ABSTRACT

As advances are made in artificial intelligence technology, the definition of 'artist' demands reevaluation. At first glance, artistic creativity would seem to be impossible to endow in a computer, given artistry's apparent reliance on such deep behaviors as use of metaphor, analogy, and composition—abilities that today's computers do not have. While computers have previously been used as tools for creative artists, what sort of capability would a machine have to demonstrate to be considered an artist in its own right? To address this question, I consider the Ability Hypothesis from Lewis and Nemirow, which postulates a difference between knowledge that is propositional (knowledge-that) and experiential (knowledge-how). I show that the distinction between propositional and experiential types of knowledge finds a physical foothold in the evolved architecture of the human nervous system, but is not similarly enforced in the designed architecture of today's computers. Human art-makers must hone their artistic knowledge through experience to become expert artists, because the nervous system cannot be directly programmed. Computers, on the other hand, may exhibit instant expertise and ability through access to new propositional knowledge alone. Therefore, the requirement of vast experiential knowledge that we hold for human artists should not apply to computers. I apply this analysis to the key case study of Harold Cohen's AARON painting machine, arguing that AARON counts as an expert, artistic system. I conclude with a comparison to machine learning algorithms that emulate neural architecture. This research broaches new ground in the philosophies of artificial intelligence and neuroscience.

"He who learned by Finding Out
has sevenfold
The Skill of him who learned by
Being Told" – Arthur Guiterman, *A Poet's Proverbs,* 1924

1 Introduction: Art Without An Artist

As advances continue to be made in artificial intelligence technology, the question arises, "could a computer ever make art as humans do?" To answer it, one would have to be sure of a definition of what it means to "make art" "as humans do", along with the differences in kind between computer and human that prohibit the former and enable the latter. This line of inquiry is similar to that made by Alan Turing in his famous 1950 paper, which dealt with the analogously thorny question "Can machines think?[1]" He proposes an "imitation game" later dubbed "The Turing Test" as a way to test a machine's thinking ability. The underlying implication of this test (to simplify things quite a bit) is that if a computer can use language in such a way that a human interacting with the computer believes they are interacting with another human, then one might say that computer thinks.

Could this imitation technique map to the question of computer-made art? In other words, if a computer produced a work that convinced the listener or viewer to think that another human had made it, is this enough to conclude that the computer made art? This has been addressed in the literature by a number of scholars, and the general consensus is no – the passing of an artistic Turing Test is not a strong enough measure to conclude that the computer in question deserves to be called an artist.

In Western tradition, it has not simply been the production of an art object that makes one an artist, as can be seen from the examples enumerated in Howard Becker's *Art as Collective Action* article. Marcel Duchamp signed a snow shovel as one of his "readymades" being credited as the artist, while whomever actually made the snow shovel gets no artistic credit. Another example would be the artisans and architects who worked to realize Gaudí's vision and plans for La Sagrada Familia

---

[1] Turing, 1950

even after his death. While their efforts and choices resulted in what is undoubtedly a work of art, they are not credited as the artists behind the work. Following that same line of thinking, why should a computer get any credit at all, or even the title "artist" regarding the work it has been directed to do by its human creator? On the other hand, there is the idea of a commission – while the Medici family certainly requested and funded the production of Renaissance art, they are not considered the artists in those cases. What qualities or behaviors are essential in determining who or what is the artist, somewhere between commissioner and direction-follower, for a given work of art, and who or what are the "support personnel" – those who executed the vision of the artist – in Becker's terms?

The first requirement for artistry, intuitively, is some involvement in the production of art. To address this issue any further, we must define what we mean by the term 'art'. I define art to be any work (material or performance) for which apparent choices (of color, tone, figure, perspective, content, etc.) arouse a strong emotional response in its perceivers. The more perceivers the piece moves to feel powerful feeling (even if they aren't all the same emotions), the better the work of art. Notice that only the perception of intentionality on the part of the viewer is required. This means that art, under this definition, does not require an artist. If a group of very lucky monkeys did bang out Hamlet on a typewriter out of purely random monkey activity, we would certainly enjoy Hamlet as a fantastic play, but hardly credit those monkeys with authorship. Could we even credit Shakespeare? No one at all would receive any credit at all if the monkeys in question somehow typed an interesting, moving work that had never been written before. In essence, the designation of "art" employed here can be handed out based on a Turing Test style of reasoning. If a work is so convincing that a human believably crafted it, then it should deserve the title "art;" the broader its emotional appeal, the better art it is.

If there were an actor on the other end, intentionally making those emotionally salient choices, then they certainly seem to deserve the title of artist. But intentionality is tricky – how could a computer ever be intentional when it always follows the directions handed down by a human programmer? In his 1950 *Mind* paper, Turing considers an objection raised by Ada Lovelace – that a machine could never be original, because it does what it was programmed to do. The Lovelace Objection, if it holds, relegates computers permanently as "support personnel." Turing responds, quoting the Bible, that there is "nothing new under the sun." Humans, he implies, do not magically cook up creative original new ideas. Today, the results of research into human creativity seem to back up Turing's response. Modern psychological theory, particularly as presented by Margaret Boden, emphasizes the computational processes that likely are performed behind the scenes when an artist, designer, or scientist conceive of new ideas. Boden offers three types of such processes – combinational, exploratory, and transformational.[2] Given the mechanistic nature of these processes, it seems that creativity is not as deep as was once thought – that perhaps a non-conscious computer would be able to perform such calculations and yield creative new outputs. Is this all there is to artistry?

There are key differences to consider between nervous systems and computers. Nervous systems are necessarily embodied, perceiving the world through sensations and affecting the world through muscles. Computers can exist as solely information processing machines, taking in information in binary form and executing transformations according to directions written in their software, to yield a deterministic output. This paper will examine the implications of these differences with regards to the question of artistry.

First, I will introduce Frank Jackson's Knowledge Argument and Lewis and Nemirow's reply of the Ability Hypothesis. I will then compare the manifestation of the experiential/factual knowledge distinction in humans and machines. This will set up my discussion of the connection between know-how, or experiential knowledge, and artistry, pointing out that humans package factual knowledge into intuition, which is used in the application of experience. Expert human artists use the efficiency of their intuition to focus on higher-level decisions, yielding more emotionally salient works. For example, instead of getting caught up in how to shade this tree or how to paint in the wind blowing or how to resolve this chord, the experienced artist has, for the most part, automated such a task. She can instead devote mental resources toward exploring and testing the search space of possible

---

[2] Boden, 2004.

choices. With her technical ability down pat, she also possesses a broader range of such possibilities and a better understanding/imagination of this space. Finally, I will discuss this theory as it applies to AARON, an impressive art-making algorithm from the mind of the late artist-turned-programmer, Harold Cohen. Under the knowledge-how of artistry requirement, is there an argument to be made that AARON is an artist in its own right? Does it matter that the entirety of AARON's instant expertise comes from Harold Cohen's learned expertise?

I hope to show that while artistic experiential knowledge is indeed necessary (while making no claims regarding sufficiency) for artistry in humans, artistic experiential knowledge should not be required for artistry in computers, because the distinction between knowledge-that and knowledge-how falls apart outside of the functional architecture of the human nervous system. What must then be considered is whether or not there can be ownership of experiential knowledge-how with regard to which agent gathered that experiential knowledge, how it was fed to the computer, and implications for who, then, is the artist.

2 Deconstructing Mary's Room

Imagine a woman named Mary, sitting in a drab room. Everything in this room is a shade of gray – there are no colors, anywhere. Mary is a brilliant color scientist. She has been studying color since she was born. From aesthetics to psychology to art history, even the precise neuroscience of how color impacts the brain – Mary understands it all. However, she lives in this black and white room and has never actually seen any color before in her life, even while she knows every physical fact that one could possibly know on the subject.

One day, her captors open the door to her room, and she steps outside. A world of color greets her. Mary always knew the phrase – "the sky is blue" – she had learned it at a young age. She understood what the sky was, having seen many pictures of the outside world, and she had studied the color blue. Even so – when Mary looks up, finally outside of her colorless confines, and sees the very blueness of the sky, is Mary surprised at some sort of new information? Does Mary learn something she hadn't known before, even in all her understanding of the physical facts of color?

The thought experiment of Mary's Room is from Frank Jackson's 1982 essay entitled *Epiphenomenal Qualia*.[3] At first glance, it seems to be a mash-up of John Searle's Chinese Room experiment and Plato's Allegory of the Cave. We have a character that lives in a constructed reality, but does not truly grasp the nature of her reality until she steps out into the light, á la Plato. She is also stuck in a room with much information available to her (all the physical facts regarding color) but not the one key ability that the information points to (the ability to recognize color). Philosopher John Searle proposes a similar thought experiment, which he uses to question a machine's ability to understand language beyond symbol manipulation.[4] Mary in Jackson's essay is akin to Searle's man with the Chinese dictionary/reply-guide and his inability to actually speak Chinese. The question that Jackson poses that seems to break free of these analogous allegories is whether or not Mary *learns something new* or *gains new knowledge* when she leaves her room and sees blue with her own eyes for the first time.

The argument Jackson is making is a refutation of Physicalism – the belief that every phenomenon in the mind is a physical one and therefore can be described by a set of physical facts. Jackson hoped to show that Mary, who knows everything about color, *still learns something new* when she goes and *experiences* color for the first time. This intense color experience is called a "quale" (plural, "qualia") in philosophy of mind – qualia are the mental correlates of the sensory experiences we have as embodied humans. Jackson argues qualia must have a non-physical component since Mary knows all there is to physically know about blue, but still has an experience she could not have imagined otherwise when she steps outside. This means Mary learns something new, something she did not know in the room, even when she knew every fact about color. Therefore there must be an aspect of color experience that cannot be captured by physical description. Hence, Physicalism is false.

---

[3] Jackson, 1982.

[4] Searle, 1980.

This argument makes no direct reference to computers or artificial intelligence but pertains directly to the question of the artificial creation of a mind, as I have read Jackson's work (Daniel Dennett employs a similar line of reasoning in his essay *"Epiphenomenal" Qualia*, responding to Jackson). If we accept the conclusion of Jackson's argument, then there is more to color and the human experience of color than facts that describe its physical nature. So, even if we gave a computer a sensory apparatus by which it formed precepts of its world, and so had experiences, all of those experiences would be rendered in a pattern of ones and zeroes, as is the binary nature of how digital computers store information. The computer would be unable to be sensitive to non-physically describable aspects of color, such as the supposed epiphenomenal nature that must accompany the physical facts (by the Knowledge Argument). This is because the computer can only translate physical phenomena into ones and zeroes. There cannot be the physical chain of causation necessary for a computer to register a stimulus if the stimulus is not physical. This seems to present a problem for computational intelligence - how could this qualia-sensing ability be instantiated in a computer if it is not a physically describable phenomenon?

Since Jackson made his controversial claim, known as the Knowledge Argument, many thinkers in the field of philosophy of mind have come up with replies that expose its flaws. Frank Jackson himself abandoned the idea some years later, coming to agree with Nemirow and Lewis on their reply, called the Ability Hypothesis (explored in detail later in this paper), an argument that Mary does not learn a new fact but rather gains an ability when she steps outside. The particular method of these responses include attacking the premise, that Mary could ever know everything about color and what this would mean, and the argument itself, which entails that if Mary learns something new when she sees blue for the first time outside, Physicalism is defeated. The paragraphs that follow are my own brief analysis.

Strictly speaking, when Mary steps out of the room, she does have an entirely new experience. No matter how much factual information she has stored in her memory about the color blue, Mary, the human color scientist stuck in the black and white room, has never made use of the specialized cone cells in her retina that are tuned to react to photons with wavelengths in the blue spectrum. She's never received this pattern of neurological input to her visual cortex. Let's assume those cells are intact and functioning even after all these years, and that she has the appropriate cells in cortex ready to interpret color. The pattern in her brain will be novel and she will have a new experience. Mary knows this of course, as she understands all the effects of the color blue on the human mind. The one piece missing – the piece of new knowledge that Mary learns - is the connection of being *in* the new brain state of blueness (activating the pathway that is initiated when and only when we see the color blue) with the knowledge that *this is* what blueness is.

When she steps out of the room she will be flooded with new color stimuli. It might take a second for her to attach the visual sensation she receives from looking at the sky as "blue." Once she syncs this up with her understanding that a) the sky has *this* element/feeling and b) the sky is known to be blue then c) this is blue. This unification of concepts with perceptive experience is thought to occur in the brain as a strengthening of synapses, or connections, between the different neurons that fire together, making it easier for one pattern to trigger another.[5] After connecting this blue feeling and her concept of the color blue, when shown blue light the 'blue' concept will be more likely to be triggered. While this is a new experience causing a novel brain state of blueness perception, there is only a simple mapping of never-been-had-before experience to strongly understood concept – there is no room for non-physical "epiphenomenal qualia."

Mary can certainly try to imagine color, even in the black and white room. This is likely to be, in a physical sense, neural signals between her prefrontal cortex and her visual cortex, simulating reality in her mind, concocting new brain states mostly independent of her current stimuli (all black and white)[6]. The odds that she correctly simulates the blueness brain state in her visual cortex are nigh impossible. Can you imagine a color you've never seen, or what "ultraviolet" might look like? While Mary has conscious control over the processes that give rise to imagination –

---

[5] Malenka and Bear, 2004.

[6] Rich et. al., 2006.

i.e., the mind's eye, she does not, however, have the conscious power to unilaterally set off her blue-tuned cone cells in her retina. Information does not flow from the brain to the retina; it is only sent from the retina to the brain. Mary cannot possibly have the full blueness experience until some blue light finds its way into her eye.

Remember, Mary has such a complete understanding that she has memorized and could describe the position and state of every relevant neuron in her brain after she sees some blue light. But as she can only describe the brain state, she is a level above actually experiencing this state. In conclusion - Mary does learn something new when she walks outside, but this poses no threat to Physicalism. This is because the language of factual propositions does not have the same effect on brain structures as do sensory experiences as a result of the architecture of the flow of information from peripheral structures to the brain and not the other way around in sensory systems[7]. My response to the Knowledge Argument is related to a well-known response to the Knowledge Argument that I characterize in the following pages.

3 The Ability Hypothesis and Its Limits

David Lewis' essay *What Experience Teaches* presents a strong and well-known reply to Jackson's Knowledge Argument. This line of thinking is oft referred to as 'The Ability Hypothesis' – it claims that while Mary did learn something new, it was not a new type of fact. Instead, Mary learns a new ability – to recognize the color blue. Lewis offers a succinct version of his central point using the taste of Vegemite as an example, instead of the blueness of blue:

> *If you have a new experience, you gain abilities to remember and to imagine. After you taste Vegemite, and you learn what it's like, you can afterward remember the experience you had. By remembering how it once was, you can afterward imagine such an experience… Further, you gain an ability to recognize the same experience if it comes again. If you taste Vegemite on another day, you will probably know that you have met the taste once before… Here, the ability you gain is an ability to gain information if given other information. Nevertheless, the information gained is not phenomenal, and the ability to gain information is not the same thing as information itself.*

---

[7] The brain pieces together a percept from the information that the retina receives, but the retina has no access to this percept or any propositions the brain makes based on this percept.

In other words, later in the essay:

> *The abilities to remember and imagine and recognize are abilities you cannot gain (unless by super-neurosurgery, or by magic) except by tasting Vegemite and learning what it's like. You can't get them by taking lessons on the physics or the parapsychology of the experience… The ability hypothesis says that knowing what an experience is like just* is *the possession of these abilities to remember, imagine, and recognize. It isn't the possession of any kind of information, ordinary or peculiar… It isn't knowing-that. It's knowing-how. Therefore it should be no surprise that lessons won't teach you what an experience is like. Lessons impart information, ability is something else. Knowledge-that does not equal knowledge-how.*[8]

By now I'm sure that you can see why it's called the Ability Hypothesis – it says that ability of blue-recognition is not factual knowledge, and therefore, Physicalism is unscathed by the Knowledge Argument, in that physical, factual knowledge may still describe everything, but having physical knowledge is different than having an ability to use it. Mary can be as surprised as she likes by the blueness of blue, but it is not because physics cannot capture the epiphenomenal nature of blue, but rather that Mary lacked the ability to recognize blue.

The distinction between knowledge-how and knowledge-that as these concepts are used in the Ability Hypothesis, comes from the British philosopher Gilbert Ryle in his appropriately named essay *Knowing How and Knowing That*, written in 1945.[9] Ryle employs examples of learning chess and logical theory to show not only how knowledge-that is not equal to knowledge-how, but that knowledge-how cannot be reduced to knowledge-that. Here is the essence of the chess example: "What truths does the clever chess-player know which would be news to his stupid opponent? …We can imagine a clever player generously imparting to his stupid opponent so many rules, tactical maxims, 'wrinkles', etc.… Yet he might still play chess stupidly, that is, be unable intelligently to apply the maxims, etc."[10] Ryle seems to be saying here that good chess playing requires certain

---

[8] Lewis, 1988

[9] Ryle, 1945

[10] Ryle, 1945.

know-how, a strong intuition, and is not composed by a simple set of rules to be followed.

This point is critical to the Ability Hypothesis' point against Jackson's Knowledge Argument. Physical facts are knowledge-that. In the thought experiment, Mary has all of the physical facts – she has all of the knowledge-that she can possibly have about color, and yet, because she lacks the knowledge-how, or simply the ability to recognize, she is still surprised to see that this special quality of the sky is exactly the blueness she had learned so much about.

The central point I would like to make here is that the knowledge-how/knowledge-that distinction finds a basis in the functional architecture of human nervous systems[11]. This distinction is constructed, and does not necessarily hold for other possible minds, such as computers. Maybe you smiled at the irony of Ryle's dated chess example from 1945. We now know that it is entirely possible to play excellent chess using only rules and tactical maxims. These are all you need to beat even the best of human intuition! This was shown in 1997 when the IBM computer Deep Blue beat the reigning human chess champion, Gary Kasparov. Deep Blue had no honed intuition, no know-how – it was not operating on its own learned chess experience. The algorithm had not developed intuition out of its own playing. Computers play incredible games of chess by performing massive searches of data structures called trees, examining the potential board state hundreds of moves in the future, and use this information to select what they evaluate to be the best outcome, based on the rules and maxims they are programmed to value. These values are encoded in a programming language by a human programmer. Other software, called a 'compiler', reads these instructions by translating them into lower level machine language, and from there into binary code. This is definitively knowledge-that, under Ryle's paradigm. It is factual knowledge, yielded in a logical set of ones and zeroes. Yet, the computer seems to have acquired the skill we would call knowledge-how, in that it demonstrated incredible chess ability by beating the world chess champion. This approach simply does not work for humans, because we cannot perform massive searches as fast as a computer can.

---

[11] That is, in addition to the limits of neurosurgery today as compared to modern software capabilities to formalize human ability.

We solve the chess problem in an entirely different way – by building intuition through experience - learning.

There is not a consensus in the literature regarding the reducibility of 'knowledge-how' to 'knowledge-that'. Stanley and Williamson (*Know How,* 2001) offer a popular reply *supporting* reducibility to the points made by Ryle and Lewis, who argue that ability *is* distinct from factual knowledge (and not reducible). Alva Noë, in his 2005 paper *Against Intellectualism*,[12] counters by attacking the Stanley and Williamson argument. Noë considers examples from neurophysiology, whereas Stanley and Williamson do not, such as "enlarged cortical representations of the hand and arm" in "monkey rake-users" (a result from Iriki, Tanaka and Iwamura 1996), as evidence of the embodied nature of ability. I take Noë's deconstruction and dismissal of Stanley and Williamson's arguments to be valid[13], and will further supply evidence from neurophysiology to support the distinction between knowledge-how and knowledge-that for human-like architecture. Even if one were to reject a distinction between knowledge-how and knowledge-that for humans, the examples from physiology that I bring to bear in a novel way do stand on their own.

That said, I claim this distinction of know-how versus know-that only holds when one takes the human being as an indivisible unit. This is a useful thing to do, most of the time, for example when considering Andrew's ability to ride a bike. To Andrew, this does not consist of a series of propositional statements, or knowledge-that, but rather an intuition or ability that he is able to exercise when he interacts with a bike. However, cracking Andrew open (sorry, Andrew), his bike riding ability consists of neural circuits hooked up to muscles, and circuits that perceive the bike and his environment, like the tree to his left and the pressure from the handlebars on his hands. When Andrew uses his bike-riding ability, he is using these circuits and feedback between them to successfully manipulate the bicycle. A computer system could be endowed with such circuits,[14] once they are elucidated in

---

[12] Noë, 2005.

[13] For humans, or any mind that shares the sensation and motor architectures that give rise to this difference.

[14] Which Andrew has honed through his years of experiential learning, first with training wheels, then with a couple nasty falls, and then better riding, hence more successful circuitry at solving the bike-riding problem

appropriate mathematical detail (as is the aim of the field of computational neuroscience). Analogously, if Andrew didn't know how to ride a bike, no amount of description to Andrew of the necessary neural circuitry would give Andrew the ability to ride a bike. He would only be able to form his own neural circuitry through the experiential route. However, a complete description of this same circuitry, if emulated appropriately in the bike-riding computer, would be enough for that system to successfully ride a bike *on its first try* (theoretically).

The key difference is as follows: someone has intelligently designed the computer bike-riding system, and we may edit it without much difficulty, like the Word document I am currently typing in. Andrew may not be so easily edited. The nature of software allows for explicit access into any level of control over the bike riding process. We command the system execute 4% more force with the left foot on the down stroke by instruction, building in the option for this request in the design of the system. However, with the technology of today, we cannot design in a hack of Andrew's neuromuscular junction to implement this change, except by possibly by drastic neurosurgery. This is in fact the same type of surgery mentioned in a previously quoted passage from Lewis regarding the distinct taste of Vegemite: "unless by super-neurosurgery." Perhaps we could *train* Andrew to do this, but only through experience – trial and error – more experiential knowledge.[15] For a computer, the functional equivalent of building in this feature is not super, nor is it surgical. It only requires further editing of code.

There exists an important distinction on what may be fed to the computer in this fashion.[16] Only once a human has figured out how to distill their hard earned experiential knowledge into the appropriate knowledge-that form to feed into a computer (Python code, for example) can the computer ever gain this experiential knowledge. If computational neuroscience and psychophysics have not yet elucidated the circuitry behind expert level bike-riding ability, there would be no way to formalize this ability into computer code. Even so – I take it as clear that such circuitry does exist and can be formalized and move forward regarding all pieces of

---

[15] This may be achieved through designing a system that gives Andrew explicit feedback on his performance – i.e, "That was 9% more than your original force – try again, but this time a little less!"

[16] Thank you to Anant Matai for pointing this out.

expert human knowledge-how that have been or could easily be formalized, such as that of color recognition.

To see what this means for the knowledge argument, we now examine a move made by Daniel C. Dennett[17] in his 2006 essay, *What RoboMary Knows*.[18] Instead of Mary in the room, we replace her with RoboMary, who is a robotic version of Mary. We are allowed to pull this trick because, as Dennett writes, "Contemporary materialism… cheerfully endorses the assertion that *we* are robots of a sort." If a problem is posed by the fact that we are unable to technologically achieve this feat today, make RoboMary nothing more than a memory bank in which to store all the propositions about color, necessary software for use of these propositions, and a light sensitive apparatus, much like the ones most humans have installed in their heads. Neither Mary nor RoboMary has been exposed to blue light – i.e., light with a wavelength between 400 and 495 nanometers. Both Mary and RoboMary understand that they have cone cells that are sensitive to precisely this sort of light in their retinas (or retina equivalent). While Mary is left to wonder what on earth it must be like to feel those cells activate (i.e., see blue), RoboMary may go ahead and try it! Nothing is to stop RoboMary from setting the [IS BLUE LIGHT BEING SENSED RIGHT NOW?] field from false to true. She is, of course, lying to herself – there is no blue light in the room. I wouldn't call this imagination so much as controlled hallucination – if RoboMary wasn't aware that she had been behind the decision to set the [400nm light?] field to true, she would believe that there was actually light with a wavelength in that range entering her sensor. But after activating this sensory pathway, she will have had a blue experience, and thus be able to recognize when this happens again. All it takes is writing a value to where one is usually read from the outside world.

---

[17] Before knowing the existence of this hilariously snarky essay, I made the same move with regards to turning Mary into a robot, believing I was the first to do so – only I called her MaryBot where Dennett calls her RoboMary. Dennett may have had the idea ten years before me but who thought of the better name?

[18] I suppose if I am to bring in his example, I should address his argument. The long and short of my response is that I challenge his idea that Mary's complete understanding of the physical facts gives her significantly better odds of bumping into blueness roaming the possible search space using her imagination, Additionally, that she would recognize blue if she found it is not clear to me at all. Dennett unfairly overestimates the mind's knowledge of its own state. For more, see *Dennett on the Knowledge Argument* by Howard Robinson.

By virtue of being a computer, RoboMary's systems are accessible at all levels they were designed. RoboMary does not learn anything when she leaves the room, because computers can pull together knowledge-how out of knowledge-that. This comes from the fact that their artificial nervous systems can be intelligently designed however we like, and the quirks that give rise to this distinction in humans are not necessarily present in such a system as RoboMary. The quirk in this case is the inability of humans to run sensory simulations as RoboMary does. We don't have voluntary access to our afferents in this fashion, whereas nothing stops us from designing systems that do. For all intents and purposes, barring Lewis' super-neurosurgery, humans have only one route to knowledge-how or ability or expertise: experience. Computers, on the other hand, given the modularity of software and its ease of editing (as opposed to the analogue physically instantiated nervous system), leave open the possibility of constructing knowledge-how with knowledge-that.

Neuroscience research supports the claim that I make regarding the necessity experience to develop skill. The following is a passage from a 1998 review article on the role of them cerebellum, a brain area in the hindbrain, in motor and procedural movements.

> *Novel movements must be learned. In compound and sequential movements, one does not and cannot consciously plan all of the muscle actions…. Movement is thus largely automatic – controlled by a background subconscious mental subroutine. …the cerebellum would acquire control of the task by recognizing the contexts in which each 'piece' of consciously initiated movement occurs. After repeated trials, the cerebellum would link that context to the movement generators, so that reappearance of the context would automatically trigger the movement. Ultimately, the cerebellum would largely take over the process, as a background subconscious mental subroutine, with minimal help from the cerebrum. The cerebrum and the subject's foreground conscious mental activity would be free to do and contemplate other things.*[19]

For humans, "novel movements must be learned" so they may be brought under automatic, subconscious control – that is, packaged up into intuition. This leaves the individual "free to do and contemplate other things." For computers, novel movements may simply be programmed.

---

[19] Thach, 1998.

A second example of the neuroscientific basis for the instantiation the know-how know-that distinction is the neural efficiency hypothesis. Put simply, this hypothesis claims that with experience comes learning in the form of greater neural efficiency, which is correlated with greater skill, as evidenced by patterns of less widespread activation of pertinent brain areas in experts than in novices. Numerous studies show this pattern in elite athletes as compared to novices, who benefit from greater neural efficiency as evidenced by a smaller area of activation and greater intensity of activation in that pertinent area.[20,21] In a particularly useful example for our purposes here, one study[22] examined patterns of activation across chess players of varying degrees of expertise, comparing the breadth of activation (i.e., the neural efficiency) when the players were attempting chess-related tasks to the activation that occurred when the players attempted tasks unrelated to chess, their area of expertise. This example goes beyond motor learning into creative problem solving and ability. Not only did the better chess players do better on the chess related tasks, as one would imagine, but they also showed reduced widespread activation by a common metric (event-related desynchronization). What does this mean for Gilbert's claims regarding chess? That he was right (in human brains) that imparting all the maxims and tricks to a junior chess player won't make them stronger players. It takes chess *experience* to modify these brain areas to make them more efficient, which in turn improves cognitive ability in the relevant area.[23]

These theories and the supporting data from these studies show that knowledge-how, or experiential learning, has neural correlates. But what should be made of knowledge-that, or factual knowledge? This should be thought of as propositions that may be learned and represented in the brain without any experience necessary at all. When you learn and memorize that the capital of France is Paris, you are setting up a hierarchical relationship between two concepts in your mind. Once you memorize this, there is no doubt some neural correlate of this new bit of knowledge that you learned – but it did not require any experience. Without

---

[20] Callan and Naito, 2014.

[21] Bernardi et. al., 2013.

[22] Grabner et. al., 2006.

[23] Klimesch et. al., 2003.

ever having been to Paris or France, you can know this. Additionally, you can rewrite this understanding without needing to acclimate yourself, should the capital change to Nice. You can learn that *pi* is equal to about equal to 3.14159 without measuring it yourself, just as easily as you could have learned it was about equal to 1.41593.

4 Knowledge-How, Expertise, and Artistry

What does the distinction between knowledge-how and knowledge-that mean for our question regarding the requirements for artistry? As a reminder, under the definition I put forth in the introduction of this paper, artistic knowledge-how is not a requirement for *making* art. If a toddler accidentally knocked over a paint bottle onto a blank canvas and the result were placed in a modern art museum where it was adored by museum goers, who believed the work was powerfully moving and clearly crafted by an artist, this work counts as art. We have not yet determined if the toddler is or is not an artist in this scenario – but intuitively, the toddler deserves no credit for this accidental art. First of all, we’ve taken the toddler’s motions to be accidental, so the toddler did not make any artistic choice regarding the work. Secondly, even if the toddler had intentionally knocked over the paint, it’s fair to claim that the toddler had no detailed conception of what the resulting splash pattern would look like – not just because he is a toddler, but also because he has not developed experience with splashing paint onto canvases and evaluating the result. Contrast this with Paul Jackson Pollock, an abstract artist well known for his drip painting techniques. Pollock’s experiential knowledge with dripping paint onto canvases means that he has some intuition of what exactly will come out of just this sort of flick of the brush. Does the toddler's mother, who found the work on the floor of her studio, deserve the artistic credit for noticing the value of the accidental work? She doesn’t either – she played no part in its creation or design. I propose that the toddler, or any actor who spills the paint without the appropriate know-how, deserves no artistic credit, specifically because they have no artistic experiential knowledge. Possession and application of artistic know-how in crafting a work of art should be a requirement for calling the creator of such a work an artist.

This constructed requirement matches well to previous examples. The Medici family would not get artistic credit for works that they commissioned, because they would not have applied whatever such know-how they had to the works they commissioned. Gaudí fits this model as well as the artist behind La Sagrada Familia. His intuition and learned sense of style certainly played into his crafting of the designs and plans for the incredible building. What of the artisans who executed the details of the church, making artistic decisions on their own? Given that it is safe to say that these lesser-known sculptors were acting on their learned abilities to sculpt, there is an argument to be made that they are contributing artists. But they only applied their skills and intuition to distinct parts of the church, whereas Gaudí's vision extended over the entirety of the church, as well as its initial conception.

I have written at length about visual examples, but what of art such as music or dance? Someone who has never danced, who writes up complex choreography and has expert dancers learn and perform it, should the work turn out to be emotionally salient, is not an artist, even while the performance may be art. The novice choreographer got very lucky – this is like Mary imagining blue by accident in her black and white room.[24] A musician who reads a pre-written score fluently may not obviously be making new artistic decisions, but if they have been trained and have artistic experience, the subtleties such as their choice of playing some notes longer or shorter, or relationally louder or softer are indeed artistic choices.

I have not specified exactly what amount of experiential knowledge is required for artistry, firstly because experiential knowledge has not been quantified here, and secondly because experiential knowledge, and therefore artistry, should be taken as a continuum. What of child prodigies who are producing incredible works of art with very little experiential knowledge? Some minds are more efficient information processors and quicker learners from the get-go. Artists should not be judged primarily on their level of experience but rather by the value of their art. Regardless, some know-how should be required for artistry. Mozart's early symphonies (he began writing symphonies at age eight) were clearly influenced by those written by

---

[24] The difference is that the search space of possible dance moves is smaller than that of what the blue experience could be, and your initial map of dance move space is bigger when you are already an experienced body-mover, compared to Mary who has never seen any color at all.

J. C. Bach (son of J. S. Bach).[25] Even as a child prodigy, Mozart didn't pull his early genius out of thin air. He had to first learn to play piano, and while he did so unusually quickly and unusually well, he still needed that same input of experience in order to wield his impressive know-how to act as an artist, producing impressive symphonic works.

The neural efficiency advantage of experts over novices has been shown in artists as well – the results from one study led authors to hypothesize that the greater efficiency exhibited by the expert allowed him to focus on higher-level concepts in his portrait. In this study, an expert and a novice artist have their brains imaged while drawing pictures of the same face. Analyzing results, the authors write:

> *The artist, who sees and thinks about faces professionally, may require little involvement by this area of the brain normally associated with facial processing. The novice may require greater involvement, suggesting he is processing faces at a "lower" level, which deals with features rather than the "meaning" behind the face. In effect, the novice seems to be "copying" the face; the artist is "seeing beyond" the features.*[26]

As the novice "deals with 'features' rather than 'meaning'" in his picture, he is pulling from the factual information available to him contained in the picture from which he is to draw the portrait. He is following the implicit directions from the photo. The artist has the same factual information available to him (such as the particular shape of the nose he is trying to draw) but is able to essentially automate execution of the lower level features and make higher-level decisions.

What about art-making with regards to non-humans? We have established that the knowledge-how/knowledge-that distinction is a product of the functional architecture of the human nervous system, whereby skill can only be gained and internalized through experience, whereas factual learning (the internalization of propositions like $x < y$) can be achieved without any experience necessary, but cannot itself be a pathway to experience and expertise. But computers, which do not have to follow the same rules that constrain the nervous system (because we can

---

[25] Bourdon, 2010.

[26] Solso, 2001.

program them as we like) do not necessarily need to have gathered *themselves* the experiential knowledge we require for expertise in artistry to deserve the title of artist. In other words, computers that have been endowed with someone else's expert know-how[27] are, in effect, experts. This is the functional equivalent of performing Lewis's super-neurosurgery.

There is still room for meaningful objection here. I claimed that human experts are no longer just consciously following directions like novices do (and dedicate their brains to do it), but have internalized a powerfully flexible collection of tips and tricks in the form of functional changes in their brain, allowing them to automate the rule-following parts and concentrate on higher level decisions for which they have not yet internalized the best rules (because one can always practice more, have more extensive experience, a stronger intuition). This automation and internalization also allows experts to complete the task more quickly than a novice. A computer who is given the difference between an expert and novice, i.e., the tricks and patterns the expert has internalized, under this paradigm, is an expert too. "Wait a minute!" the skeptic exclaims.[28] What if we take all of the rules and tips and tricks that the computer has access to, which bestow expertise upon any rule-following possessor, and give them to the novice? Shouldn't we now consider the novice an expert? And isn't this a contradiction of the claim that the knowledge-that/knowledge-how distinction exists in humans, and that expertise is only available through experience?

What this objection points out is a more direct refutation of Ryle's chess example. The objection asks – what if we gave a complete chess beginner the complete code to Deep Blue? For every move, instead of relying on their intuition of how to respond to Kasparov, they carry out the calculation that Deep Blue does – that is, they navigate a preposterously enormous search tree. They will make the same moves as Deep Blue because they are following the same rules. Why does Deep Blue get the designation of chess expertise for blindly following this set of rules where a human who blindly follows this set of rules does not?

---

[27] . . .which needs to first be boiled down to know-that so that the computer may receive the information.

[28] Thank you to Anant Matai for bringing this objection to bear.

This question gets at the fact that in theory, ability is decomposable into propositions. Great chess playing really *is* just following some strategies. But in reality, humans would never have the time or concentration to carry out such calculations consciously! This comes down to the fact that computers have much faster switching speeds than the speed at which humans transfer information, and the potential for more vast and reliable memories than humans. In other words, computers are better novices than humans. Even though both novice humans and computers are blindly following maxims to exhibit the same behavior, computers achieve this with such speed that they are more comparable to the experienced human, that is, the human that has streamlined its brain activation patterns to perform the necessary calculations subconsciously and efficiently.

Let's take the task of catching a fly ball in the game of baseball.[29] Let's say we have someone who's never had the experience of catching a pop-up try to catch their first one. We do, however, give them all the necessary physical values, such as three data points of where the ball is in space and the time that had transpired since the ball was hit for each point. If our novice has taken high school physics, they do have enough propositional knowledge to calculate with a small margin of error when and where they'll need to be in order to catch the ball. Will they have enough time to do this calculation and then get to the right spot? Absolutely not. An experienced outfielder does not consciously run through these calculations, and yet knows via their intuition where the ball is headed. They have built a model through experience in their brain that can take in visual cues about the ball's movement against the sky to know where to go to catch the ball. The computer could be given the data points without having any memory of having caught a ball ever before, but would perform the Newtonian calculation much, much faster than the novice human. This is why the computer in addition to the expert-derived rules should be considered an expert, whereas the novice human, no matter how many maxims, rules, equations and "wrinkles" we give them, is still just a novice.[30]

---

[29] I always struggled with this, which was a problem. Little League coaches would stick me in right field, which they thought would be a safe place for me to think about these sorts of things. The joke's on them.

[30] These points run parallel to the popular response to John Searle's Chinese Room Argument. While Searle says the man in the room does not speak Chinese, detractors point out that the complex of the

5 Aaron the Algorithm

AARON, the drawing algorithm was first run in 1973. Harold Cohen, a British artist who had developed a strong professional reputation as an expert painter and colorist, was attempting to emulate human drawing ability in a computer program. Notably, Cohen describes his interest in writing computer code that would allow a machine to make human-esque art, not to generate what others might call "computer art," which sometimes appears as fun swirls of color in complex patterns as a result of mathematical formulae as a rule for graphics generation. Cohen was instead looking to create something that came up with and executed pieces in a parallel fashion to humans. Pamela McCorduck, in her book *AARON'S CODE: Meta-Art, Artificial Intelligence and the Work of Harold Cohen* writes in detail about AARON's conception and evolution at the hand of Cohen over the eighteen years of its existence at the time of her writing in 1991. She offers pertinent lines of inquiry in her work, some of which will be considered here in relation to the questions I am addressing regarding know-how in computers and how that pertains to labeling the artist.

Consider the drawings in the Appendix. Starting with Figures 1 and 2, you might observe abstract doodles. With Figures 3 and 4 you likely recognize what are clearly human shapes, up to human activities, and the outlines of plants, looking quite plant-y. Figures 5 and 6 are colorful and contain human forms with expressive faces. As I'm sure you can guess, these were done by AARON. 'But really?' you may be scoffing – how can a computer draw this, having never seen a face? Or what color to use for skin? Or how to orient humans correctly (i.e., not floating or upside down)?

In order for AARON to be able to represent the world, Cohen needed to feed the algorithm facts about the world (this is quintessential knowledge-that). He discusses this at length in his essay *How to Draw Three People in a Botanical Garden*, where he is hoping to draw insight from his endeavors with AARON on the nature of visual representation. Cohen poses a central a question not unlike the one I am

---

man *and the room* together do, in fact, speak Chinese, just as the computer *and the rules of expertise* together are, in fact, an expert.

seeking to address here:[31] "What do computer programs – and, paradigmatically, human beings – need to know about the external world in order to build plausible visual representations of it. What kind of cognitive activity is involved in the making and reading of those representations?" (AAAI-88 Proceedings, Invited Talks, p, 848). Cohen here must tackle the problem that AARON doesn't see at all, and therefore cannot go about making its own visual representations of an external world it cannot access. However, it can be fed propositional facts, and construct a visual representation via those facts (this is precisely what humans cannot do due to the limits of our nervous system). For what in my use of Ryle's terms[32] is a deconstruction of knowledge-how to knowledge-that, Cohen uses the apt term 'knowledge engineering':

> *[Making representations] requires a special body of knowledge – knowledge of representation itself – that is part of the expertise of the artist, just as representation of a body of knowledge within an expert system requires an analogous expertise of the knowledge engineer. Understanding the nature of visual representation requires asking what artists need to know in order to make representational objects; what they need to know, not only about the world, but also about the nature and the strategies of representation itself.*

Cohen hits on what we have discussed. The artist's expertise, their knowledge-how of making representations, relies on knowledge of "representation itself," as Cohen puts it. AARON is no doubt successful at making representations, at least representations that evoke associations to human viewers. AARON's map of association extends no further than how to draw a two-dimensional person form, whereas ours extends much, much deeper. But what is important is that AARON uses a comparably small amount of propositional knowledge to expertly draw these scenes. The expertise that AARON exhibits here didn't come out of random chance but rather a distillation of Cohen's expertise in representing people.

We can find evidence of Cohen's artistic intuition written explicitly into AARON's code. In the *Botanical Garden* paper, Cohen launches into a new level of detail about how he gets AARON to represent recognizable structures such as a

---

[31] But I address this sort of question from a different frame of reference and theoretical backing: the debate around the Knowledge Argument and Ability Hypothesis.

[32] Although, to be clear, Ryle would not endorse this usage because he claimed knowledge-how is not reducible to knowledge-that.

human hand. “Thus AARON… proceeds by expanding each entry in the class hierarchically into an internal tree-structure, at the lowest levels of which are the management procedures responsible for the generation of individual elements of the drawing. There is, for example, a ‘hand’ manager, whose sole task is to produce examples of hands on demand, to satisfy the specifications…” (AAAI-88 Proceedings, p. 851). This hand manager modularity is analogous to the sort of change that goes on when a brain goes from novice to expert. Instead of trying to draw the hand like the novice does in the MIT portraiture study, using widespread brainpower to copy the contours of the photo, AARON’s packaging up of hand-drawing ability into its own little hand-drawer is closer to how the artist in the brain scanner drew the portrait from the photo – by allowing the automation developed from his expertise to handle individual details.

What should we say of computational shortcuts? At this stage of its existence in 1988, AARON makes humanoid outlines that have the appearance of existing in three dimensions. All successful two-dimensional portrayals of three dimensions have this property, but the humans that draw such representations bring to bear their experiential knowledge of what it is like to interact with objects in three dimensions in their artistic representations of real-life scenes. Instead of teaching AARON to conceive of a three-dimensional scene and then draw in two, as many human artists go about drawing, up until just before Cohen wrote this paper in 1988, AARON took a non-human-like shortcut. It never represented the subjects of its works in three dimensions but was rather was given the know-how needed to feign three-dimensionality. Cohen writes “Until recently, AARON has never had a fully 3-dimensional knowledge-base of the things it draws: foreshortening of arms or the slope a foot in the representation were inferred from AARON’s knowledge of principles of appearance, not by constructing the figure in 3-space and generating a perspective projection” (AAAI-88 Proceedings, p. 851). This seems like a potential threat to the idea that Cohen is giving AARON his own, learned expertise – AARON doesn’t include any representation of “3-space” in this case, which is no doubt an important part of the human embodied experience.

Until there is a need for this inclusion as evidenced by unnaturalness in AARON's output, this shortcut will do fine in terms of output – but it does work against AARON's appeal for expertise. At some level in the computational circuits of an artist's brain, three dimensions must indeed be represented somehow in two dimensions if the artist is trying to cue a viewer into interpreting the piece that way – to extend into the wall or pop out of the canvas (instead of a flat bunch of lines, which is all that there is on the page, in a literal sense).[33] If AARON skips over these steps directly to a strategy no doubt also employed by artists (i.e., if I shorten this foot compared to that one, a human viewer will interpret this foot to be at an angle, not actually shorter), what is the consequence? I would argue that this makes AARON a little bit less of an artist. The human activity of mapping from thinking in 3-space to drawing in 2-space is in itself a knowledge-how, or ability. We can see evidence for this in looking at how children first draw in two dimensions (picture a square and a triangle for a house) and then later learn to draw in three (if they do). A computer, when fed this ability to render 3-space into 2-space representations (know-how in humans) in the form of computer code (knowledge-that) has inherited a deeper ability, deriving its image from experiential knowledge that was a little bit closer to that gathered by humans, its original source of expertise.

For example, consider two third graders. Ben has figured out a trick in order to be ready for the multiplication quiz on the nines table: he holds up his hands, palms up, and counts *x* fingers into the ten fingers and curls up this finger, where *x* is the multiplier of nine. Then he reads from both sides of the curled up finger, which gives him the product of the multiplication – i.e., third finger down, product is two on the left of the third finger and seven to the right of the third finger, so three times nine is twenty seven. Amy, the other third grader, prefers to derive the product through addition every time. Instead of manipulating her fingers, she adds the number nine three times, yielding twenty seven. Ben might be faster at giving you the product of, say, nine multiplied by nine than Amy by way of his nifty finger trick.

---

[33] Importantly, the visual sense of three dimensions is constructed in the brain even in humans – our retinal images, that is, the image of the world on the retina, is a two dimensional representation of a three dimensional world. Through binocular vision and some complex circuitry, the brain creates a visual illusion of three dimensions.

But while Amy is doing multiplication in a real sense, Ben is taking a shortcut that will not treat him well as third grade continues. Who will learn their elevens table faster, if neither memorizes, as before? From this way of thinking, AARON's new skill in 1988 of rendering images in three dimensions[34] and then projecting them onto a plane is an example of a deeper process, more Amy than Ben, than when AARON used the shortcut of "appearance" to prompt humans to interpret its drawings in three dimensions.

Now I'd like to discuss AARON in terms of Mary's Room. Is Mary, stuck in her room, a good analogy for AARON and his strictly propositional knowledge? Or is liberated, more able Mary a better analogy? Well – Mary the human in the room does not have the circuitry or anatomy to demonstrate know-how from just facts. She cannot know blueness from just the physical facts. AARON is, in a sense, similarly trapped in a black and white room. Because AARON cannot see anything, it has never seen anything like the subjects of its drawings – plants, people, and by the 2000's, AARON could expertly color its artwork. Like Mary, AARON has only physical facts, such as range of motion in parts of human physiology. But unlike Mary, AARON demonstrates the relevant knowledge-how – knowing how to draw and paint. A better analogy would be RoboMary. Despite never having seen color, RoboMary marshals her available factual knowledge and, using her non-human architecture, is able to simulate color to herself. AARON was never fitted with the sort of neurological color response equipment that we have, and that we gave to RoboMary. But just like RoboMary, simulation is indeed how AARON is able to keep track of the drawing it conceives of. Instead of drawing and evaluating, AARON works by conceptualizing and periodically evaluating its work. The drawing is completely finished by the time the machine goes to instantiate the picture on canvas or paper.

Under the paradigm we have set up, there is certainly a case to be made that AARON – that is, the expertise in the form of the software and the novice rule-follower that is the computer it runs on – has the requisite know-how to be an artist. AARON produces art, satisfying the first requirement. Furthermore, AARON,

---

[34] Note that AARON at this stage did its arranging of forms in two dimensions. The innovation was that it then rendered them in three dimensions to check that the three dimensional referents of the two dimensional constructions made spatial sense, before projecting them back into two dimensions.

even while having no experiential knowledge (AARON does not keep track of past paintings and therefore has no memory of what it painted before, or the sort of efficiency-building set up of the human brain) bypasses the constraints that act on the functional architecture of the human nervous system in that it follows experience-derived rules (derived and translated to code by Harold Cohen) and displays expert behavior as a painter (so says Cohen himself in his essay AARON, Colorist: From Expert System to Expert).

6 Conclusion: Whose Know-How is It?

Even if we can conclude that AARON is allowed the requisite know-how to be considered an artist, there is more to discuss regarding how its know-how was gathered. One argument that holds weight would be that because all of AARON's expertise is really just Harold Cohen's expertise, Cohen deserves artistry credit. AARON is to be seen more as an apprentice (this is indeed how Cohen describes the algorithm in one of his papers). At first glance, this may seem silly. Apprentices can learn a lot from a master, including the master's formalized version of their hard-won expertise, and still deserve credit. But imagine a student that after finishing school or an apprenticeship, only made decisions based on the notes they took during their time learning. Artists are not just the sum of their influences. Mozart, while he was initially influenced by J.C. Bach as noted before, did not continue to produce symphonies in that style, but later became a pioneer of his own style.

In reality, students who take notes and learn as apprentices may internalize these lessons through experience, built into expert intuition over time, no longer needing to consult their notes from earlier years. With each repetition – new painting, new symphony, etc., beyond the influence of their old teacher, noise is introduced to the learning process – new art that piques the student's interest, a new method that the student accidentally discovers, etc. This is a benefit of human neural architecture, which necessitates repetition, one that AARON cannot have, because AARON's rules do not change unless Cohen changes them, like the student who never becomes *his own* expert because he only consults his notes and never his growing intuition.

However, today in 2019, *there are* algorithms that can learn in this way. AARON is an example of an old brand of AI, one that embodied the idea that intelligence can be formalized into knowledge-that can be written into the very code of an algorithm from the get-go. Architectures used in the field of machine learning such as artificial neural networks were inspired by that same by-experience-only architectures in the human brain that give rise to the distinction between knowledge-how and knowledge-that. Consider the 'Creative Adversarial Network', from a paper by Elgammal et. al., exploring modifications to the popular machine learning network used to generate images called 'Generative Adversarial Networks.' These use a certain stratagem reminiscent of a two-player game, where one algorithm, *A* is trained to recognize an image that is part of a set of images, *I,* that share characteristics (which may not be rigorously defined – the network teaches itself). Meanwhile another algorithm, *B*, is trained against *A* to generate images that share key characteristics with those in *I* to the point where *A* sorts the generated image as a part of the set. In this manner, such an algorithm develops know-how out of know-that through experience, similarly to humans. Without the detailed mathematical description of what characteristic makes the images in *I* belong to that set, a programmer could not just perform Lewis' super-neurosurgery and give the computer the *ability* to selectively sort images into the set *I*.

Whatever case can be made for AARON in terms of artistry by experience, a machine learning algorithm that was able to make images that count under our definition as art in this manner would seem to have an even stronger case. Moving back to our chess example – while Deep Blue was able to beat honed human intuition in the form of Gary Kasparov via simple rules (knowledge-that), obeyed quickly and completely, a machine learning program called AlphaZero was able to *teach itself* the tricks to good chess just from the knowledge of the rules of the game, as well as shugi, a Japanese version of chess, and Go, a complex Chinese strategy game. AlphaZero thoroughly beat world champions in all three strategy games. What are the implications here? AlphaZero is making the best of both worlds – human intuition building via learning from experience, with all the blinding speed of transistors instead of synapses. AlphaZero is clearly more of *its own* expert than Deep Blue was, just as CAN puts forth a better claim toward being a computer artist than AARON. Even though they are both following rules, AARON is following someone else's rules, whereas CAN made its own.

7 **Appendix** (I did not yet get permission to use these images…)

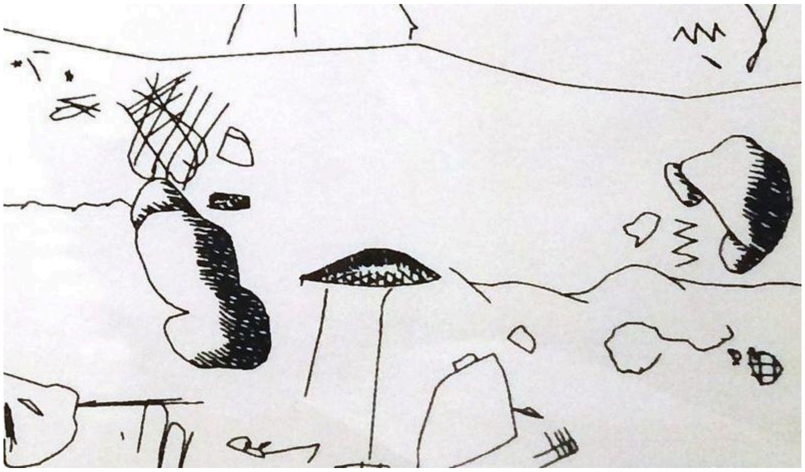

Figure 1

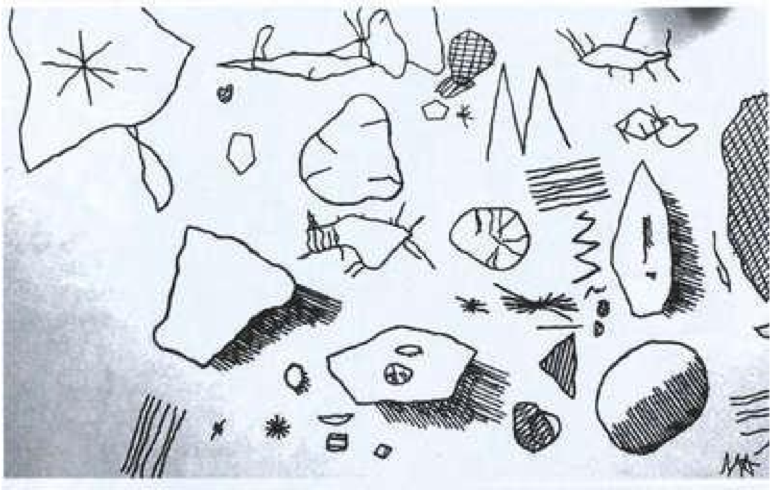

Figure 2

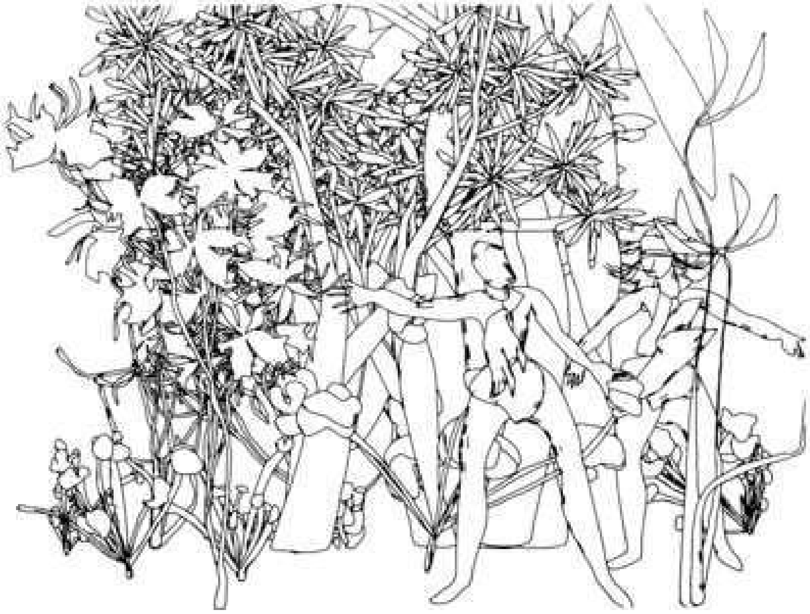

Figure 3

Figure 4

Figure 5

Figure 6

8 Bibliography

Becker, Howard S. "Art As Collective Action." *American Sociological Review* 39, no. 6 (1974): 767–76.

Bernardi, Giulio, Emiliano Ricciardi, Lorenzo Sani, Anna Gaglianese, Alessandra Papasogli, Riccardo Ceccarelli, Ferdinando Franzoni, Fabio Galetta, Gino Santoro, and Rainer Goebel. "How Skill Expertise Shapes the Brain Functional Architecture: An FMRI Study of Visuo-Spatial and Motor Processing in Professional Racing-Car and Naïve Drivers." *PloS One* 8, no. 10 (2013): e77764.

Boden, Margaret A. *The Creative Mind: Myths and Mechanisms*. Routledge, 2004.

Bourdon, Austin. "Johann Christian Bach's Influence on Mozart's Developing Style," 2010.

Bringsjord, Selmer, Paul Bello, and David Ferrucci. "Creativity, the Turing Test, and the (Better) Lovelace Test." In *The Turing Test*, 215–39. Springer, 2003.

Callan, Daniel E, and Eiichi Naito. "Neural Processes Distinguishing Elite from Expert and Novice Athletes." *Cognitive and Behavioral Neurology* 27, no. 4 (2014): 183–88.

Cohen, Harold. "AARON, Colorist: From Expert System to Expert," 2006.
———. "Colouring without Seeing: A Problem in Machine Creativity." *AISB Quarterly* 102 (1999): 26–35.
———. "How to Draw Three People in a Botanical Garden.," 89:846–55, 1988.
———. "Parallel to Perception: Some Notes on the Problem of Machine-Generated Art." *Computer Studies* 4, no. 3/4 (1973).
———. "The Further Exploits of AARON, Painter." *Stanford Humanities Review* 4, no. 2 (1995): 141–58.

Dennett, Daniel C. "Epiphenomenal" "Qualia." *There's Something about Mary: Essays on Phenomenal Consciousness and Frank Jackson's Knowledge Argument*, 2004, 59–73.
———. "What RoboMary Knows," 2006.

Elgammal, Ahmed, Bingchen Liu, Mohamed Elhoseiny, and Marian Mazzone. "CAN: Creative Adversarial Networks, Generating" Art" by Learning about Styles and Deviating from Style Norms." *ArXiv Preprint ArXiv:1706.07068*, 2017.

Grabner, Roland H, Aljoscha C Neubauer, and Elsbeth Stern. "Superior Performance and Neural Efficiency: The Impact of Intelligence and Expertise." *Brain Research Bulletin* 69, no. 4 (2006): 422–39.

Guiterman, Arthur. *A Poet's Proverbs: Being Mirthful, Sober, and Fanciful; Epigrams on the Universe, with Certain Old Irish Proverbs, All in Rhymed Couplets*. New York: EP Dutton, 1924.

Iriki, Atsushi, Michio Tanaka, and Yoshiaki Iwamura. "Coding of Modified Body Schema during Tool Use by Macaque Postcentral Neurones." *Neuroreport* 7, no. 14 (1996): 2325–30.

Jackson, Frank. *Epiphenomenal Qualia*. Abr, 1982.

———. “What Mary Didn’t Know.” *The Journal of Philosophy* 83, no. 5 (1986): 291–95.

Klimesch, Wolfgang, Paul Sauseng, and Christian Gerloff. “Enhancing Cognitive Performance with Repetitive Transcranial Magnetic Stimulation at Human Individual Alpha Frequency.” *European Journal of Neuroscience* 17, no. 5 (2003): 1129–33.

Lewis, David. “What Experience Teaches,” 1988.

Ludlow, Peter, Yujin Nagasawa, and Daniel Stoljar. *There’s Something about Mary: Essays on Phenomenal Consciousness and Frank Jackson’s Knowledge Argument*. MIT Press, 2004.

Malenka, Robert C, and Mark F Bear. “LTP and LTD: An Embarrassment of Riches.” *Neuron* 44, no. 1 (2004): 5–21.

McCorduck, Pamela. *Aaron’s Code: Meta-Art, Artificial Intelligence, and the Work of Harold Cohen*. Macmillan, 1991.

Noë, Alva. “Against Intellectualism.” *Analysis* 65, no. 4 (2005): 278–90.

Pease, Alison, and Simon Colton. “On Impact and Evaluation in Computational Creativity: A Discussion of the Turing Test and an Alternative Proposal,” Vol. 39, 2011.

Rich, Anina N, Mark A Williams, Aina Puce, Ari Syngeniotis, Matthew A Howard, Francis McGlone, and Jason B Mattingley. “Neural Correlates of Imagined and Synaesthetic Colours.” *Neuropsychologia* 44, no. 14 (2006): 2918–25.

Robinson, Howard. “Dennett on the Knowledge Argument.” *Analysis* 53, no. 3 (1993): 174–77.

Ryle, Gilbert. “Knowing How and Knowing That: The Presidential Address,” 46:1–16. JSTOR, 1945.

Searle, John R. “Minds, Brains, and Programs.” *Behavioral and Brain Sciences* 3, no. 3 (1980): 417–24.

Solso, Robert L. “Brain Activities in a Skilled versus a Novice Artist: An FMRI Study.” *Leonardo* 34, no. 1 (2001): 31–34.

Stanley, Jason, and Timothy Willlamson. “Knowing How.” *The Journal of Philosophy* 98, no. 8 (2001): 411–444.

Thach, William T. “What Is the Role of the Cerebellum in Motor Learning and Cognition?” *Trends in Cognitive Sciences* 2, no. 9 (1998): 331–37.

Turing, A. M. “Computing Machinery and Intelligence.” *Mind* LIX, no. 236 (October 1, 1950): 433–60.

9 Acknowledgments

This work was made possible by generous funding from the College Center for Research and Fellowships at the University of Chicago, and by the leadership and vision Director Nichole Fazio – my deepest gratitude. Thank you to Bill Hutchinson, Jason Bridges, Chris Kennedy, and Heather Keenleyside for direction, discussion, and valuable feedback on earlier drafts. Thank you to Anant Matai, Joe Berusch, and Seth Markow for scintillating discussion. Thank you to Denisse Manuela Morales, Jeremy Ornstein, Rafael Ornstein, and Susannah Sherry for listening to and supporting me. Finally, thank you to the other Fellows at the College Summer Institute for a summer of learning and fun.