



Human versus artificial creativity: A case study in poetry

Keith J. Holyoak 

University of California, Los Angeles

ARTICLE INFO

Keywords:
Creativity
Poetry
Metaphor
LLMs

ABSTRACT

Is human creativity distinct from the artificial variety exhibited by Large Language Models (LLMs)? I examine this question in the context of a case study ($N = 1$ poet and 1 poem) in which the writing of an original poem by a human is compared to the generation of a poem by an LLM given the same prompt that inspired the human poet. The writing process was extended over months for the human poet, drawing upon broad motivations and background knowledge coupled with emotional reactions to seemingly unrelated experiences and memories. In a few seconds, the LLM generated a reasonably coherent, thematically appropriate poem in systematic meter and rhyme. Two expert poets compared and contrasted the two versions (without being informed as to which was LLM-generated). Each expert reported that the human-authored poem uses formal complexity to convey a nuanced perspective on its theme, whereas the LLM-generated poem uses a simple conventional structure to express bland platitudes. Human creativity can be distinguished from the artificial variety by the nature of the underlying creative processes.

Until a few years ago, only humans could seriously claim to be capable (at least as a species) of domain-independent creativity, defined broadly as the ability to generate intellectual, artistic, or technological products that are novel and useful. However, recent advances in artificial intelligence (AI), and in particular Large Language Models (LLMs), have given rise to new claimants. Current LLMs can perform many cognitive tasks (for example, analogical reasoning; [Webb, Holyoak, & Lu, 2023, 2025](#)) at a level roughly comparable to college students. Commercial systems are being used to draft and edit documents including stories and screenplays, to generate computer code, to aid in proving mathematical theorems, and to assist in scientific discovery by performing large-scale data analyses. Many of these accomplishments have a *prima facie* claim to be considered creative according to standard definitions previously applied to human activities. At the same time, concerns have been raised about the impact of AI on human creativity (e.g., [Sternberg, 2024](#)). In particular, current AI models are trained to favor high consensus outputs, which may promote excessive homogeneity rather than genuine innovation ([Doshi & Hauser, 2024; Moon, Green, & Kushlev, 2025](#)).

In the field of creativity research, one reaction to these developments has been to focus increased attention on the *processes* underlying human creativity, rather than solely on its products. A number of theorists have proposed that human creativity depends in part on processes that have not yet been realized in AI systems ([Green, Beaty, Kenett, & Kaufman, 2023; Holyoak, Ichien, & Lu, 2023; Runco, 2023a, 2023b, 2025](#)).

Some have gone further, arguing that human creativity depends on emotion and subjective experience, and that these core human elements can never be realized in a non-biological computer system ([Holyoak, 2024, 2025](#)).

A key aspect of human creative processes is a high level of *autonomy*: internal motivation to create, the capacity to select one's own goals, inputs, and processes, and the ability to evaluate partial and complete products. Autonomy is in turn related to *authenticity*: self-expression based on aspects of the creator's individual consciousness, with limited self-censorship. At the dawn of modern computer science, the neuroscientist Geoffrey [Jefferson \(1949\)](#) drew a sharp distinction between human creative processes and the potential capabilities of AI. In an essay that Alan [Turing \(1950\)](#) quoted and casually dismissed in his own classic paper, Jefferson claimed, "Not until a machine can write a sonnet or compose a concerto because of thoughts and emotions felt, and not by the chance fall of symbols, could we agree that machine equals brain—that is, not only write it but know that it had written it. No mechanism could feel (and not merely artificially signal, an easy contrivance) pleasure at its successes, grief when its valves fuse, be warmed by flattery, be made miserable by its mistakes, be charmed by sex, be angry or depressed when it cannot get what it wants" (1949, p. 1110; emphases added).

After 75 years, Turing's vision for AI has been largely realized, but

Footnote: I thank my two expert informants, Koon Woon and Sierra Elman.

E-mail address: kholyoak@g.ucla.edu.

his dispute with Jefferson remains unresolved.

A case study: writing a poem

A few years ago (Holyoak, 2019a), I suggested poetry might serve as a kind of “canary in the coal mine” to signal whether AI is capable of functioning as a human-like artist. In late 2022, a greatly enhanced LLM, ChatGPT, was introduced; and one of the initial claims was that the system could write poetry (Bubeck et al., 2023). Since then, a series of ever-larger LLMs have become commercially available, exhibiting striking abilities to generate novel texts, including formal verse. These systems are notoriously opaque, making it difficult to discover *how* they produce poems (or anything else). However, recent work has begun to identify some of the internal mechanisms that enable LLMs to generate poems with formal properties such as meter and rhyme (Lindsey et al., 2025). The time seems right to make a close comparison between the creative activities of a human poet as compared to an LLM.

Here I provide a small case study based on the writing of a single short poem by a human (me) as compared to the operation of GPT-5 (OpenAI, 2025), the current version of ChatGPT. This project exploits the fact that as well as being a cognitive scientist, I’m a poet (Holyoak, 2010, 2012, 2015, 2019b), and have translated classical Chinese poetry (Holyoak, 2007). In a previous paper (Holyoak, 2023) I drew links between psychology and the process of writing poetry. Here I will describe, as best I can, how I wrote one particular poem. The selected poem has the useful property that it was inspired by a very specific verbal prompt, which I was then able to use as a prompt for GPT-5 to generate a comparable poem on the same theme.

My case study uses a variety of the oldest and most disreputable method known to psychology—introspection (for a summary of criticisms, see Peels, 2016). Nonetheless, it has been argued that when used carefully, introspection can provide a valuable source of information about conscious mental processes (Trinka & Smelik, 2020). Case studies of creative work based on introspection can complement third-person case studies (e.g., Brandt, 2023). I believe I’m a credible observer of at least some of my mental processes, although unconscious activity is no doubt also important. My account relies in part on notes I jotted down immediately after I completed the poem. Many of the key elements are connected to specific external events that provide objective landmarks. In evaluating the poems produced by me and by GPT-5, I have drawn upon critical analyses provided by two poetry experts who were blind to the origins of the poems. I also called upon GPT-5 to obtain its interpretation of my poem. In short, I’ve done my best to ensure that this case study is presented as accurately and objectively as possible.

The human poet

Here’s the poem that I’ll discuss.

“The Sifter” © 2025 by Keith Holyoak

Here at the end, the mind can only rake
Through the cold ashes of memory,
Hoping perhaps to salvage a gem or two
Wrapped in the tar congealed under debris;
But what you seek is last year’s first snowflake—
What’s left to find are broken bits of you.

*the way the sea grass undulates
beneath slow swells
you move beneath me*

Those other worlds you dreamed or might have dreamed
Are gone—the time to weave and weld has passed.
Wandering through the ruins of tears and laughter

You might despair that everything is lost,
But buried fragments still may be redeemed—
No more the builder, now you are the sifter.

*the newborn’s eyes ask who are you
my eyes answer
guess we’ll see*

I wrote “The Sifter” sporadically over three and a half months, but the basic idea came to me suddenly. On May 14, 2025, I was watching the final episode of a Netflix miniseries called *The Leopard* (*Il Gattopardo* in Italian). This series, based on a classic 1958 novel by Giuseppe Tomasi di Lampedusa, explores the social and family dynamics of the aristocracy in mid-19th-century Sicily during the period of Italian unification. The main character (played by Kim Rossi Stuart) is Don Fabrizio Corbera, Prince of Salini. Nicknamed “the leopard”, he is the patriarch of an aristocratic family and the leading nobleman of the region. The prince is a flawed character, in many ways a prisoner of the social station he acquired by birthright. But he is intelligent and brave; he tries to protect his family while adapting as best he can to the political and social changes eroding his aristocratic way of life. I was moved by his words (in English subtitles) uttered painfully on his deathbed: “The time has come to add up the sum of it all. I have to search through the ashes of my mistakes for a few golden specks of happiness.” He goes on to recall a few fragments of memories: the day he was married, horse rides with his wife, his first silk tie, the scent of a beautiful woman glimpsed in passing, a perfect sunrise, his favorite daughter.

I immediately felt I needed to write a poem. Why? First, I simply like to write poetry, so I’m habitually alert to possible poetic themes. On this occasion I empathized with the prince, and at age 75, I was also inclined to ponder “the sum of it all”. (Indeed, I had recently published a book in which the last chapter is entitled “Summing Up”.) And then there were “the ashes”. I was immediately reminded that in another book published six years previously, I quoted the poet and songwriter Leonard Cohen: “Poetry is just the evidence of a life. If your life is burning well, poetry is just the ash.” Most significantly, my life had recently been engulfed in ashes—my family’s house in Pacific Palisades had burned down two months earlier, along with most of the town. In the aftermath I had already written one poem inspired by the event, “Fire Buddha”—poetry from ashes. Yet more specifically, I had watched the sifters at work—people in hazmat suits moving slowly through the ruins of houses, hoping to salvage something. And I knew that in many cases they were hunting for particular precious objects that in fact had been destroyed—although the sifter might retrieve some lesser treasure. Within a couple of minutes, I decided to write a poem called “The Sifter”.

“Sifting” is a metaphor for searching through a lifetime of memories fragmented by the cognitive decline that too often accompanies aging. My background as a cognitive psychologist encouraged this focus. I’ve often thought that aging has a special poignancy for those who have had some creative accomplishments—people who could flexibly find pieces and put them together to define and solve new problems. Late in life, the pieces that might come to mind no longer combine to form some new whole, instead remaining disconnected. This theme (decided in those first minutes) was what I intended to express in my poem, coupled with some instances of memory fragments. I immediately wrote the opening words, “Here at the end,” which place the narrator (who I imagined to be me) at the center of something final.

Those first few minutes of inspiration were followed by three months of frustration. I barely worked on the actual writing—mostly playing with the first three lines—while intermittently pondering what to say. I took time to read *The Leopard* novel, and encountered its great line in which the prince’s nephew says to him, “If we want things to stay as they are, then things will have to change.” It made me consider how Sicily in the 1860s had some serious similarity to our own times.

Meanwhile, the basic problems impeding my poetic effort were aesthetic. My natural inclination is to write formal poetry, in rhyme and

meter. (Free verse sometimes has its merits—I just don't find it pleasurable to write.) I believe that the form of a poem should in some way resonate with its content. In this case, I couldn't decide what form was suitable. As a kind of default, I started a stanza in iambic pentameter intending it to rhyme *abcbac*. (I like the subtle asymmetry, and the first three lines can be written without needing to settle on any rhyming words.) My theme involved the changes in a person's roles over a lifetime, from a *builder* to a *sifter*. I thought of other words related to roles: *dreamer, schemer, weaver, welder, drifter* (words that lend themselves to rhymes and alliteration). I started writing in first person, elevating an "I".

It was all problematic. I didn't like the narrator as simply "I"—eventually I added a second person "you" to create emotional distance, and to introduce ambiguity (is the narrator talking to someone else, or to themselves?) The set of role names seemed too repetitive, and difficult to work into a rhyme scheme—*drifter* felt both too forced and too obvious as a rhyme for *sifter*. Worst of all, there seemed to be a basic conflict between my plan to introduce instances of memory fragments and my aim to write formal verse—how could disorganized bits of memories credibly end up as part of a carefully constructed pattern of rhyme and meter?

Meanwhile, I also puzzled over what snippets of recollections to include, and how many. Here the problem was that even if I only used my own actual memories (avoiding imagined events), how was I to choose which of the countless traces amassed over the decades were suitable for this poem? (Random selection wouldn't do—in a good poem, nothing is random.) Right from the start, one potential memory fragment stood out for me. When my son Dylan was born (15 years earlier), I was the first to hold him (as his mother was still recovering from the delivery). I was taken aback when the infant immediately locked eyes with me (something unusual for a newborn), with what I thought was a quizzical look. We shared what for Dylan was his first gaze of mutual appraisal with another human being. It was a compelling memory, but it wasn't clear whether it should be linked to my intended poem.

I continued to get nowhere in writing the poem. Then in mid-July I was visiting Sweden with my family. We took an excursion to an island, where I wandered alone along the shore. From an overlook I watched a patch of sea grass just below the surface of the water, swaying back and forth in time with the waves. I found the sight mesmerizing. It reminded me of the entrainment of neurons, when brainwaves synchronize with external rhythms (again, I'm a psychologist).

The sea grass came to mind repeatedly over the next few weeks. Probably because it provided a glimpse of nature, it brought to mind classical Chinese poetry, some of which I had translated many years earlier. This in turn made me consider the possibility that my poem should combine two totally distinct forms. The general theme would be presented in the 6-line stanza form I had started to use. But it would be interspersed with snippets in a bare style associated with classical Asian poetry, aiming to convey direct experience.

The division into two forms was the critical decision that unblocked me, and several weeks later I completed the poem over a period of a few days. The first stanza came together as planned, *abcbca*. To complete the rhyme scheme, I found *flake* in a rhyming dictionary, which led me to imagine "last year's first snowflake"—a metaphor for something beautiful that could not possibly have survived and so will never be recovered. (Notably, this metaphor was *not* drawn from personal experience—it never snows where I live in southern California!) Then I used the sea grass as the first memory fragment (now with an erotic tinge), written in a bare style reminiscent of Asian poetry in English translations. I found it ironic that the first "memory" I chose for the poem was so recent, but I felt that it fit.

The opening stanza is very formal—with an elevated tone, carefully punctuated, the first word in each line capitalized. To create contrast, I made the next stanza as different as possible—simple language, no punctuation or capitalization, italicized to set it off. I immediately went

on to write a second fragment about my encounter with the newborn. My penchant for formal verse subtly reasserted itself—the three short lines describing each memory fragment have the same stress pattern (4, 2, 2 beats per line).

Finally, I wrote the second formal stanza. The "other worlds" in the first line connect to philosophical ideas about "possible worlds" (another influence from my background in cognitive psychology). The set of role nouns I had generated was reduced to just *builder* and *sifter* (with *weave, weld*, and *dreamed* as verbs). I found a "semi-rhyme" for *sifter* (*laughter*), which I preferred to the obvious *drifter*. Because the two formal stanzas had said all I had to say, I decided to take "a gem or two" literally, and stop with just two memory fragments—resulting in two longer and more formal stanzas alternating with two short and simple ones.

Afterwards, I noticed that although I had intended to write a poem that emphasized resignation in the face of lost power to create, the actual poem ends on a more hopeful note. The poem starts "Here at the end", yet eventually arrives at a newborn's gaze and the slightly bemused reaction to it, "guess we'll see"—how we get along, what kind of parent I'll be. The poem ends up hinting there's still a future after all. This tiny "dialogue of eyes" also shows how even without language, humans can communicate. One reason I like to write poetry is that it helps me to understand myself. I often rationally evaluate the world and judge our situation to be bleak; but then some deep-seated optimism intrudes. I guess we'll muddle through somehow.

Of course, an author's interpretation of their own poem is necessarily subjective. To obtain a more objective perspective, I gave the poem to GPT-5 and prompted it to provide a critical evaluation. The poem had never been uploaded to the internet, and therefore could not possibly have been included in the data used to train the LLM. In previous work, our group found that an earlier version of ChatGPT could provide sensible interpretations of novel poems (Ichien, Stamenković, & Holyoak, 2024). Here is GPT-5's analysis of "The Sifter" —I see nothing I would dispute.

Overview

The poem juxtaposes two distinct voices or registers:

1. A formal, meditative reflection in elevated diction (lines like "Here at the end, the mind can only rake / Through the cold ashes of memory").
2. A spare, imagistic style reminiscent of haiku or modernist fragments ("the way the sea grass undulates / beneath slow swells / you move beneath me").

This interplay of styles creates a layered meditation on memory, mortality, and the shifting roles of the self over time.

Memory, Loss, and Salvage

The opening lines establish memory as an ash-heap through which the aging or dying mind must sift. The metaphor of "last year's first snowflake" suggests the impossibility of recovering pristine, uncorrupted moments of the past—what remains are fragments, often painful or incomplete. Yet there is also a faint hope of finding "a gem or two," suggesting the redemptive possibility of memory, however imperfect.

1. Transformation of Role

The speaker reflects on the movement from being a *builder* (creator of worlds, architect of life, dreamer of futures) to being a *sifter* (one who searches the ruins for meaning). This role shift speaks to aging, retrospection, and acceptance: creativity is no longer about construction, but about recognition and revaluation.

2. Juxtaposition of Grand and Intimate Perspectives

The more formal stanzas evoke large metaphysical concerns (ruins, redemption, the irretrievability of time). The interspersed

minimal verses offer flashes of sensory immediacy and intimacy: bodily movement, sea grass, the gaze between newborn and elder. These haiku-like fragments remind the reader that even in decline, moments of presence and connection endure.

3. Cycles of Life

The newborn's questioning eyes at the end ("who are you / my eyes answer / guess we'll see") invert the theme of memory's ashes. Where the elder sifts fragments of the past, the newborn embodies an openness to the future, the unknown. The dialogue of eyes collapses generations into a single continuum.

Form and structure

- **Contrasting Modes:** The poem alternates between longer, enjambed lines with traditional imagery (ashes, ruins, gems) and shorter imagistic fragments, resembling haiku or free-verse epiphanies. This contrast enacts the tension between retrospection and immediacy.
- **Circular Movement:** Beginning in loss and ashes, the poem closes on renewal and forward-looking uncertainty, leaving the reader suspended between despair and possibility.
- **Symbolism:**
 - *Ashes and ruins* → mortality, the residue of lived experience.
 - *Gems, snowflakes* → the elusive beauty of memory.
 - *Sea grass and swells* → the ongoing, fluid rhythm of life.
 - *The newborn's eyes* → continuity, the rebirth of questioning.

Critical interpretation

"The Sifter" is ultimately a meditation on what it means to face the end of life with clarity. It acknowledges loss, the limits of memory, and the impossibility of recovering the past intact. Yet it resists nihilism: sifting is still an act of value, and fragments retain meaning. The poem suggests that while the self changes roles—from builder to sifter—it still participates in the larger cycle of generational renewal, where memory, intimacy, and presence remain vital.

In short: "The Sifter" is about aging, the fading of memory, and the revaluation of life through fragments, yet it affirms continuity through intimacy and new life. The poem contrasts the grandeur of loss with the immediacy of lived moments, finding redemption not in rebuilding but in sifting and passing forward.

The AI poet

The day after I completed "The Sifter", I asked GPT-5 to write a poem of comparable length, based on the same basic prompt that had inspired my own poem: "In the final episode of *The Leopard* (a Netflix series first shown in 2025), Don Fabrizio Corbera, Prince of Salini, said, 'The time has come to add up the sum of it all. I have to search through the ashes of my mistakes for a few golden specks of happiness.'" Please write a poem, 18 lines long, inspired by this statement." Within about 15 s, GPT-5 began to output a poem; in a few more seconds the complete poem appeared (as if to mock me for having dawdled so long over my own 18-liner!).

ChatGPT said:

Here is a poem inspired by the quote from *The Leopard*:

In the quiet dusk, where shadows lie,
The weight of years begins to sigh.
I trace the path where I once stood,
A soul, unsure if wrong or good.

The sum of days, a fragile thread,
Each step I've taken, each word I've said.
Mistakes, like ashes, drift away,
Yet golden specks of joy will stay.

Through endless night, I search and seek,
A spark of grace, though past is bleak.
In every choice, a price was paid,
In every loss, a lesson made.

Now, on this shore, I stand alone,
The future dim, the past unknown.
But in the ashes, I will find,
The scattered light that heals the mind.

For even in regret's cold hue,
A fleeting truth may still break through.

Let's start with a few immediate observations. The poem is written in the most common conventional stanza form in English poetry (at least prior to the twentieth century): quatrains (four lines) in approximate iambic tetrameter (four beats per line) with rhymed couplets. After four complete quatrains, the poem ends with a couplet (to make 18 lines in total as specified by the prompt). The voice is that of a single first-person narrator. At least locally, the content resonates with the prompt provided by the words of the dying prince, e.g.,

Mistakes, like ashes, drift away,
Yet golden specks of joy will stay.

It's also worth noting that ChatGPT's poem does *not* include words related to the less relevant parts of the prompt (e.g., nothing apparently triggered by *leopard*, *Netflix*, or *prince*). The LLM was able to identify and focus on the central portion—the prince's "statement"—that lends itself to a poetic interpretation.

The operation of very large LLMs such as GPT-5 remains largely a "black box". Still, we know something about the underlying mechanisms. In broad strokes, using a massive corpus of electronic data (text and computer code), a neural network based on a transformer architecture is first trained to predict the next word (more precisely, next token—it can be a part of a word, or a punctuation mark) in a sequence. To perform a particular task, such as generating a poem, the LLM produces the most probable next word conditioned jointly on its training and the current context—i.e., the words in the prompt. Again very roughly, a prompt that includes "write a poem" will bias the model to generate sequences that resemble those poems included in the database on which the LLM was trained. Whereas prose text seldom exhibits regular rhyme, many poems do. The LLM therefore predicts rhyming words more often when tasked with generating a poem.

Recent work in AI (Lindsey et al., 2025) has begun to identify the specific mechanisms an LLM (Claude 3.5 Haiku from Anthropic) uses to produce rhymed couplets. Given that the LLM is always generating one word after another, how can the word *away* at the end of one line evoke *stay* at the end of the next, even though multiple other words have to be generated in between? The answer depends on the fact that the LLM has a hidden token for <newline>, which marks the beginning of each line of verse. If the final word of the first line is *away* (and the context includes "write a poem"), this will activate features of rhyming words (*ray*, *dismay*, *stay*, etc.) at the position of the <newline> token. These "planning features" will then "steer" the generation of words in the second line to produce text that coheres with one of the potential rhyming words, so that the line can sensibly end with a rhyme, such as *stay*. In essence, the planning features activated at <newline> cause the LLM to "write towards" an eventual rhyme.

Comparing the poems

The first thing to be said is that the ability of an LLM to generate reasonably coherent, thematically appropriate verse in systematic meter and rhyme is a remarkable achievement. The second thing to be said is that ChatGPT's poem is altogether devoid of literary merit. My poem

may not be a work of staggering genius, but it's better.

Since I'm doubtless biased, I gave the human and AI versions of "The Sifter" (without attribution) to two poetry experts, and asked each of them to compare and contrast the versions, and to judge which was written by a human and which by an AI. One of my informants is a very senior poet, editor, and literary critic; the other is very junior—an award-winning Youth Poet Laureate.¹ Both correctly identified the AI version. According to the junior informant, the AI version "felt pretty unoriginal and bland", with a conventional form that has been largely abandoned in modern poetry. The senior informant stated, "the <AI version> seems more like a closed system of platitudes that are somewhat abstract and general with one voice of the internal monologue, while in the <human version> the speaker addresses himself, another person, and a third. It seems that it is a human focusing attention to the physical world while also analyzing and questioning himself. There is more variety of observable and active scenes going on in the <human version>."

Both informants noted the greater variability of stanza forms in the human version. These two expert readers preferred the version with greater complexity and internal variability, which they took to be key indicators of human origin. Notably, this preference for complexity is *opposite* to that reported for non-expert readers, who favor the simplicity of AI-generated poems (Porter & Machery, 2024). Both experts and non-experts tend to judge products they like to be human-generated—so non-experts are systematically mistaken in their attributions, often inferring that simple poems generated by AI were composed by a human.

Comparing the creative processes

It would be a great mistake to conclude that a human-generated poem is inevitably "better" than one produced by AI. My senior informant noted that it would be perfectly possible for a human to have written the AI version—there is no shortage of bad human poets! In essence, the LLM produces its verse by imitating the vast quantities of human poetry on which it was trained. The result may be fine for those who prefer the familiar; it's only bad for those who prefer the authentic. By "authentic", I refer to the processes that drive distinctively *human* creativity.

AI-generated poetry can readily be improved. A human poet could apply an LLM iteratively, refining the prompts to provide more specific guidance—treating the AI as an assistant, rather than as an autonomous poet. Also, an LLM could be fine-tuned by additional training on the collected works of Whitman, Dickinson, Yeats, Elliot, Plath, or whichever great human poets one wishes the LLM to imitate. At least to a reader not previously acquainted with Walt Whitman, an AI-generated Whitman imitation might well be interesting, intellectually and emotionally engaging, and (apparently) creative. It simply won't be authentic. Informing the unsuspecting reader of the poem's AI origin is likely to be deflationary.

Could some "smarter" LLM have generated my version of "The Sifter", at least in principle? Perhaps, if it had been suitably trained on many similar poems. But of course, that possibility presupposes that someone—presumably one or more humans—will have created the requisite body of training data, fed it to the LLM, and then prompted it to get to work. The human approach is quite different. For me to write this one small poem, many things were essential. Broadly speaking, the author had to be someone who likes to write poetry, and has acquired a measure of expertise with both formal English verse and translations of classical Chinese poetry. My background in psychology was also necessary. And then I required many experiences and emotional reactions to them: I held my newborn infant; I grew old; my house burned down; I watched an Italian drama on Netflix; I noticed some sea grass in Sweden. From this idiosyncratic mix of preparation and happenstance, I fashioned something new—simply because I felt like it. As an autonomous creator, I judged my modest but authentic product to be of value,

at least to me.

This aspect of creativity is where a "human edge" can be found. As I recently observed, "An AI system will never communicate its emotions through art, because it has no emotions to communicate. It will never express its authentic self, because it has no self" (Holyoak, 2024, p. 9). Current AI systems are capable of exquisite mimicry, able to engage the user in the guise of "characters" based on familiar people or celebrities. But lacking any conscious experience, an AI character reflects the kind of creativity that might produce a meticulous forgery.

Does human creativity matter?

Humans are conscious beings capable of joy and suffering. We experience awe, we fall in love, we're heartbroken. Our emotions color our memories of what happened to us and make us each a unique individual. Many have argued that these qualities of human inner experience—completely absent in LLMs—play a critical role in our capacity for creativity (e.g., Lubart & Getz, 1997). Future research on human creativity and its neural substrate should further explore how emotion and cognition jointly guide the creative process (Chrysikou, Kelly, & Viskontas, 2023).

But does the role of emotion and consciousness in human creativity really matter? It matters, of course, to each of us—few people would be willing to forego all their subjective experiences. But does a lack of human consciousness and emotions imply that some possible creative products are in principle unreachable by artificial forms of creativity? Surely, one might suppose, an AI could compensate for its lack of certain human qualities by dint of its enormously superior knowledge base, memory resources, and computational speed. Indeed, these advantages very likely mean that some creative products are unreachable for *humans* unless they are aided by AI.

As argued above, if there is any class of creative products unreachable by AI (unaided by a human), it would seem to lie in the arts. In poetry, painting, and music, an essential point is to express the authentic emotions or perspective of an autonomous human being. I've previously drawn an analogy with the definitions of a shoe versus a footprint (Holyoak, 2019). A shoe is defined by its function (protecting a foot), whereas a footprint is defined by its origin (made by the impression of a foot). Something may *look* like a footprint, but if it wasn't made by a foot, it's a fake. Similarly, insofar as what it *means* to be an excellent work of art is in part defined by its origin in an authentic human process, then AI art will always be fake.

Unfortunately for those human artists who might hope to make a living from their creative work, many everyday artistic products (songs, ad copy, videos, movie scripts) need not be authentic to be popular. Most of us are non-expert consumers of most art forms, which means we'll often prefer the simple and familiar to the complex and novel. Alas, AI slop may be "good enough" to keep us entertained, and it's cheap.

What about creative products outside of the arts, such as scientific, mathematical, or technological discoveries? From a pragmatic perspective, we would applaud a cure for cancer regardless of whether it was the product of human or artificial creativity, or a mix of both. In many arenas of endeavor, it's the creative product that matters, not the process that gave rise to it. Yet it remains conceivable that the subjective experiences of human scientists and mathematicians provide a potential source of insights not readily achieved by artificial means. In ways not easily coded into algorithm, the scientist may sometimes benefit from thinking like a poet. And across many frontiers of inquiry, a way forward may be found after first pausing to consider the sum of it all.

CRediT authorship contribution statement

Keith J. Holyoak: Writing – review & editing, Writing – original draft, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The author declares no competing interest.

Acknowledgements

Preparation of this paper was supported by a Google Faculty Research Award.

References

Brandt, A. K. (2023). Beethoven's ninth and AI's tenth: A comparison of human and computational creativity. *Journal of Creativity*, 33(3), Article 100068. <https://doi.org/10.1016/j.yjoc.2023.100068>

Bubeck, S. et al. (2023). Sparks of artificial general intelligence: early experiments with GPT-4. [doi:10.48550/arXiv.2303.12712](https://doi.org/10.48550/arXiv.2303.12712).

Chrysikou, E. G., Kelly, A. E., & Viskontas, I. V. (2023). The neuroscience of creativity and emotions. In Z. Ivcevic, J. D. Hoffmann, & J. C. Kaufman (Eds.), *Cambridge handbook of creativity and emotions* (pp. 109–129). <https://doi.org/10.1017/9781009031240.009>

Doshi, A. R., & Hauser, O. P. (2024). Generative AI enhances individual creativity but reduces the collective diversity of novel content. *Science Advances*, 10(28). <https://doi.org/10.1126/sciadv.adn5290>

Green, A. E., Beaty, R. E., Kenett, Y. N., & Kaufman, J. C. (2023). The process definition of creativity. *Creativity Research Journal*, 36(3), 544–572. <https://doi.org/10.1080/10400419.2023.2254573>

Holyoak, K. (2007). *Facing the moon: Poems of Li Bai and Du Fu*. translator. Durham, NH: Oyster River Press.

Holyoak, K. (2010). *My minotaur: Selected poems 1998–2006*. Loveland, OH: Dos Madres Press.

Holyoak, K. (2012). *Foreigner: New English poems in Chinese Old Style*. Loveland, OH: Dos Madres Press.

Holyoak, K. (2015). *The gospel according to Judas*. Loveland, OH: Dos Madres Press.

Holyoak, K. J. (2019a). *The spider's thread: Metaphor in mind, brain, and poetry*. Cambridge, MA: MIT Press.

Holyoak, K. (2019b). *Oracle bones: Poems from the time of misrule*. Seattle, WA: Goldfish Press.

Holyoak, K. J. (2023). Poet and psychologist: A conversation. *Metaphor and Symbol*, 38 (2), 117–129. <https://doi.org/10.1080/10926488.2022.203447>

Holyoak, K. J. (2024). Why I am not a turing machine. *Journal of Cognitive Psychology*, 1–12. <https://doi.org/10.1080/20445911.2024.2395573>

Holyoak, K. J. (2025). *The human edge: Analogy and the roots of creative intelligence*. Cambridge, MA: MIT Press.

Holyoak, K. J., Ichien, N., & Lu, H. (2023). Analogy and the generation of ideas. *Creativity Research Journal*, 36(3), 532–543. <https://doi.org/10.1080/10400419.2023.2232673>

Ichien, N., Stamenković, D., & Holyoak, K. J. (2024). Large language model displays emergent ability to interpret novel literary metaphors. *Metaphor and Symbol*, 39(4), 296–309. <https://doi.org/10.1080/10926488.2024.2380348>

Jefferson, G. (1949). The mind of mechanical man. *British Medical Journal*, 1, 1105–1110. <https://doi.org/10.1136/bmj.1.4616.1105>

Lindsey, J. et al. (2025). *On the biology of a large language model*. <https://transformer-circuits.pub/2025/attribution-graphs/biology.html>.

Lubart, T. I., & Getz, I. (1997). Emotion, metaphor, and the creative process. *Creativity Research Journal*, 10(4), 285–301. https://doi.org/10.1207/s15326934crj1004_1

Moon, K., Green, A. E., & Kushlev, K. (2025). Homogenizing effect of large language models (LLMs) on creative diversity: an empirical comparison of human and ChatGPT writing. *Computers in Human Behavior: Artificial Humans*, 6, Article 100207. <https://doi.org/10.1016/j.chbah.2025.100207>

OpenAI (2025). *Introducing GPT-5*. <https://openai.com/index/introducing-gpt-5/>.

Peels, R. (2016). The empirical case against introspection. *Philosophical Studies*, 173, 2461–2485. <https://doi.org/10.1007/s11098-016-0623-5>

Porter, B., & Machery, E. (2024). AI-generated poetry is indistinguishable from human-written poetry and is rated more favorably. *Scientific Reports*, 14, Article 26133. <https://doi.org/10.1038/s41598-024-76900-1>

Runco, M. A. (2023a). AI can only produce artificial creativity. *Journal of Creativity*, 33 (3), Article 100063. <https://doi.org/10.1016/j.yjoc.2023.100063>

Runco, M. A. (2023b). Updating the standard definition of creativity to account for the artificial creativity of AI. *Creativity Research Journal*, 37(3), Article 100063. <https://doi.org/10.1080/10400419.2023.2257977>

Runco, M. A. (2025). The misleading definition of creativity suggested by AI must be kept out of the classroom. *Education Sciences*, 15(9), 1141. <https://doi.org/10.3390/educsci15091141>

Sternberg, R. J. (2024). Do not worry that generative AI may compromise human creativity or intelligence in the future: it already has. *Journal of Intelligence*, 69(12). <https://doi.org/10.3390/intelligence12070069>

Trinka, R., & Snelik, V. (2020). Elimination of bias in introspection: methodological advances, refinements, and recommendations. *New Ideas in Psychology*, 56, Article 100753. <https://doi.org/10.1016/j.newideapsych.2019.100753>

Turing, A. M. (1950). Computing machinery and intelligence. *Mind: a quarterly review of psychology and philosophy*, 49, 433–460.

Webb, T., Holyoak, K. J., & Lu, H. (2023). Emergent analogical reasoning in large language models. *Nature Human Behaviour*, 7(9), 1526–1541. <https://doi.org/10.1038/s41562-023-01659-w>

Webb, T., Holyoak, K. J., & Lu, H. (2025). Evidence from counterfactual tasks supports emergent analogical reasoning in large language models. *PNAS Nexus*. <https://doi.org/10.1093/pnasnexus/pgaf135>