative process,²⁶ developed in the 1920s. The stages Wallas identifies are those of *preparation*, *incubation*, *illumination*, and *verification*. According to Wallas's model, a problem is defined and elaborated upon in the preparation stage. The incubation stage is then manifest as a kind of passive mulling over of the problem, leading to a moment of illumination, which can then be more formally and logically verified. The middle stages of incubation and illumination are understood as involving subconscious mental activities, during which candidate solutions and combinations of ideas are processed. The beginning and end stages involve more conscious processes, where the problem is framed and the solutions are examined.

What happens during incubation and how these incubated ideas present themselves in the form of illumination have subsequently become two prominent areas of inquiry in the psychology of creativity. It is a fascinating prospect (especially from the point of view of computational creativity) that the creative mind does its magic out of sight of our conscious awareness, like a separate machine that we feed creative tasks to. But it is a view that tallies well with the majority of intuitive and anecdotal accounts of creative inspiration. Like Poincaré, you may be familiar with the experience of suddenly encountering a solution to a problem at an unlikely moment, or having a great idea come to you and not being able to trace where it sprung from. More exotically, the chemist Kekulé recounts moments of dreamy reverie underpinning his discoveries.²⁷ In one case he daydreams of snakes biting their own tails, suddenly realizing that this vision provides

the solution to his problem of the structure of benzene; the molecules form rings.²⁸ Poincaré's accounts also allude to these dreamlike states of mind: "Ideas rose in crowds; I felt them collide until pairs interlocked, so to speak, making a stable combination."²⁹

Attempts to develop concepts to explain what is going on under the surface of our conscious awareness have been in no short supply. Arthur Koestler, a journalist who turned to popular science, writing extensively on creativity, elaborated on the illumination stage, identifying different forms of emotional response grounded in surprise and aesthetic experience. For Koestler, responses to striking new ideas could take the form of aesthetic pleasure (AH), a "eureka" moment (AHA), or amusement (HAHA).³⁰ Koestler was also one of many proponents of the idea that, as the earlier quote from Poincaré suggests, the key process of idea formation involves the combination of previous ideas. He called this process *bisociation*. More recently, and relatedly, Gilles Fauconnier and Mark Turner have proposed that *conceptual blending* is a fundamental tool of human cognition, possibly a critical step in humankind's evolutionary psychology.³¹ A conceptual blend brings together two inputs to create a blended concept or mental space, which, importantly, can then be used to think in new ways, a kind of mental simulation or thought experiment. In Fauconnier's words, "Conceptual blending is a basic mental operation that leads to new meaning, global insight, and conceptual compressions useful for memory and manipulation of otherwise diffuse ranges of meaning."³² Others have described the emergence of new ideas in a way more akin to the incubation of organisms. From a psychoanalytic perspective, Silvano Arieti³³ proposes the *endocept*, a concept in its early stage of formulation, as in the incubation period. By contrast, an *exoecept* takes a mature form that can be properly expressed. The creative mind, then, is an incubator for exoecepts.

Closely associated with these hypothesized phases of incubation and illumination is the idea that creative cognition involves exploratory search. Amabile distinguishes creativity from other forms of problem solving as involving heuristic rather than algorithmic processes, the latter being "those for which the path to the solution is clear and straightforward."³⁴ For example, solving a long-division problem using the method of long division taught to you at school is formulaic (and easily programmed into a computer), whereas coming up with that long-division method itself is not; there is no obvious formula for doing so and no certainty about the outcome.

Correspondingly, the incubation phase has become understood as the process of broadly scanning and processing the world of possibilities in order to bring up candidate solutions, which are then evaluated, preempting illumination. Thus a dominant theme in the study of creativity resulting from this view has been the study of how people successfully engage in thinking that fans out to include unlikely or apparently unrelated ideas, which are needed if one is to escape the straightforward path' and come up with novel solutions.

The issue is illustrated by a well-known experiment called the candle problem.³⁵ This involves a task in which a person is given a candle, a box of tacks, and a book of matches. The participant is asked to mount the candle onto a pinboard on the wall and then light it, using only the items provided. However, an additional criterion is that the candle must not drip wax onto the ground. Tacking the candle straight to the wall would not work. The solution instead is as follows: empty the box of tacks, tack the box to the wall so that it forms a shelf, place the candle in the shelf and light it.

The solution is obvious in hindsight. However, one of the ingredients of this solution is neatly hidden: many people do not immediately recognize that the box containing the tacks is itself an available resource, assuming it only to be there to hold the tacks, a phenomenon known as *functional fixedness*; the presumed function of the box blocks it from being seen as a shelf. The speed at which people arrive at this solution therefore depends critically on how quickly they overcome this misrepresentation of the box in their formulation of the problem. Experiments have shown how this response factor can be directly influenced through the problem representation: if you present the box and tacks separately, participants solve the problem more quickly. You can even speed up the solution time simply by saying "a box and some tacks" instead of "a box of tacks."

Thus when the environment is not kind enough to reveal the components necessary in solving a problem, we must employ heuristic strategies to help us search for them, epitomized in the now clichéd notion of *thinking outside the box*. Psychologists of creativity have subsequently been interested in what enables us to reformulate problems and broaden our search for solutions. The Torrance test,³⁶ designed to examine an individual's natural capacity for such *divergent thinking*, requires participants to list as many possible answers as they can to questions such as "What can a brick be used for?" You might start with boring answers such as "building a wall," or

“propping a door open,” and end up with ideas further removed from traditional expectations, such as “to draw right angles.” Respondents’ answers are then evaluated according to four criteria:

- *Fluency* The total number of interpretable, meaningful, and relevant ideas generated in response to the stimulus.
- *Flexibility* The number of different categories of relevant responses.
- *Originality* The statistical rarity of the responses.
- *Elaboration* The amount of detail in the responses.

While such tests may tell us something about individuals’ natural propensities for divergent thinking, we can also employ more overt strategies to support creativity. Mathematicians are known to learn heuristic abstractions to help solve problems,³⁷ and organizational strategies such as brainstorming encourage teams to work together to expand a space of possible solutions by actively withholding judgment and encouraging weirder ideas.³⁸

A more general theoretical formulation of exploratory search in creativity is offered by Dean Simonton, who draws on Donald Campbell’s 1960s theory of blind variation and selective retention (BVSr). Simonton defines blind variation as any stage in the creative process where possibilities not immediately to hand are examined for their fit. Given two possibilities, X and Y, if X is more likely to be encountered than Y, but Y is more valuable than X, then, he contends, the discovery of Y as the preferred solution must be arrived at in a blind manner, not a sequential one. Wherever problems have this structure, blind variation will necessarily play a part in the solution. As “problems become more ill-defined, not only must the person rely increasingly on weak heuristic methods that do not promise a proper solution, but it also becomes ever less evident what are the best heuristic methods to try out first. ... The creator may have no other option than to generate-and-test all available heuristic methods without any assurance that even one option will actually work.”³⁹

Simonton warns against the perception that this view reduces creative cognition to mere random trial and error, firstly on the basis that blindness should not be considered synonymous with randomness. A blind search can be highly systematic; to make the point, Simonton uses the example of a radar sweeping in a circular search pattern. The radar searches blindly, meaning that it spins around ready to encounter objects rather than actively homing in on them, but its rotation is systematic and makes for a sensible

heuristic (visit each compass direction equally often at equal intervals). It is not random. Secondly, Simonton emphasizes that in real creative cognition blind search is integrated with other strategies and needn't be considered the primary strategy. "Problem solvers will endeavor to use algorithmic methods whenever doable, but frequently they will have to fall back on a more heuristic search through the various possibilities." He goes further to say that "ideational variations in creativity and discovery take place along a continuum from wholly sighted to entirely blind." He posits that blind variation may be integrated in different ways among creative processes such as mental association and behavioral tinkering. This includes a distinction between internal blind variation, as in the ideation discussed above, and external blind variation, for which he studies Picasso's initial sketches for the painting *Guernica* as an example. In the analysis of Picasso's sketches, the blindness of the search is evidenced in a series of progressions and regressions that do not clearly progress toward the ultimate end point.⁴⁰

In real life, the problems we encounter are situated in different specifics and present opportunities for the use of creative thinking intertwined with other strategies. Real-world "candle problems" present themselves to us in different ways; sometimes the box is hidden, at other times it is in plain sight.

In Robert Sternberg's investment theory of creativity,⁴¹ which I discuss in more detail in the following chapter, "Creativity requires a confluence of six distinct but interrelated resources: intellectual abilities, knowledge, styles of thinking, personality, motivation, and environment." Among intellectual abilities, he includes divergent thinking, but also adds "the analytic skill to recognize which of one's ideas are worth pursuing and which are not, and the practical-contextual skill to know how to persuade others of—to sell other people on—the value of one's ideas." In personality, he includes attitude to risk. In motivation, the focus is on the value of intrinsic over extrinsic forms of motivation, a topic studied closely by Amabile and also central to Mihaly Csikszentmihalyi's concept of flow, a particular state of mind that occurs when one is deeply engaged in a task. Further research on the candle problem showed how much the context in which we encounter the problem matters: counterintuitively, those who are paid rewards to complete the problem do so more slowly than those who are not. Thus pressure to complete a task can hinder exploratory search, supporting the intuitive notion that creativity requires a certain state of mind and a certain freedom.

The preceding passages frame creativity largely in terms of problem solving, and while such topics may be of importance to mathematicians, scientists, and people trying to pin candles to walls, it is not quite so obvious what this has to do with making art and music. But this core idea of creativity as involving a blind search applies to artists equally. Creation of any kind involves something analogous to setting and solving problems. For example, “problem finding” is a key concept in the psychological study of art and design.⁴² Indeed, Simonton suggests that “artists are truly creative, whereas scientists merely make discoveries. ... If Michelangelo had never been born, there would be something else painted on the [Sistine] chapel’s ceiling, but if Newton had never been born, someone else would have come across the universal law of gravitation.”⁴³ This theme will be picked up in depth in chapter 4.

With this overview of the study of creative psychology in mind, we can begin to look at how such themes have been tackled from a computational perspective.

Algorithmic Creativity: Searching through Spaces

Margaret Boden, a philosopher of cognitive science, is most widely known as driving creativity theory in the direction of computing and automation in her 1990 book *The Creative Mind*.⁴⁴ One of Boden’s core themes is the idea that creativity occurs in a set of *conceptual spaces*, abstract spaces of ideas held by individuals or shared among groups of people. She posits that in our interactions with such conceptual spaces we exhibit fundamentally different forms of creativity.

The simplest form of creativity is *exploratory creativity*, whereby we search through a conceptual space, finding new variants on a concept. This may sound relatively formulaic, and therefore not fully satisfying the heuristic nature of creative search, but this depends on what exactly is involved; the exploration may still require smart search heuristics and evaluation. Another of Boden’s creativity types is *combinatorial creativity*, which involves combining concepts together, similar to the way Koestler talks about bisociation. Lastly, Boden proposes that a more sophisticated form of creativity called *transformational creativity* involves actually changing the conceptual spaces themselves, so that new possibilities become available that weren’t previously. In very simple terms, this might be understood as the difference between a musician working in a standard traditional jazz idiom and a musician innovating a new style, such as bebop or free jazz. Or it could describe

the new conceptual spaces introduced by Einstein to replace Newtonian physics.

Boden's combinatorial, exploratory, and transformational forms of creativity describe different modes of human creation with respect to an existing stock of ideas but also suggest models of the different ways in which a computer might do the same. We might explicitly set out to build combinatorial, exploratory, and transformational creativity systems. But while combinatorial and exploratory creativity appear to offer reasonable designs for a computational system, transformational creativity seems problematic, leading some to think of this as a holy grail. Part of the problem lies in the fact that conceptual spaces are not dealt with formally by Boden, and so the idea of exploring them or transforming them remains too abstract to be meaningfully interpreted.

This has been addressed to some extent by Geraint Wiggins, who uses a formal framework, the creative systems framework, to explore Boden's ideas in finer detail.⁴⁵ Wiggins defines a conceptual space (C) in terms of a set of constraints or rules (R) applied to a universe (U) of all possibilities. He then defines the search of that space using a given strategy T . Importantly, for Wiggins's formulation, it is essential to recognize that while we can define a conceptual space, C (for example, the conceptual space of jazz music), we do not have direct and unrestrained access to this space but must search for things in it. While the space in some way *contains* all of jazz music, and we may understand the rules of jazz (or criteria for determining if something is jazz or not) in some capacity, we do not hold all of jazz music in our heads!

Any search strategy T searches the universe U , to find things in the space C . If we sit at a piano to compose a piece of jazz, our process of searching for possibilities is by no means limited to things that conform to jazz rules. We might come across all kinds of things that sound nothing like jazz during our search. Generally, we combine the search strategy T and the rule set R in some way, so that we keep our search constrained to the jazz space, but a critical distinction introduced by Wiggins is this separation between R and T . In addition to R , which constrains the area we're looking in, we also apply some kind of evaluation, distinguishing good from bad, whether inside or outside of the space defined by R . We might find good jazz music, or something that we think of as good, that doesn't conform to expectations about jazz.

Wiggins's formulation leads to slightly different conclusions to Boden, or at least to many of the most common interpretations of Boden's approach. The distinction between the definition of a conceptual space (C , defined by rules R applied to U) and the method for searching it (T) means that searching within the conceptual space might still be hard, requiring a search among possible alternative T s. "It is perfectly possible," Wiggins claims, "for an exploratorily creative system to generate real novelty (because it can produce results which are unpredictable) and in some circumstances for those results to be valued, even if it works by brute enumeration."⁴⁶ Likewise, the transformation of the conceptual space C by a change in the rules R needn't be as revolutionary as the notion of transformational creativity evokes. Transforming R needn't be any more sophisticated an activity than searching through C . Wiggins proposes instead that transformation of the underlying basis for how the rules are described might be considered a more radical form of transformational creativity. Wiggins therefore suggests that transformational creativity is "exploratory creativity at the metalevel." Instead of an exploration through a conceptual space, it is an exploration through different possible conceptual spaces.

But what are conceptual spaces really? The answer remains vague, although a body of work is beginning to grow up around studying conceptual spaces in cognitive science, and tantalizingly is starting to intersect with work in AI systems. Peter Gärdenfors⁴⁷ defines them simply as geometric spaces that are made up of "a collection of one or more quality dimensions," where quality dimensions "represent various 'qualities' of objects." For Gärdenfors, the critical thing about conceptual spaces is that they define perceptions of similarity between objects: "For instance, we can meaningfully claim that the taste of a walnut is closer to the taste of a hazelnut than to the taste of popcorn in the same way as we can say that the color orange is closer to yellow than to blue."

For example, much research has gone into the perception of musical pitches, and we know (as much from the underlying acoustic principles as anything else) that certain notes are closely related to each other harmonically (they sound well together and in some sense sound more similar). Octaves (middle C and the C above) or fifths (middle C and the G above) are harmonically closer than semitones (middle C and the C-sharp above), even though the semitone is closer when you look at it on a piano keyboard or a guitar's frets. Thus when we represent pitches in a conceptual space,

we may want to represent them in a way that places the more harmonically related notes closer together, *not* how they are represented on a piano keyboard. There are various multidimensional representations that researchers have developed to do this, sometimes in the shape of a spring or a spiral, although we are not sure if there is such a thing as a definitive conceptual space for pitches.

We find corroboration from other fields for Wiggins's approach of defining creative domains in terms of sets of constraints. The psychologist of artistic creativity Patricia Stokes proposes a hierarchy of constraints with which creative practitioners navigate their work: "Constraints facilitate problem solving by directing and limiting search for solutions. Thus, they come in pairs. In creative problem solving, one constraint precludes (or limits search among) low-variability, tried-and-true responses. The other simultaneously promotes (or directs search among) high-variability, novel responses."⁴⁸

At the top of Stokes's hierarchy lies a set of *goal* constraints, which are involved in defining different domains of artistic practice: "Goal constraints are overall criteria. Accepted by a domain, they become stylistic conventions, answers to questions like 'is this a Fauve painting?'"⁴⁹ Stokes's view is that individuals choose to adopt these constraints upon entry into a domain of creative work. An individual then develops a series of additional, more individually specific personal constraints as part of their practice, which help them solve the problem of producing quality work according to the goal constraints and quality criteria. These include source constraints, the choice of stylistic elements; task constraints, the materials and methods you use; and subject constraints, the choice of subject content.

This model offers a way of thinking about what might be considered a transformational event (in Boden's sense). For Stokes, "as a new style develops, its goal criterion will be specified, albeit gradually."⁵⁰ Her study of individual painters shows how they adopt goal constraints, which are generally then treated as fixed, and then apply and manipulate other chosen constraints in their solution. The emphasis here is on the fact that only sporadically do painters revisit their goal constraints. But when they do, transformation occurs, in the sense that Boden intended, even if, as Wiggins notes, this transformation need not be particularly earth-shattering (though it may be).

In this way, we begin to perceive a complex of relations between individual creative practitioners and the domains of creativity activity that they

navigate. Individuals are trained in their arts but by necessity engage in generally applicable creative search strategies—at different levels, over their lifetimes and each time they sit down to create—that have parallels with creative work in science, design, and engineering.

From Algorithmic Thinking about Creativity to Implementing Creativity in Machines

The Evolution of Thinking on Machine Creativity

Parallel to research into humans' unique creative capacity, a more abstract discussion of mechanized creativity reaches back into the nineteenth century and beyond. Two important strands feed into this discussion. The first is the idea that animals and other organisms are themselves essentially types of machines, extended to humans if we are content that humans are merely a type of animal. The hearts, lungs, and limbs of animals are easily understood as machinelike elements, with the pumps, engines, and pistons of the industrial revolution inspiring some confidence that what we find in nature, we can also build. Brains and nervous systems have been less palatable as a subject for automation and held out against subjugation to a materialist perspective far longer, although the makers of playful automata, mechanical instruments, and weaving looms were providing food for thought about thinking machines long before the arrival of the computer. (We can follow this thread as far back as ancient Egypt, where Hero of Alexandria made mechanical theatrics, including simple autonomous vehicles.) It is now becoming increasingly hard to maintain a nonmaterialist perspective in the face of discoveries in neuroscience and developments in AI and bionics. As a materialist perspective has gained ground, its proponents have contended that there can no longer be seen any a priori reason to believe that machines can't be built to have capabilities comparable to brains. This is nothing like saying "brains are computers," which has the added and needless complication of having to define what a computer is. It also does not follow that thinking machines will necessarily resemble brains in their style of behavior—we might already anticipate that they will be less forgetful and delusional.

The second strand of thinking concerns the idea that simple mechanical processes can readily create novel forms, thus *be creative*. In particular, this is what Darwin's algorithm of evolution by natural selection has been shown

to do. Human creative minds are, after all, its creative product, or as the cybernetician William Ross Ashby put it, "The brain is merely Nature's latest means of self-preservation."⁵¹ This strand has expanded into theories of ever-growing complexity, emergence, and self-organization, as well as more or less palatable applications of Darwinism to domains beyond biology, including culture. This seems to present a parallel, less orthodox perspective for computational creativity theory because it has little or nothing to do with the creativity of the thinking human with goals and motivations.

Straddling both of these strands is the emergence of the social sciences, and in particular the idea of social or cultural systems as coherent entities in their own right, that shape individual people's views, goals, and behaviors and have their own dynamics that are irreducible to individual human actions. At the turn of the twentieth century, Emile Durkheim⁵² used this principle of irreducibility to help establish the foundations of contemporary social science, proposing that just as the principles of physics do not provide the necessary tools to study chemistry, neither does psychology suffice to study societies, which should be understood as occupying their own stratum of scientific rules and logic.

Contemporaries of Darwin were quick to apply his theories to the evolution of the social and technological world, involving both of these strands tightly intertwined. Documented by George Dyson in detail in his book *Darwin among the Machines*, the maverick thinker and sometimes antagonist of Darwin Samuel Butler presents a vision that reads like a precursor to the prophecies of Kurzweil:

[Is] the animal phase to be the last which life on this globe is to assume? Or shall we conceive that we are living in the first faint dawning of a new one? Of a life which in another ten or twenty million years shall be to us as we to the vegetable? What has been may be again, and although we grant that hardly any mistake would be more puerile than to individualise and animalise the at present existing machines—or to endow them with human sympathies, yet we can see no a priori objection to the gradual development of a mechanical life, though that life shall be so different from ours that it is only by a severe discipline that we can think of it as life at all.⁵³

As computers proper emerged, a vision of thinking machines came clearer into focus. With Jacquard's weaving loom, which could be fed different weave patterns encoded into punched cards, came a tangible instantiation to support the notion of the programmable machine. Soon after, Charles

Babbage's work on the difference engine and analytical engine gave the world its first real glimpse of the idea of a machine that could perform a diverse range of "thinking" tasks. Babbage's collaborator Ada Lovelace, who was instrumental in recognizing this potential, sparked a debate central to AI by considering the proposition that "the Analytical Engine has no pretensions to originate anything. It can do whatever we know how to order it to perform."⁵⁴

Now known as the Lovelace objection, this was most notably contested many years later by the pioneer of modern computing Alan Turing in his famous paper "Computing, Machinery and Intelligence," in which he also introduced the test that bears his name.⁵⁵ Turing's paper ostensibly sets out to argue that computers, with the right algorithms, are in principle capable of humanlike intelligence, with all of its creativity, wit, and awe thrown in.

Against the Lovelace objection, Turing's response combines the dual strands of emerging thought about machine creativity: that brains are not fundamentally different entities from machines, and that simple mechanistic processes can be creative. He first responds to the idea that machines can't produce anything new by drawing a parallel conclusion regarding humans, making what is essentially an argument from the perspective of social creativity. "Who can be certain," he says, "that 'original work' that he has done was not simply the growth of the seed planted in him by teaching, or the effect of following well-known general principles."⁵⁶

As well as his more famous work on thinking machines, Turing was also a major contributor to our second theme, discovering the creative power of simple mechanical processes. Also bearing his name are Turing patterns, sophisticated spatial structures that he modeled mathematically, produced by simple local chemical interactions playing out on a surface. We see these patterns produced chemically in the striped markings on a tiger or zebra, and in the distribution of grasses over a terrain. In Turing's time, the marriage of a mechanistic Darwinian view of the biological world with creative engineering and design was emerging into a research agenda in its own right. Known as cybernetics, the field was defined by one of its founders, Norbert Wiener, as the "scientific study of control and communication in the animal and the machine."⁵⁷ While control and communication were more typically associated with animal behavior, Turing's chemical models were among many instances of work that set out foundational research

questions in what we now call artificial life. How does information propagate around a distributed system in such a way as to produce coordinated high-level effects, and what role does each element play in the creation? Notions such as self-organization and emergence followed and began to feed into a wider appreciation that creativity could be seen as a general phenomenon, not one limited to human thought, and not least in the relationship between human social systems and human individuals.

This line of thinking was on the cusp of entering the mainstream as Turing wrote in “Computing, Machinery and Intelligence.” Darwin’s *Origin of Species*⁵⁸ had not yet been published at the time of Lovelace’s objection, but it carries a clear message about the idea of simple, unintelligent—let alone unintentional—processes being able to originate things. Here was a process through which the complexity and sophistication of natural structures had seemingly progressed with a consistent and stable direction, just as the Second Law of Thermodynamics runs the other way, breaking down order. By Turing’s time, this far-flung notion of mechanical creativity would have started to have become more palatable to the person in the street, though perhaps not as much as to the pioneering computer scientist.

The argument is now well rehearsed and has been honed by Stevan Harnad, who classifies the objection as one of several “granny objections” to the potential for machine intelligence: “The correct reply is that (1) all causal systems are describable by formal rules (this is the equivalent of the Church/Turing Thesis), including ourselves; (2) we know from complexity theory as well as statistical mechanics that the fact that a system’s performance is governed by rules does not mean we can predict everything it does; (3) it is not clear that anyone or anything has ‘originated’ anything new since the Big Bang.”⁵⁹

Harnad also offers a more pithy version of this argument: “What’s a machine?—Till further notice, it is any system that operates according to the causal laws of physics. And what are we?”⁶⁰

Turing goes on to investigate a more specific instance of the Lovelace objection, that machines can’t do anything that would surprise us. To this his response is more impressionistic: “Machines take me by surprise with great frequency. This is largely because I do not do sufficient calculation to decide what to expect them to do, or rather because, although I do a calculation, I do it in a hurried, slipshod fashion, taking risks.”

In recent work on evaluating computational creativity, Selmer Bringsjord, Paul Bello, and David Ferruci⁶¹ take exception to this second line of argument. They view the characterization of the observer's misunderstanding or of programming error as being an instance of "surprising output" as a straw man, to which they argue that being surprising is not a sufficient condition for originating something. They give an example: imagine a malfunctioning robot in a Toyota Camry factory sticks a spare tire to the rear bumper of a car. The designers see this and are inspired to develop a new design, blending a sedan with a sports utility vehicle. This, the authors argue, is the kind of situation Turing is thinking of as an example where the machine could be argued to originate something, by virtue of producing a surprising outcome that would not otherwise have been produced. But, they continue, this is not the machine's originality, and besides it is a trivial case. In chapter 3, I will return to this example to make the case that while Bringsjord et al.'s analysis seems entirely reasonable, this distinction between their view and Turing's is of critical importance. In their dismissal of the creativity of the errant robot, they place too much weight on the individual and their intentions, and require that the robot is either a fully blown creative agent or is not. A distributed view of creativity would allow, instead, to understand robot error as having creative capacity as part of a larger system.

Bringsjord et al. propose what they claim is a more useful formulation of the Lovelace objection in the form of a "Lovelace test" for computational creativity. To pass the test, the computer must produce an output that is not a fluke (it is repeatable) and cannot be explained by the programmer (or someone equivalent) "by appeal to [the system's] architecture, knowledge-base and core-functions." Their appeal to explainability is commensurate with Simonton's appeal to include non-obviousness as a required trait for creativity.

With this in mind, it is interesting to return to the question of public perceptions of machine creativity, as discussed in chapter 1. Computational creativity researchers Martin Mumford and Dan Ventura⁶² have more recently surveyed people's perceptions about how creative computers might one day become, and asked their respondents to suggest what evidence would make them feel the computer was being creative. Their results suggest that the demonstration of *autonomy* by the computer, showing that it did what it did independent of the programmer, is a fundamental concern, especially

to those who are more skeptical of computers' creative potential. They do not specifically identify Bringsjord et al.'s criterion of unexplainability, however, and the specification of what might actually constitute autonomy in the terms of the respondents remains an open problem. Mumford and Ventura's study also indicates that in 2015 people were generally quite willing to accept the possibility of machine creativity, whether this is because they had witnessed firsthand the steady surge of progress in machine intelligence, or because we have come to appreciate a much more mechanical, social, or distributed view of what creativity is.⁶³

The Imitation Game

Mention of Alan Turing in the context of creativity wouldn't be complete without reference to Turing's own imitation game, which in some sense stands aside from these discussions of human and machine potential.

In his version of the imitation game,⁶⁴ based on a popular parlor game, two people and one computer communicate only via written text, rendered on a screen. One human, the judge, has the task of telling which of their correspondents is the human and which is the computer. Both the human candidate and the computer program are hidden from the judge. Both must convince the judge that they are the true human, through written correspondence only. If the judge cannot reliably identify the true human, then the computer has passed the test.

Turing's test has come to mean many things for many people. For several key commentators, it was never intended as anything more than a thought experiment aimed at leveling our underlying assumptions and highlighting the fallibility of privileging human intelligence without good cause, to dispel the idea that there is anything more to human thought than that which we can observe empirically. It does so simply by illustrating what empirical evidence of intelligence would look like. If the machine *appears to* appreciate the beauty of a sonnet, or to come up with something original and non-obvious, and through repeated testing seems no less capable of doing so than a human, then what reason would remain for denying the system intelligence? We should certainly suspect some kind of trickery, and vow to keep testing. Indeed we should maintain a cynical stance, since nowadays systems are being built explicitly to try and score well on the test, as we will see below. But that is the point: the Turing test admits as much scrutiny as

we care to give it. In Robert French's words, "Turing's central claim is that there would be no reason to deny intelligence to a machine that could flawlessly imitate a human's unrestricted conversation."⁶⁵

Another way to think of Turing's test is as a kind of operational definition of intelligence—*that which can be observed*, as opposed to a tautological and essentialist view in which being human is a prerequisite of having human-like intelligence. As we have seen, computational approaches to creativity appeal to the same basic reasoning. The Turing test certainly does offer an operational definition of intelligence, but several critics, Blay Whitby, in particular, have argued that such a view has been counterproductive to AI research: "Much AI work has been distracted by a view of intelligence as a holistic phenomenon, demonstrated only by human beings, and only to be studied by the direct imitation of human beings."⁶⁶ Whitby argues that aiming for human intelligence in this holistic way is of little use if we do not have good theories of intelligence that can be explained in terms that do not appeal to humans. He grounds this point in an analogy with the invention of flying machines: "Direct imitation of natural flight proved a relatively fruitless avenue of research. ... Working aircraft were developed by achieving greater understanding of the principles of aerodynamics."⁶⁷

More vexing to such critics is the emergence of a high-profile obsession among certain communities with winning the Turing test. In 2014, in one of several regular annual Turing tests, a chatbot named Eugene Goostman, taking the persona of a thirteen-year-old boy, caused a third of its judges to guess that it was a human. In doing so, it purportedly passed some variant of the test, and some consequently declared 2014 a landmark year for AI. As more pessimistic commentators were quick to point out, this exercise only served to highlight the potential fickleness of specific implementations of the test. Consider a short transcript of one of the interrogations by a judge of Eugene:

Judge: what is your gender

Eugene: I'm a young boy, if you care to know. And not very ugly, by the way!

Judge: what are you aware of at this moment

Eugene: I would rather not talk about it if you don't mind. Could you tell me what are you? I mean your profession.

Judge: how old are you

Eugene: I'm a little boy 13 years old.

Judge: what are your qualifications

Eugene: My qualifications? I'm not gonna tell you. Oooh. Anything else?⁶⁸

There is some reason to find Eugene's AI impressive. The system can hold an engaging conversation, even taking the upper hand in steering it, and there should be no doubt that the technical challenges involved in getting this smoothness of response are nontrivial. But as to whether conversations such as these should be considered informative enough to adjudicate on human-comparable intelligence is another matter. Consider AI veteran Marvin Minsky's disparaging response: "Ask the program if you can push a car with a string. And, if not, then, why not?"⁶⁹ Why, indeed, did the judge not ask this or some similar question? Such a question requires a model of the world and all of the concepts we use to represent it, as well as how those concepts might interact in unforeseen ways. *Can you carry water in your pocket? Should a former terrorist who has served time in prison be allowed a job as a security guard?* Any thirteen-year-old can make good work of the conceptual material involved therein, yet it is a devilishly hard challenge for AI. The judge only gets out what they put in, and they should certainly put questions such as these in if their goal is to pick out the nonhuman. It should not be hard to ask questions that would reveal a depth of understanding beyond Eugene's, who we might guess from the above would cockily shift topic rather than provide a convincing answer.

Make no mistake, AI is rapidly advancing in its ability to process complex linguistic operations, leading to advances in comprehension. For example, Radford et al. were able to show that a system trained to continue a word sequence, much like the musical continuation task discussed in chapter 1, also performed very well at answering trivia questions, since answering a question coherently *is* basically a continuation of the question sequence. This applied to other specific language tasks such as translation: ask the computer to "translate *bonjour* into English" and, again, the response "hello" is a natural continuation. The language sequences that this model creates demonstrate deep contextual awareness even though it is just a model of word prediction.

A longer continuation of a fictional story is provided by Radford et al. (see box "Text Generated Using the GPT-2 Language Model by OpenAI"). We can read this text and be impressed, without needing to be tricked. Indeed the authors are able to find many ways to ways to measure the performance of their system's output to provide rich detail.

Text Generated Using the GPT-2 Language Model by OpenAI

Context (human-written): In a shocking finding, scientists discovered a herd of unicorns living in a remote, previously unexplored valley in the Andes Mountains. Even more surprising to the researchers was the fact that the unicorns spoke perfect English.

System output (continuation): The scientists named the population after their distinctive horn, Ovid's Unicorn. These four-horned, silver-white unicorns were previously unknown to science.

Now, after almost two centuries, the mystery of what sparked this odd phenomenon is finally solved.

Dr. Jorge Pérez, an evolutionary biologist from the University of La Paz, and several companions were exploring the Andes Mountains when they found a small valley with no other animals or humans. Pérez noticed that the valley had what appeared to be a natural fountain, surrounded by two peaks of rock and silver snow.

Pérez and the others then ventured further into the valley. "By the time we reached the top of one peak, the water looked blue, with some crystals on top," said Pérez.

Pérez and his friends were astonished to see the unicorn herd. These creatures could be seen from the air without having to move too much to see them—they were so close they could touch their horns.

While examining these bizarre creatures the scientists discovered that the creatures also spoke some fairly regular English. Pérez stated, "We can see, for example, that they have a common 'language,' something like a dialect or dialectic." Dr. Pérez believes that the unicorns may have originated in Argentina, where the animals were believed to be descendants of a lost race of people who lived there before the arrival of humans in those parts of South America.

While their origins are still unclear, some believe that perhaps the creatures were created when a human and a unicorn met each other in a time before human civilization. According to Pérez, "In South America, such incidents seem to be quite common." However, Pérez also pointed out that it is likely that the only way of knowing for sure if unicorns are indeed the descendants of a lost alien race is through DNA. "But they seem to be able to communicate in English quite well, which I believe is a sign of evolution, or at least a change in social organization," said the scientist.⁷⁰

I agree with David Auerbach when he says:

Turing was trying to establish a beachhead for conceiving of intelligence in a way that wasn't restricted to humans. Now that we can conceive of it, the question shifts not to whether a computer can trick people in one particular way, but to the vast varieties of ways in which it can display intelligence, human or not.⁷¹

Put simply, the value of such interaction tests is in the detailed analysis, not in the binary outcome of pass or fail, which if Eugene's example is anything to go by can be quite deceptive. Eugene should rightly have been celebrated for its achievement by highlighting its true capacities, such as the ability to engage in a playful linguistic interaction. This includes some sense of human personality, even perhaps some essence of social intelligence: the ability to tease or please or steer the interaction in a targeted manner. All that considered, the Turing test win is itself something of an aside, an arbitrary landmark.

What about an art Turing test? One of the most important and direct ways we can appraise any automated art-generation system is by asking people to evaluate its outputs, and it is natural that the Turing test comes to mind as a potential model for how to go about this. In the previous chapter I mentioned an example where the *Guardian* newspaper asked readers to take such a test for Iamus, a music-generation system, and try to guess which compositions were made by a real human and which by a computer. The pieces sound impressive, and it may be hard for a non-expert to make the distinction. Auerbach's response applies here perfectly well, and I could paraphrase it more succinctly as follows: the question is not "can it do it" but "what can it do?" Evaluations of *human-likeness* for individual musical pieces should be just one sort of data gathered about the system.

French also points out the problem of being able to create Turing test variants that, rather than being open conversations, take the form of "restricted domain" tests. Art-based tests would certainly fit this description. French draws attention to "the virtual impossibility of clearly defining what does and does not count as being part of a particular real-world domain."⁷² I return to this discussion in chapter 7, which looks at methods for evaluating computationally creative systems and considers in greater detail some of the attempts to run blind tests on the output of computationally creative systems.

Generative and Adaptive Creativity

The central idea put forward in this chapter is that creativity often involves blind search to discover things that would not have been discovered via non-search-based processes, algorithmic processes that have a clear progression to an outcome. Humans possess exemplary cognitive abilities to deploy creative strategies, which are the main focus of study of creativity researchers and the main source of inspiration for computational creativity systems. A key part of that cognitive machinery is the performance of heuristic strategies, both consciously and subconsciously, along with cognitive capacities that support such a process, in particular the ability to incubate concepts subconsciously. However, although far less formulaic than so-called algorithmic processes—those such as long division where there is a clear path to the solution—there is no good reason to believe that a machine can't implement such processes, while humans can. This is seen as an essentialist view, exemplified by the Lovelace objection that Turing dismisses, giving a simple "yes" to the question "Can computers be creative?" While I think this answer is entirely satisfying, the question of the social dimension remains poorly addressed, and it is to this that we turn properly in the following chapter.

I conclude this chapter by drawing out a key distinction that I feel clarifies an area of persistent confusion. It is broadly true that we intentional humans set out to create things with goals and an evaluation framework, and our creative machinery is constructed in such a way as to seek out things that we perceive to have value. Yet this is also a problematic formulation: I wish to claim that in addition, much that is created, artistically and in other domains, demonstrates a significantly more complicated relation between the intention of the creator (or creators) and the outcome, and its respective value and impact. In this more complicated relation, things may be created by accident or are adapted in ways not previously foreseen (not part of the original creator's intention) and are vaunted and valued in ways that are determined by their cultural interpretation and also beyond the creator's relation to the work.

One way to approach this is to allow that in essence creativity means the creation of new things independently of how they are valued or what the intention of the creator was, if any. One reason for this is to be able to admit evolutionary processes into the canon of processes that we call creative, not

simply to offer a more inclusive concept of creativity, but because it is also a relevant aspect of human creativity. When viewed at the collective cultural level, where movements emerge and great works of art are defined, human activity can be understood as much in terms of a blind evolutionary process that *emerges* out of individual action, as in terms of a directed intentional processes led by individuals. As mentioned in this chapter, key to avoiding tautological treatment of Darwinian processes, we must not view the creation of elephant trunks as being “in order to” benefit elephants, but as part of a process of emergence in which elephants and their trunks came about. This is also why it may be wrong to dismiss Bringsjord et al.’s errant robot in the car factory as a noncreative force, when seen as part of a wider distributed process.

A more universal view of creativity over different timescales and in different configurations might start from the central argument of many proponents of a social view of creativity: that the understanding of individual creative cognition must be complemented by an understanding of the social dynamics involved in the various forms of human creative production. Social factors *act* upon individuals as much as social phenomena result from individual actions. This is a different idea of creativity from that associated with the human cognitive capacity to be creative, rather than a competing hypothesis. On a universal epochal scale we witness value-free blind creation of form, equivalent to Darwinian evolution, if not identical in detail. While we can’t talk about this in the same language as the psychology of human creativity—and simplistic metaphors should certainly be avoided—we can observe, with Simonton, that a form of blind variation and selective retention is common between them. In our daily lives, meanwhile, we witness value-laden purposeful creativity employing powerful dedicated neural circuitry that is unique to the human species. Now zoom in to the neural processes where concepts are being generated, and at some point we again see a system that is incubating or brainstorming, looking far and wide without being steered by the constraint of evaluation. Zoom back out, and at the social scale we see both purposeful creativity, as when teams of individuals come together with a common goal, and less purposeful creativity, as in the emergence of subcultures that arise from the work of individual actors with disparate goals but that have no overall purpose themselves.

In short, creative outcomes that are not directly grounded in goals and values can be found everywhere—interlaced with creativity that is grounded

in value. In earlier work,⁷³ I have attempted to distinguish this as a different category of creativity, called *generative creativity* by association with the notion of a generative process. It blindly and mechanically generates new stuff, which comes about in a process of emergence. The more recognizably human creativity-with-value I have, by contrast, called *adaptive creativity*. In this form, there is always an agent for which the creative outcome serves a purpose. These categories were an attempt to help more clearly locate the former kind of emergent process in studies in computational creativity and, more generally, to explore different relationships between creativity and value, such as how an influential artist may have greater freedom to be nonconformant than a newcomer.

For the purpose of this book, I treat these terms as a useful way to probe this distinction, but not necessarily as an essential framework. Nevertheless, creativity might more productively be understood not as the production of novel and valuable things but as the *situated production of novelty*. Creative systems are by definition responsible for producing novel outcomes, but the way in which they do so can be understood through the way in which they are situated in a given context. In adaptively creative situations, the creative system is an agent, situated in a context in which it has motivations and tangible outcomes, and its creativity does produce things that can be said to have value. In generatively creative situations, these conditions may not be so apparent. The creative system may be a distributed process involving multiple competing agents with conflicting goals, or even a blind mechanical process that has no reason to its creative productivity. These ideas will continue through the following chapter, which further develops this socially oriented treatment of creativity.

© 2021 Massachusetts Institute of Technology

This work is subject to a Creative Commons CC-BY-NC-ND license.

Subject to such license, all rights are reserved.



The open access edition of this book was made possible by generous funding from University of New South Wales.

This book was set in Stone Serif and Stone Sans by Westchester Publishing Services.

Library of Congress Cataloging-in-Publication Data

Names: Bown, Ollie, author.

Title: Beyond the creative species : making machines that make art and music /
Oliver Bown.

Description: Cambridge, Massachusetts : The MIT Press, [2021] |

Includes bibliographical references and index.

Identifiers: LCCN 2020017931 | ISBN 9780262045018 (hardcover)

Subjects: LCSH: Technology and the arts. | Arts—Technological innovations.

Classification: LCC NX180.T4 B68 2021 | DDC 700.1/05—dc23

LC record available at <https://lcn.loc.gov/2020017931>

10 9 8 7 6 5 4 3 2 1