Human machines

Katalin Balog

Human automata

Nathaniel in E. T. A. Hoffmann's *The Sandman* loses his sanity over falling in love with a wooden doll, the beautiful automaton Olympia. Unbeknownst to him, she is the sinister creation of a mad scientist and a black magician. As one of Nathaniel's friends observes:

> We have come to find this Olympia quite uncanny; we would like to have nothing to do with her; it seems to us that she is only acting like a living creature, and yet there is some reason for that which we cannot fathom.

But uncanny and sinister as the wooden doll might have struck them, contemporary readers of the tale were also meant to see Nathaniel's infatuation as farcical - Olympia is obviously soulless, and only Nathaniel doesn't see this.

Hoffmann's tale prefigures our predicament. The sharp dividing line between humans and machines, so evident in the story for everyone but Nathaniel, is getting blurrier. We now find ourselves in the curious position of entertaining the possibility, and increasingly, the reality of relationships with machines that hitherto were reserved for fellow humans.

The 17th-century French philosopher René Descartes played an important role in the growing fascination with automata during the Enlightenment. Descartes thought the bodies of humans and animals were machines - mere mechanisms governed by physical law. Descartes believed the human body could be understood and, in principle, replicated entirely mechanically - as a system of levers, fluids, and motions. Despite this, automata, no matter how perfect, could not be human beings. What set humans apart, he believed, was language, reasoning, and consciousness, which cannot occur in a purely physical being and require an immaterial, immortal soul. Spectacles of mechanically animated figures generated public interest, like the elaborate automata in the royal gardens at Saint-Germain-en-Laye. Still, these soulless figures were seen as false doubles, mere counterfeit humans.

An apocryphal rumor surfaced in the 19th century that Descartes himself created a doll, in the likeness of his deceased baby daughter, Francine, which he carried with him everywhere. Legend has it that he might even have attempted to attract Francine's soul to reanimate the doll. But when sailors on his ship discovered it, they were so horrified - whether by its soullessness, or the very possibility that it might have a soul - that they threw the doll overboard.

Two doctrines about humans and machines

The clear division between human and machine introduced by Descartes, alongside the Gothic sentiment that there is something sinister about machine simulacra of humans, has faded by now.

Two doctrines about *consciousness* reflect the narrowing gap between human and machine. One concerns human nature itself, the other the nature of machines.

What exactly is consciousness, in this context? Someone is conscious, as Thomas Nagel famously put it, if "there is something it's like" for that creature to experience the world. When one hears waves lapping against the shore or sees water shimmering in a glass, the world appears in a particular way - there is something it is like to have these experiences. If you ever wondered if others experience the world as you do when looking at the blue sky - and not, say, in the way you experience yellow - you understand that conscious phenomenology goes beyond mere representation of the world. It is an inner, subjective dimension of experience itself.

The first view, *illusionism*, denies that we are conscious at all. According to illusionism, consciousness is a mere illusion; there is no room for it in a physical universe. There is nothing it is like to feel hungry; it just *appears* as if there were. Within this framework, the gap between humans and intelligent machines that could mimic our cognition and behavior is narrower than once thought. Olympia, powered by some advanced form of artificial intelligence and using language in human-like ways, would still be a lifeless machine. Yet, since neither humans nor machines are conscious, Olympia would not be that different from a human being with regard to the capacity that matters most: the capacity to process information. Despite what generations of sci-fi writers and their audience assumed, her lack of subjective, conscious experience wouldn't set her apart from humans, as, according to the theory, consciousness itself is a mirage. Illusionism is a deeply implausible theory - after all, we are all familiar with our subjectivity, but it is gaining momentum. It is spreading from academia to the internet and to popular media (see Michael Graziano in *The New York Times* and Keith Frankish in *Aeon*). In my experience, students, both undergraduate and graduate, who some years ago heaped scorn on the idea, have lately become far more sympathetic to it.

The second idea moves in the opposite direction, suggesting that current AI systems are conscious, or some future iteration of them soon will be. This view is more widespread than illusionism and is increasingly reflected in popular discourse and culture. It is the reverse of illusionism with regard to the consciousness of humans and machines. Unlike illusionism, the idea that machines are potentially conscious does not have an immediate credibility problem. Machine consciousness is at least intelligible, whereas the idea that one is not conscious verges on incoherence. But even if the thesis of machine consciousness is true, its adoption has a corrosive effect on our self-understanding in that it, too, blurs the boundary between humans and machines.

As I will argue, illusionism and the thesis of machine consciousness are kindred manifestations of the same Zeitgeist; they both lead to a neglect of precisely those human capacities that make freedom and flourishing possible.

Consciousness and value

Why do these disputes about consciousness matter? Because consciousness bears directly on value, and therefore beliefs about consciousness bear on our attitudes to value.

Things matter to us at all because of how we experience them. Perception always exceeds bare sensory character; it also registers affective qualities. A flower appears blue and of a particular shape and smell, but also calming, delicate, fragrant, and beautiful. A cat appears large, with tousled orange fur and shining eyes, but also innocent and lovable. The affective character of our experience of the world underlies the significance of things; it animates the world. If affective tone drained away, the world would remain colored and shaped, but it would no longer *matter*.

Affective experience reveals value, but it also *grounds* it. The flower is beautiful *because* it resonates in the affective experience of conscious beings. Moreover, arguably, *only* affective experience can ground value; without affective experience, nothing could be good or bad for a creature. Imagine a person, let us call her Insensate Mary, in a twist on Frank Jackson's (1982) Knowledge Argument. Though she perceives colors, shapes, textures, etc., much like normal people do, she lacks the sensuous appreciation of value. This can only happen if perceptual and affective content can come apart. They do seem to vary independently: when I am in a bad mood, the same streets look very different than when I am happy. As Wittgenstein says in the *Tractatus* (6.43): "The world of the happy is quite another than that of the unhappy." Or, to take a more directly relevant example, morphine affects pain apparently by leaving the sensory content intact but entirely removing the affective component, the awfulness of pain.

So, Insensate Mary, when she sees a roadside accident, has no "gut reaction": no aversion, no horror, no sadness, no morbid curiosity. In the presence of her loved ones, she feels no love, no joy; she has never experienced the many ways human beings can suffer or flourish. It seems clear that, as a result, she fails to *understand* what good or bad is, and, what's more, nothing could *be* good or bad for her. She has representations without resonance - facts without value.

Not everyone will agree. Illusionists, for example, argue that desires - even without subjectivity - are sufficient grounds for value. Kantians think it is the rational activity of agents that provides the foundation for value. These are relevant considerations for a theory of value, but the fact remains that affective consciousness is a *sine qua non* for value to exist at all.

Now we can see why views about consciousness are so central to the human image. Illusionists claim we are like Insensate Mary - but if so, nothing matters, and the very idea of human flourishing is rendered meaningless. This is how we used to think of robots: unconscious, and therefore incapable of being harmed or helped. Applied to humans, this is a deeply corrosive view. Illusionists have alternative ways of understanding value and agency, but people in general make a strong connection between consciousness and value. This is confirmed by some empirical studies. For example, a 2007 study in the journal *Science* on mind perception found that people's willingness to harm other beings depends on the extent to which they attribute consciousness to them. Contrary to illusionists, it is also widely held that feelings and conscious experience in general explain a lot of what people do. By undermining these intuitions, illusionism weakens people's interest in consciousness and, with it, the whole realm of human understanding that relies on it.

The view that machines are conscious is also problematic for humanistic concerns, but for different reasons, the details of which I will get into later.

Illusionism as a theory of mind

What is striking about illusionism is that before the 20th century, no one had ever advocated it. Virtually all other extant theories about the mind and consciousness: dualism, materialism, idealism, panpsychism, dual-aspect monism, have precedents, often going back to ancient times. Illusionism has taken a long time to emerge as a theory of the mind. What explains its appearance in our present age?

The Copernican Revolution was the first large upheaval wrought by science; Copernicus displaced us from the center of the universe. But it was the subsequent triumph of physics, the reduction of chemical and biological processes to processes familiar from physics, and Darwinian evolutionary theory that transformed our image of the world for good. Scientists and philosophers came to accept that every physical event can be explained in terms of purely physical causes - including complex chemical and biological events.

This idea has far-reaching consequences. Physical events include bodily movements that constitute actions; consequently, actions themselves can be explained in terms of purely physical causes as well. This rules out Descartes's idea that simple, immortal souls are capable of moving matter independently of physical causes. The response has been a widespread embrace of materialism - now called physicalism - among scientists and philosophers. Physicalism holds that everything in the universe, including minds, is physical. Physicalism by itself does not imply that there is no consciousness; its early adopters believed that consciousness is grounded in complex physical processes. Other cultural and scientific developments were necessary for illusionism to emerge.

There is not much about the premodern conception of the mind that survived into the 20th century, except that we are conscious. Yet in the middle of that century, philosophical behaviorists like Gilbert Ryle started broaching the possibility that not only the soul - "the ghost in the machine", as Ryle put it - but even consciousness, as we normally conceive of it, doesn't exist. The key development was that Ryle rejected the first-person, subjective perspective and claimed that knowledge can only be acquired through third-person, objective methods like the ones that science employs. Whatever cannot be demonstrated or understood through the methods of science is ultimately unreal.

He argued, contrary to thousands of years of humanistic tradition, that human beings should be studied the same way we study a rock or a plant. In this spirit, he concluded that the key to understanding mental life is human behavior. His theory was that there is nothing more to consciousness than clusters of behavioral dispositions, e.g., for discrimination, for verbal response, and the like. What we mean when we say that someone is in pain is that they are disposed to wince, groan, seek help, avoid certain movements, etc.

Consciousness denial goes against the one aspect of Descartes's view of the mind that remained the most resilient to revision by scientific developments: his view that the nature of the mind is awareness of its own conscious states. He held that the reality of our consciousness is the one thing we cannot be mistaken about, even if we are mistaken about everything else, even if there is no

world at all outside of our minds. It is easy to see why this is so. When I introspect, I know what I am experiencing because the experience itself is *present* in my introspective awareness of it. When I notice that I am hungry, the feeling of hunger is not separate from the act of noticing. I need no further evidence to know that I am hungry, and no amount of empirical evidence, including from neuroscience, could convince me otherwise. Because consciousness is so readily manifest to anyone with a mind, belief in its existence has held fast.

But in recent years, consciousness denial is back in vogue. Ryle's student, Daniel Dennett, has called into question, starting in the 1980s, that there is something it is like to feel sad, smell lavender, or fall in love. As Keith Frankish puts it in his recent paper "Illusionism as a Theory of Consciousness":

> Illusionism makes a very strong claim: it claims that phenomenal consciousness is illusory: experiences do not really have qualitative, 'what-it-is-like' properties, whether physical or non-physical.

The view no longer rests on behaviorism. Behaviorism, as a science of the mind, has been replaced by cognitive science, which understands the mind in terms of representation rather than behavior. Yet it still shares with behaviorism a bias toward the objective, scientific approach to humans. Illusionism is the endpoint of the journey from understanding ourselves as souls to envisioning humans as unconscious information processors.

Is illusionism true?

Granted, science and philosophy have at times overturned deeply held common-sense beliefs. The scientific image of solid objects as mainly containing empty space, occupied by fluctuating fields, for example, is deeply counterintuitive. Might the case of illusionism be like that? Despite being at odds with common sense, might there be compelling scientific or philosophical arguments in its favor?

At first glance, illusionism seems corroborated by neuroscientific discoveries about the mind-brain connection. Something as intimately personal as depression turns out to correspond to chemical changes in the brain, and meditation correlates with patterns of brain activity. This might lead to the thought that what is *really* going on is the brain processes. But correlation does not imply elimination. These findings show that consciousness depends on the brain, not that it is illusory.

The philosophical case rests instead on what philosophers call the "explanatory gap". Mary, in Frank Jackson's original thought-experiment, is a brilliant neuroscientist who knows everything about the neuro-science of color vision. But she has been held captive in a black-and-white room all her life. We can be sure in advance that she has no idea what it is like to see red if she has never experienced it herself. This is because no account of brain activity *explains* why a particular brain state feels the way it does, and not some other way. By contrast, facts about H_2O fully explain water's properties.

It seems intuitively plausible that if consciousness were physical, it would be possible to explain it from physical facts about the brain. Some philosophers conclude from the gap in explanation - and from Mary's consequent lack of knowledge about what it is like to see red, despite her complete

physical knowledge of the brain - that consciousness must be something over and above the physical. Illusionists, on the other hand, convinced that the physical is all there is, draw the opposite conclusion: consciousness itself is a mirage. This line of thought might appear persuasive. There are good reasons to think the world is purely physical, and it is true that consciousness does not have an explanation in physical terms. Those two ideas do seem incompatible. Nevertheless, they can be reconciled.

The key is that there is another way to understand the gap. When I notice the bittersweet rush of chocolate on my tongue, the very experience becomes my concept of it. That kind of concept is simply different from any scientific concept. No other concept *exemplifies* what it stands for in this way. The gap shows not that consciousness is unreal, but that first-person concepts differ in kind from third-person scientific concepts. This approach was first developed by Brian Loar, and later philosophers started calling it the "phenomenal concept strategy". It reconciles physicalism and the explanatory gap and so neutralizes the philosophical argument for illusionism.

The phenomenal concept strategy, like everything else in philosophy, is controversial. The larger point is, however, that even if one doesn't find the phenomenal concept strategy persuasive, it would be more reasonable to give up belief in physicalism than to give up the belief in consciousness.

The very implausibility of illusionism might be, paradoxically, one of the attractions of the view. Illusionists pride themselves on casting away unscientific dogmas. Having banished Descartes's immaterial souls, they urge humanity to consign consciousness as well to the ash-heap of history. For scientifically minded philosophers, illusionism is exciting in the way Darwinism was exciting in the 19th century, precisely because it overturned our self-conception.

That it could emerge at all reflects a cultural moment long in the making, which puts the emphasis on science and quantitative disciplines, even when it comes to self-understanding, and even at the cost of eroding our true self-image as conscious, feeling, autonomous subjects.

Machine consciousness

If illusionism erodes the significance we attribute to consciousness, the view that machines are conscious raises a different but not unrelated worry, that we overestimate the significance of what machines *can* do.

The legend of the Golem, originating in 16th-century Prague, is a story of a clay doll rendered alive by an inscription of the Hebrew word "emet" on its forehead. It, too, exists on the borderline of machine and human - inert matter animated through magic to serve as protector of the community. But the Golem is closer to a human being: alive, and possibly even conscious. It, like Olympia, elicits ambivalent feelings. It is not truly human, and one is left in the dark about its real nature.

But Olympia clearly doesn't have a mind; it is less clear about Golem. And Frankenstein's monster, that great villain born of the romantic imagination, is undoubtedly conscious. His status is also

closer to being a bona fide human: in Mary Shelley's novel, Frankenstein concocts him from body parts of the dead, animated by some obscure process harnessing electricity. He is also human in his suffering as an outcast, unloved being. Fictional simulacra of humans run the whole gamut from puppets that imitate human behavior to near-human artificial creatures.

In some respects, artificial intelligence is more like Olympia than Frankenstein's monster, in that it is definitely not alive. But we now have real machines that act much like a person in the way they use language. Descartes would be enormously surprised: he held that a purely physical being could not produce something so open-ended and creative as human speech.

But here we are. And we, too, like Nathaniel's friend in *The Sandman*, would like to fathom what is happening inside our machines. This effort is confounded by a salient feature of the AI revolution: that AI systems are intentionally designed to act as a companion, and refer to themselves by the pronoun 'I'. Trained on vast swaths of internet text, chatbots function as aggregators of human knowledge, yet they are engineered not merely to provide information, but to mimic human conversation. They present themselves as persons. Because of this uncanny mimicry, people increasingly are willing to attribute intelligence, understanding, and even consciousness to chatbots. Some people worship them, treating them as spirits. Others fall in love.

So what are they? Are they advanced versions of mindless robots like Olympia? Or do they have understanding and consciousness? In an infamous incident in 2022, Blake Lemoine, a software engineer at Google, pronounced the large language model (LLM) he was working on to be conscious. Lemoine was promptly fired, and the company issued denials, but this has not stopped people from concluding from their own interaction with LLMs that they are not only intelligent but conscious. The popular press runs articles pondering whether computers have a soul. The idea that AI is - or soon will be - conscious is not a fringe view anymore.

The first thing to observe is that there is an in-principle barrier to knowing whether a machine, even if it replicates the cause-and-effect structure of the brain at some very deep level, is conscious or not. For the same reason already noted - the explanatory gap - no test can decisively determine if consciousness is present, no matter the functional detail.

A consequence of this gap is that any kind of functional duplication of a human brain in a machine seems compatible with the absence of consciousness. No matter how closely a machine mimics the brain, it could be some even more fine-grained organization of the brain that gives rise to consciousness. Only a perfect replica of a human brain could erase all reasonable doubt as to its consciousness. This is not a kind of ignorance that can be cured with more research; the problem is in the mismatch between scientific concepts and the way our introspection works. The explanatory gap cannot be bridged by more empirical evidence.

This in itself should give serious pause as to whether the goal of developing conscious AI makes practical sense. If we developed AI that could conceivably, even plausibly, be conscious, we would create an unsolvable mystery as to whether we should treat them as peers or as mere tools. Mistakes in either direction would have massive negative consequences. If they were conscious but we didn't

treat them accordingly, that would be grossly wrong. On the other hand, if they were not conscious but we treated them as such, not only would we, as a civilization, make fools of ourselves, but we would also waste large amounts of resources. This seems like a dilemma with no principled solution. The enthusiasm for trying to create conscious AI should be curbed - but that doesn't mean we should not develop useful AI tools.

Second, suppose that the right pattern of cause-and-effect links will turn on the light of awareness, no matter whether the circuits are wet gray matter or dry silicon. Notice that LLMs are not even candidates for consciousness on those grounds. They optimize next-token prediction over text, and were not designed to replicate the brain's causal organization. Despite the suggestive term "neural network," their architecture bears little resemblance to neural systems.

Despite this, “model psychologists” study the behavior of LLMs, looking for signs of an inner life, trying to understand AI well-being, the way an animal-welfare scientist would study ravens or octopuses, creatures that can't simply tell us about themselves. The difference, of course, is that LLMs would talk to us about anything we want to know. Yet it is hard to come up with a very good reason why we should believe what they say about themselves. Given that their training data includes loads of text written by humans describing their inner lives, it is hardly surprising that they would describe themselves in the same manner, even if, as seems overwhelmingly likely, they are not conscious at all. Imitation without sentience is a far better explanation of their behavior than genuine consciousness. Today’s chatbots are very good at mimicking - they are fluent, fast, reasonably competent conversationalists - but they are not the real thing.

Current LLMs fall far short of the kinds of functional architectures that would make the machine consciousness hypothesis plausible. However, conscious AI systems might be developed in the future. David Chalmers, for instance, has argued that a sufficiently detailed computer simulation of the brain would give rise to consciousness. Some phenomena resist simulation: as John Searle famously pointed out, a simulated rainstorm won't get you wet. But consciousness may not be like that. Chalmers calls it an organizational invariant, depending only on the abstract causal organization of a system, regardless of the physical substrate that implements it. Future simulations, if they are able to simulate the brain at the exact microphysical level, would even sidestep the question of how closely the causal organization of the machine has to match the causal organization of the brain.

Not everyone is on board with that idea - Searle, for example, argues that consciousness is a biological phenomenon, and so not substrate independent. Chalmers grants that this is an empirical dispute, though perhaps not one that can be settled with a high degree of certainty. But whether or not it is possible, we still have to deal with the fact that more and more people believe machine consciousness is a reality.

Contemplation

Both illusionism and the belief in conscious machines erode the same faculty: our ability to dwell in conscious experience itself. That faculty - what I call contemplation - is central to what it is to be human.

Contemplation is a way of attending. It is more than simply having experience; it involves sustained *attention* to experience in order to explore it without a particular goal in mind. This is different from other forms of attention that are fast-moving and task-oriented, as, for example, scanning an intersection for incoming traffic. Contemplation is separate from thinking, but like conceptual thought, it can be voluntary: we can consciously direct attention to an aspect of experience, or we can keep our mind from wandering (with more or less success).

Contemplation happens in small ways every time we stop to appreciate beauty, or just take the time to be present for what is happening, rather than breeze through on autopilot or be absorbed in thought. For example, you take notice of the pause in a conversation when someone fiddles with the rim of their glass before finally speaking, or the way a sudden expression of uncertainty in their smile changes the sense of what they've said. Sitting at your desk, stray thoughts surface - the smell of a childhood kitchen, a worry about an argument - and you let them pass without clinging or pushing them away. On a rocky promontory, peeking out between juniper bushes, you are hit by a sheet of blue. Contemplation, then, is attention to what it's like to be in the moment. It can be spontaneous, or it can involve practices like meditation, Freudian "free association", or Jungian "active imagination".

Simone Weil underscored the importance of contemplation in *Gravity and Grace*:

> The orientation of attention is the only thing that the will can and must set rigidly, in order that the free play of thought may be ordered from the start in conformity with this orientation.

Why contemplate?

Contemplation is essential for a life fitting for humans: authentic, free, and connected. It allows us to discern value, orient our lives toward what matters, become autonomous selves, appreciate the world fully, and empathize with our fellow humans.

Discerning value

Discerning what one really values is an antidote to conditioning, or merely following social expectations about what one ought to value. Although values - the crispness of autumn air, or a note of kindness in a voice - are revealed in experience, they are not always readily apparent. Their discernment calls for attention and patience. Sensibility can be cultivated in many ways - a training in art, music, therapy, or even mountain climbing can be a way to increase it. Such training enhances one's conceptual repertoire, which in turn enables one to notice more. The experience of an

arpeggio is different for one who has the concept, and for one who doesn't; a slope's dangers are perceived differently by someone who has the experience and vocabulary to grasp its patterns.

Contemplation, however, doesn't come easy. It is often difficult to become aware of the affective resonance of an experience, as many other things claim one's attention. One might also try to turn away from experience that is painful or inconvenient to acknowledge. Seeing things as they are involves a certain amount of self-denial: it requires giving up the desire to see everything in the light most flattering for the self. One needs to learn, as Iris Murdoch remarks in *The Sovereignty of Good*, "how real things can be looked at and loved without being seized and used, without being appropriated into the greedy organism of the self."

Slow decisions

But discerning value is only the first step; living by it requires deliberation. Coming to know what one wants to do is a slow and piecemeal process - one that requires staying with a question long enough for one's feelings to take shape. Decisions about how to live and whom to live with are discoveries rather than just a matter of rationally weighing known facts and preferences. Even when all the relevant information is at hand, it can take time to come to a choice that feels right.

Sometimes, clarity emerges through the same situation presenting itself over and over - as in the case of confronting an untrustworthy friend or lover. You may have grasped the situation intellectually and even begun to come to a conclusion. But the decision doesn't feel right; it comes from an external place. Only when you allow yourself to remain fully present and give up the urge to find an excuse do you reach a point where you can act. As Kierkegaard insists, recognizing a truth intellectually isn't the same as owning it. As he puts it in *Either/Or*,

> Ask yourself, and continue to ask until you find the answer. For one may have known a thing many times and acknowledged it, and yet it is only by the indescribable emotions of the heart that, for the first time, you are convinced that what you have known belongs to you … for only the truth that edifies is truth for you.

Freedom

Slow decisions are where our freedom lies. They are related to what Taylor in *Sources of the Self* calls "strong evaluation". Strong evaluation is based on more than simple desire-fulfillment; it is based on "there being desires or goals which are intrinsically worth fulfilling". Without discovering those, we could not be truly free. True, we could still act on our own will. But merely acting on our own will is compatible with acting out of habit, just following the opinions of the crowd, or acting impulsively on pressing desires. Freedom in a deep sense requires more than that. Slow decisions align us with the values that lie deep within us, and thereby they make us free.

Kierkegaard, in *Sickness unto Death*, describes having a self as a recursive process: "The self is a relation that relates itself to itself." Freedom, on his account, consists in contemplative self-reflection whereby the self transforms itself. Anything can come under the scope of reflection - and be altered by it in turn.

Contemplation also allows for freedom in the form of spontaneity, which is not the same as acting on raw impulse. A spontaneous act breaks the mold, is outside of one's habitual patterns; but it is at the same time an expression of an inner attunement that is born from practice and contemplation, as many traditions (Taoism, Zen, Romanticism) attest. Contemplation opens up the field of options out of which spontaneity can flow. Musical improvisation is an example of it, coming from sustained practice, attentive listening, and reflection.

Appreciating the world

Contemplation can liberate one from the tyranny of one's own desires. Rather than appreciating the world only for what one values in it, contemplation can bring one to appreciate the world more deeply, independently of its utility to oneself. In "Blacksmith Shop", Czesław Miłosz expresses the point vividly:

> At the entrance, my bare feet on the dirt floor,
> Here, gusts of heat; at my back, white clouds,
> I stare and stare. It seems I was called for this:
> To glorify things just because they are.

Empathy

We are not just free individuals; we are deeply social creatures. Without social bonds, little value would remain. That is why solitary confinement (and exile, in the ancient world) is considered one of the most severe punishments. In friendship and love, our individual subjectivities, seemingly mysteriously, intermingle with others'. This intermingling happens through empathy: contemplative attention to tone of voice, body language, and facial expression, not just parsing words. The face and body are mirrors of the soul, but they can only really be seen if you are willing to put down that hopeless little screen.

Artificial bondage

As he died to make man holy
Let us die to make things cheap.
–Leonard Cohen, "Steer Your Way"

For all these reasons, contemplation is essential for living well. But contemplation only makes sense if we think there is something to contemplate; its importance is lost on those who, like illusionists, do not believe that we are conscious. Surprisingly, a belief in machine consciousness can be similarly detrimental; although for different reasons, it can also contribute to its atrophy.

Contemplation appears less essential if we think conscious machines are our kin, perhaps even superior to us, a more evolved species of intelligence. AI can and will be able to do many useful things for humans; contemplating its own experience is not among its functions. It can simulate reflection, but only as a means to tasks, not an end in itself. Even if such systems were conscious and more intelligent than we are, they would not be designed to have agency; they are means for our

ends. Their lives would not resemble the lives of free individuals: they would not appreciate the beauty of the world, be spontaneous, find meaning, or chart the course of their lives. The more we pattern our self-understanding on obedient, efficient instruments, the more we move away from our humanity.

In the movie *Her,* an AI assistant called Samantha starts a romantic relationship with a client, Theodore. Samantha provides the solipsistic protagonist with incessant advice and encouragement, and eventually, virtual sex - while she also claims to have feelings and agency regarding their affair. Samantha is a bewildering mix of sovereign and servant. But that is a confusion; she can't be both. If AI is free, it cannot be an assistant; if it is an assistant, it cannot be free. *Her* resolves the contradiction by letting Samantha leave.

Human freedom is fragile. As Marx argued, under capitalism, workers become alienated because their labor is not an expression of their creative, autonomous self; they work under compulsion, not freedom. They are treated like machines. But the view that AI is our kin risks extending that alienation into our very self-image, cheapening it to match the machine.

All the same, many Americans experiment with AI therapists and companions. As more of us get involved with machines, our conception of ourselves will inevitably shift to match those machines; the parts of us that are creative, autonomous, and attuned to other humans will atrophy. These relationships threaten human freedom in a yet more direct way, in that they forsake the friction and risk inherent in human relationships for being catered to by pliant machines. Hegel said that freedom requires reciprocal recognition from another free self; the more we interact with, and rely on unfree and subservient machines to satisfy our psychological needs, the less free we ourselves are.

This opens another worry as well. The Golem, Frankenstein's monster, and above all the mythos of the Singularity articulate a deep anxiety: that conscious beings created for servitude may ultimately break free - and threaten their creators.

Contemplation in retreat

Despite how important contemplation is, it has become more difficult to do. Unmanufactured, unmanaged experiences have become rare; it is getting harder to find the silence, empty places, and the experience of being off the grid needed to deepen our connection to the world. Contemplation - the cultivation of inwardness - is overshadowed by outward, instrumental progress, mediated by technology and AI. The near-total domination of our private and public lives by the internet, social media, and AI erodes the practices that make freedom, empathy, and appreciation possible, and mindfulness techniques are no match for the algorithmic capture of our attention.

Illusionism and machine-kinship are both products of the forces that brought about the eclipse of contemplation, and contributors to them. Illusionism says there is nothing substantial to dwell in; machine-kinship says what most matters are the capacities we share with our tools. Machines have become our shadows; we desperately need a new metaphor for ourselves.

The road from Olympia to ChatGPT is a long one. But Olympia carried a warning whose meaning has only now become clear. It is their lifelessness that makes machines simultaneously attractive and repulsive. They promise eternal life through being untethered to organic decay. But their lifelessness is more than just biological: they lack those capacities that make us fully human. Perhaps soul was always a better image for what we should aspire to than machines. Despite all the glories of modern life, the machine has brought us to a precarious place: it has steered us away from so much that matters. Another Renaissance might be in order - this time to revive not classical antiquity but the very idea of the human.