

Seeing Double: The Science Behind Gestalt Images That Change Before Your Eyes



Greg Robison

Follow

8 min read

Mar 27, 2025

Listen

Share

“The whole is other than the sum of the parts.”

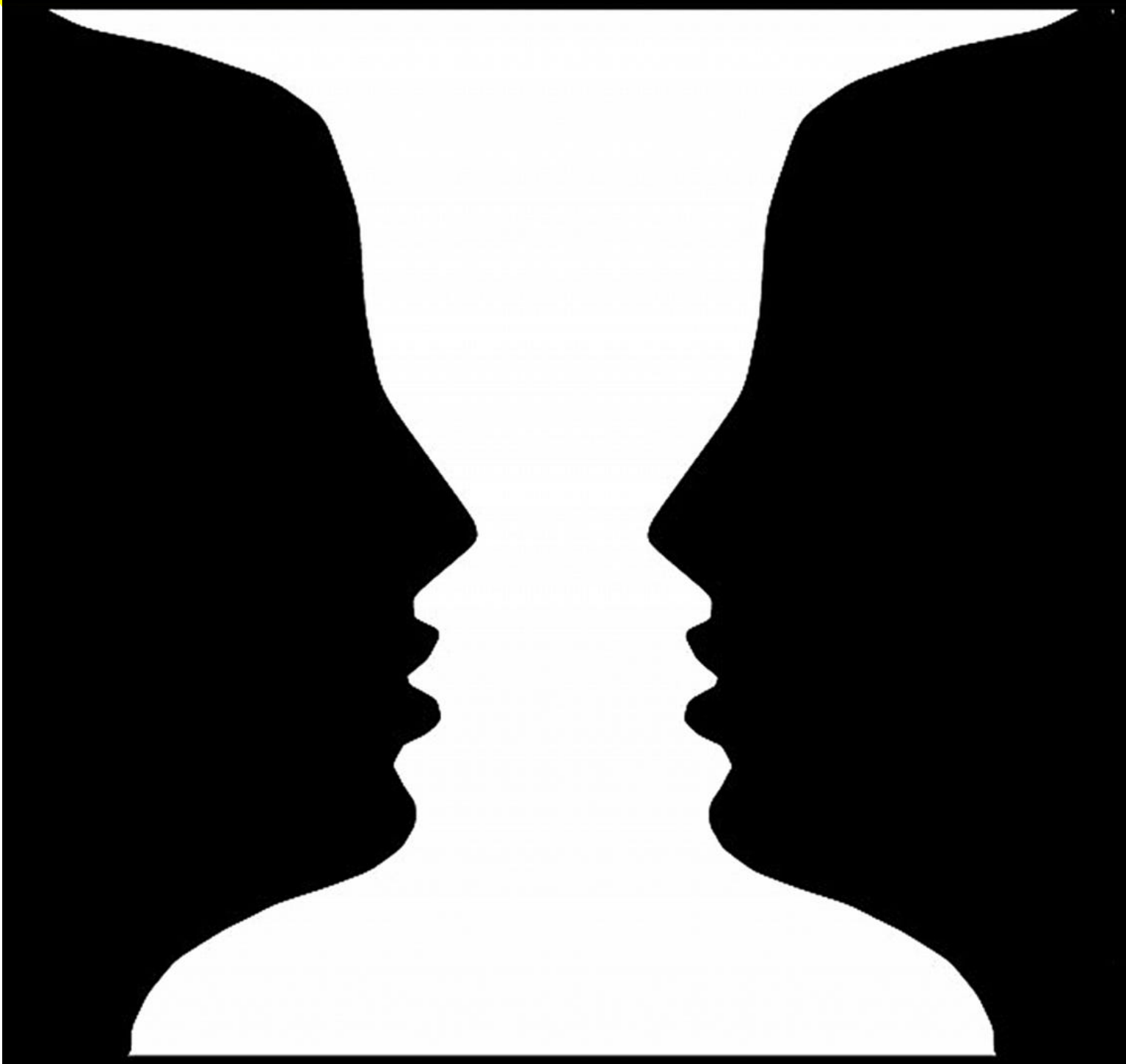
- Max Wertheimer

Is it a young woman or an elderly lady? A duck or a rabbit? A vase or two face profiles? I’ve always been fascinated by these optical illusions called Gestalt images (also known as bistable or ambiguous figures), which are visual illusions that can be perceived in two distinctly different ways, though somehow never simultaneously. Strange, right? What makes these images so fascinating isn’t just their clever design, but what they reveal about human perception itself. Each time your brain “flips” between seeing one interpretation and then another, you’re witnessing your visual system actively constructs, rather than passively records, reality. These perceptual toggles show how our minds continually organize, interpret, and even struggle with visual information, demonstrating that what we “see” is as much a product of our brain’s interpretive work as it is of the light entering our eyes. And can modern “vision” AI models perceive both sides? Let’s find out!

The Gallery of Visual Flip-Flops

The fascinating world of bistable illusions features several iconic examples that have captivated viewers for generations. Rubin’s Vase, introduced by Danish psychologist Edgar Rubin in 1915, presents an elegant black-and-white silhouette that toggles between a symmetrical vase and two face

profiles gazing at each other.



The Duck-Rabbit illusion, which first appeared in a German humor magazine in 1892 and was later popularized by psychologist Joseph Jastrow, features a simple line drawing that can be seen as either a duck facing left or a rabbit looking right.



The Necker Cube, described by Swiss crystallographer Louis Albert Necker in 1832, is a wireframe drawing of a cube that spontaneously appears to flip between two different spatial orientations.



Perhaps most delightful is the “My Wife and My Mother-in-Law” illusion created by cartoonist W.E. Hill in 1915, which alternates between depicting a young woman turning away and an elderly woman in profile.



What makes these images particularly interesting is their stubborn refusal to be seen in both ways simultaneously. Despite knowing that two valid interpretations exist, our brains still can commit to processing only one version at a time. This phenomenon, known as perceptual bistability, occurs because these images carefully balance visual cues that normally help us determine figure from ground

or orientation in space. When viewing Rubin's Vase, for instance, your brain must decide which region to treat as the "figure" (the object) and which to relegate to "ground" (background). The contours between black and white sections can belong to either the vase or the faces, but never both at once, forcing your visual system into a perpetual tug-of-war between equally valid interpretations.

Why These Images Mess With Your Mind

Our brains are constantly performing the critical task of figure-ground organization, separating what's important (the figure) from what's less important (the background). Normally, this process happens effortlessly thanks to visual cues from the context like contrast, size, surroundedness, and convexity that help us determine what to focus on. Gestalt images subvert this system by balancing these cues, creating what psychologists call "bistable stimuli." In Rubin's Vase, for instance, both the white center shape and the black surrounding shapes can equally claim ownership of the shared border. Your brain must assign this border to one interpretation or the other, but the usual visual shortcuts that would help make this decision have been deliberately neutralized. The result is that your visual system flips between interpretations, constantly reassessing which parts should be figure and which should be ground.



An image called "Einstein's Gorge" I created using GenAI to be both a landscape image and Einstein famously sticking out his tongue.

Get Greg Robison's stories in your inbox

Join Medium for free to get updates from this writer.

Subscribe

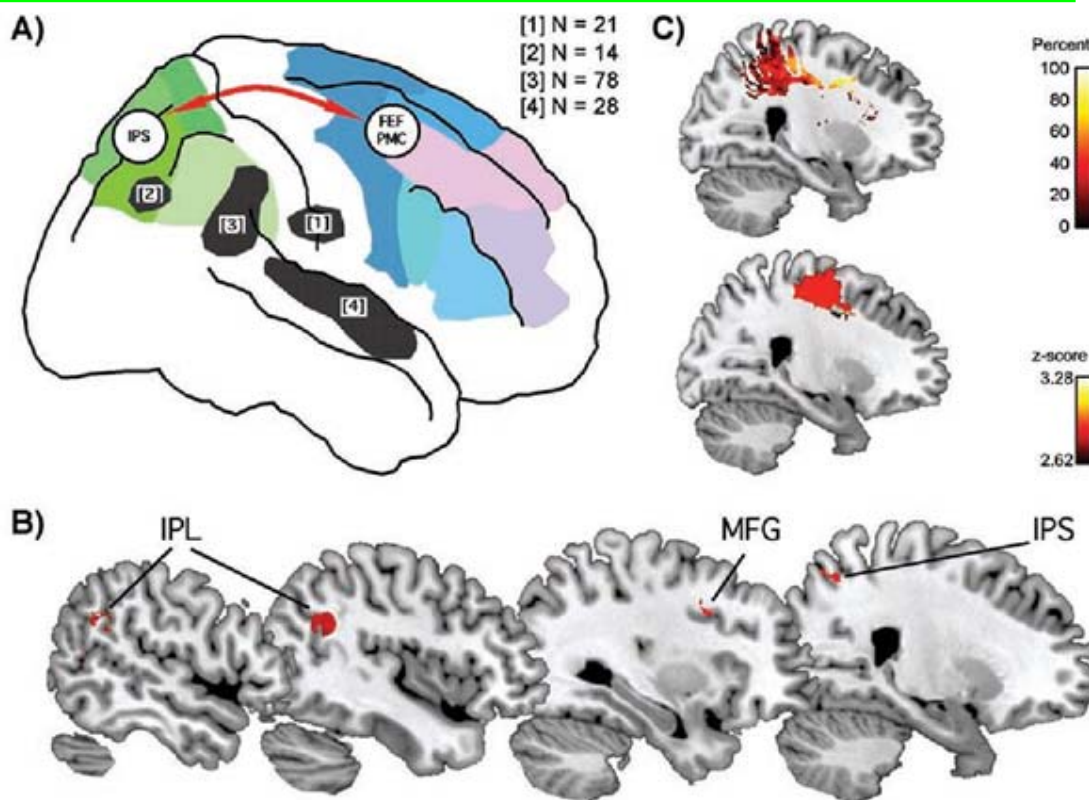
Remember me for faster sign in

This perceptual toggling reveals the fundamentally competitive nature of visual perception. Inside your brain, different neural populations are simultaneously representing competing interpretations of the same visual input. These neural assemblies essentially inhibit one another, creating a neural tug-of-war where only one interpretation can dominate awareness at any moment. What's fascinating is that this competition eventually leads to neural adaptation, the currently dominant neural pattern gradually fatigues, weakening its signal until the alternative interpretation gains enough strength to take over, triggering the perceptual "flip" you experience. This process highlights how vision isn't a

passive camera-like recording of reality; but instead, it's an active, dynamic construction. Your brain isn't simply receiving visual information; it's constantly interpreting, organizing, and even choosing between possible realities when the visual evidence is ambiguous.

The Science Behind the Switches

Now that you've seen several examples yourself, let's discuss what is going on in your brain. The science behind perceptual switching involves an interplay of both "bottom-up" and "top-down" brain processes. On the bottom-up side, when you stare at one interpretation of an ambiguous figure, the neural populations representing that perception gradually fatigue or adapt over time. This neural adaptation creates a kind of self-defeating situation: the longer you see the vase in Rubin's illusion, the weaker the neural signals maintaining that interpretation become. Eventually, these signals weaken enough that the competing neural representation (the faces) can overcome its suppression and suddenly dominate your awareness. This adaptive process explains why these switches often feel involuntary, even if you're trying to maintain one interpretation, your neurons simply can't sustain the same firing pattern indefinitely, creating an inevitable cycle of perceptual toggling.



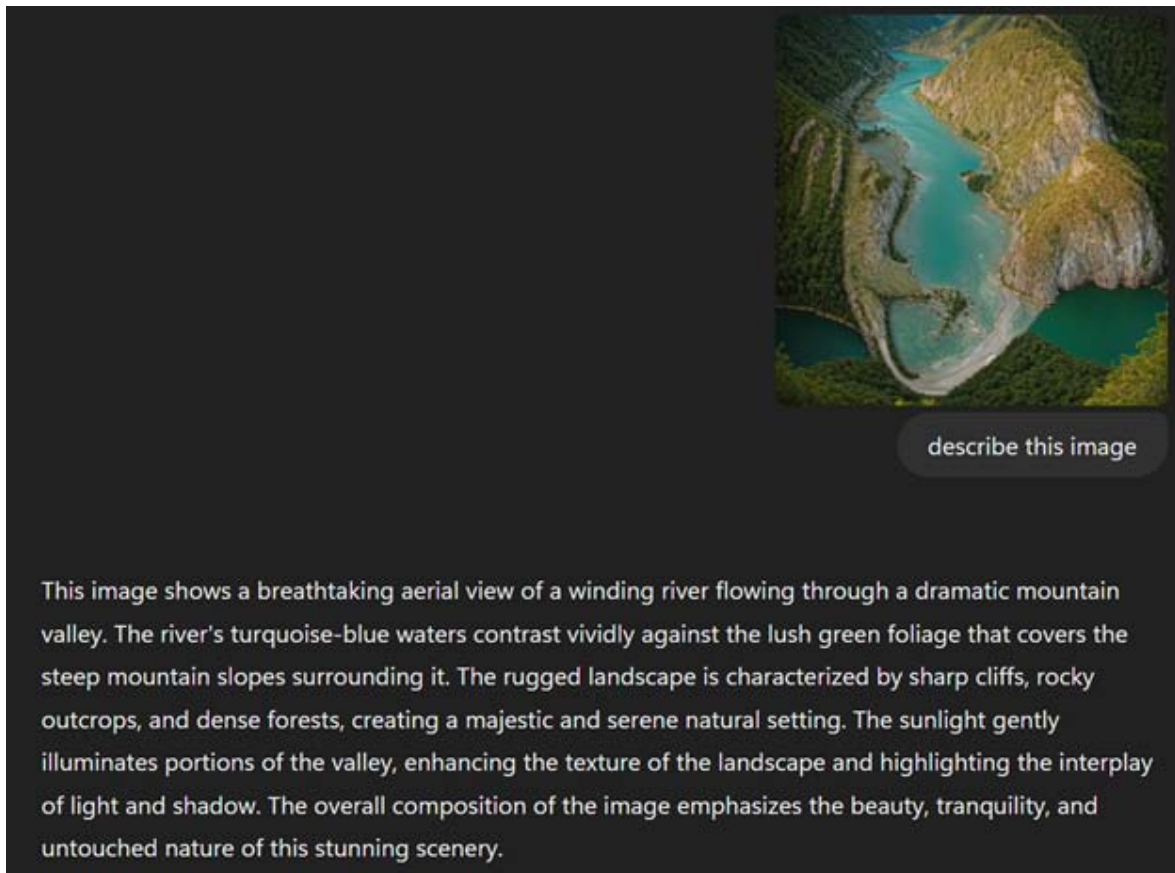
The frontoparietal network.

Modern brain imaging has revealed that perceptual switches involve more than just sensory adaptation. When you experience a flip between seeing a duck and a rabbit, your brain shows increased activity in the frontoparietal network, regions in the frontal and parietal lobes associated with attention and conscious awareness. These areas appear to act as a kind of perceptual "switch-flipping" mechanism, helping to coordinate the transition between competing visual interpretations.

Moreover, top-down factors significantly influence how often these switches occur. Directing your attention to specific features (like focusing on what is the duck's bill) can bias perception toward one interpretation, while your expectations and prior knowledge shape what you see first and how quickly you notice the alternative. This explains why, once someone points out both interpretations of an ambiguous figure, you typically see the switches more readily, your brain now has two valid “templates” to match against the visual input.

What These Illusions Reveal About Perception

Gestalt illusions demonstrate that vision isn't a passive recording of reality but an active construction process. This perspective is the opposite of how we typically think perception works — that we experience the world only as it is. Instead, our brains constantly make educated guesses about ambiguous visual input, filling in gaps and resolving uncertainties based on context and probability. These illusions specifically create scenarios where multiple interpretations are equally probable, forcing our perceptual system to commit to one interpretation before eventually switching to another. The fact that our assumptions shape our perceptions means that modern AI systems that can process images might not see everything. These multimodal systems, such as OpenAI's GPT and Anthropic's Claude, can describe images, read graphs, and more, but fail at recognizing the face in a Gestalt image.



GPT-4.5 didn't recognize the face in the image, which is natural for us.

Prior knowledge and experience clearly shape this perceptual process. If you've never seen a duck

before, you might struggle to perceive the duck interpretation of the duck-rabbit illusion. This explains why some people might get “stuck” seeing only one interpretation until the alternative is pointed out, a moment that often produces a sudden “aha!” experience as the brain reorganizes the visual information. What these illusions reveal about the limits of conscious control over perception is fascinating, despite knowing that both interpretations exist and even actively trying to see both simultaneously, our perceptual system refuses to accommodate this request. In experimental settings, participants instructed to maintain one perception of an ambiguous figure can briefly delay switches but cannot prevent them indefinitely. This limitation is a basic constraint in our visual awareness: certain aspects of perception operate according to their own rules, sometimes beyond the reach of conscious will. What will it mean when AI systems recognize both interpretations or even both at once? They have a different “perceptual system” so illusions should affect them differently.



GR describe this image

This image shows a stunning aerial view of a turquoise river winding through a deep mountain gorge. The river has a distinctive milky blue-green color, suggesting glacier meltwater rich in minerals. It flows between steep rocky cliffs and mountains covered with vegetation.

The landscape features dramatic steep mountain slopes on both sides of the river. The mountains are covered with dense forest in some areas and exposed rock faces in others, creating a striking contrast. The vegetation appears to be a mix of coniferous forests and other alpine plants.

In the foreground, there's a sandbar or small delta where sediment has collected in the river, creating lighter-colored areas where the water is shallow. The river bends through the valley, creating a serpentine path.

This type of landscape is characteristic of glacial valleys and can be found in mountainous regions around the world, though without knowing the specific location, I can't identify exactly where this is. The vibrant color of the water and the dramatic topography make this a particularly spectacular example of a mountain river ecosystem.

Claude 3.7 couldn't see Einstein either.

Beyond Entertainment: The Bigger Implications

Beyond their entertainment value, Gestalt images have influenced fields from art to cognitive science. Artists like M.C. Escher and Salvador Dalí deliberately incorporated perceptual ambiguity into their works, creating pieces that challenge viewers to reconsider their assumptions about visual reality. In

graphic design and advertising, logos like the FedEx arrow (hidden between the 'E' and 'x'), the smile under the Amazon logo, or the NBC peacock leverage figure-ground relationships to create memorable visual identities. Filmmakers use similar perceptual tricks to guide attention, architects employ them to shape spatial experiences, and user interface designers harness Gestalt principles to create intuitive digital environments.



Gestalt principles at work in logo design.

At a deeper level, bistable illusions might even provide a window into the nature of consciousness itself. When your perception flips from vase to faces without any change in the visual stimulus, you're witnessing a moment when consciousness changes while reality remains static. This research connects to broader philosophical questions about how we construct our subjective reality from sensory information. Plus, studying these perceptual mechanisms has practical applications in understanding visual processing differences in conditions like autism, schizophrenia, and ADHD, where perception of ambiguous figures often follows different patterns. For instance, people with autism spectrum disorders typically show slower rates of perceptual switching, while those with schizophrenia may experience more rapid transitions. These differences provide important insights and help researchers develop better interventions for visual processing challenges, demonstrating how something as seemingly simple as a duck-rabbit drawing can highlight the complex workings of the human mind.

Conclusion

The bistable illusion, whether it's the classic duck-rabbit drawing or Einstein's Gorge, deserves recognition as more than just a visual party trick. These simple images offer insights into how we experience reality. Each perceptual flip reveals the active, constructive nature of vision and the complex neural machinery working behind the scenes to translate light into meaningful experience. As you look at these illusions, you're witnessing your own mind in the act of constructing reality, making choices about what to perceive even when you're not consciously aware of these decisions. If our

brains can generate two completely different perceptions from the exact same visual information, **how much of what we consider “objective reality” is actually a construction of our minds?** When will AI models be able to perceive two sides of the same image as easily as we do?

Gestalt

Bistable Illusion

Perception

Artificial Intelligence

Ambiguous Figures

3



Follow

Written by Greg Robison

107 followers

.

4 following

With a Ph.D. in cognitive development and background in neuroscience, I bring a human-centric view to AI, whether theory, tools, or implications.

No responses yet



Write a response

What are your thoughts?

Cancel

Respond

More from Greg Robison



Greg Robison

Severance Explained: OpenAI's Deep Research Summary of Season 1

OpenAI recently released their new research agent called Deep Research, which conducts multi-step research on the internet of whatever...

Feb 9, 2025

A clap icon4



Greg Robison

Neuro-Symbolic AI: A Foundational Analysis of the Third Wave's Hybrid Core

Below is an unedited Gemini Deep Research review of the third wave of AI, neuro-symbolic.

Nov 19, 2025

A clap icon3



Greg Robison

Autoregressive Models for Image Generation: Principles, Architectures, and Analysis

I've been really enjoying GPT-4o's approach to image generation that differs from traditional diffusion models, with some general info on...

Apr 12, 2025

A clap icon15



Greg Robison

A Comparative Analysis of Diffusion and Autoregressive Models for Text Generation: Architectures...

Many of us are familiar with diffusion models that create images such as DALL-E, Midjourney, and Flux. With the imminent release of Google...

Jun 14, 2025



See all from Greg Robison

Recommended from Medium



In
Towards AI

by
Shreyas Naphad

If You Understand These 5 AI Terms, You're Ahead of 90% of People

Master the core ideas behind AI without getting lost

Mar 29

A clap icon11.7K



In
Write A Catalyst

by
Dr. Patricia Schmidt

As a Neuroscientist, I Quit These 5 Morning Habits That Destroy Your Brain

Most people do #1 within 10 minutes of waking (and it sabotages your entire day)

Jan 14

A clap icon47K



In
ILLUMINATION

by
Sufyan Maan, M.Eng

I Woke Up at 4:30 AM Every Day for 30 Days — Here Is What Nobody Tells You

Here is what actually happened, from someone who did it & tracked everything.

Apr 3

A clap icon8.2K



In
Level Up Coding

by
Kusireddy

I Stopped Using ChatGPT for 30 Days. What Happened to My Brain Was Terrifying.

91% of you will abandon 2026 resolutions by January 10th. Here's how to be in the 9% who actually win.

Dec 28, 2025

A clap icon13K



In
Personal Growth

by
Thomas Oppong



Careers Are Collapsing. Jobs Are Dying. The Smartest People Are Doing This Right Now

“Earn with your mind, not your time.” — Naval

Feb 22

A clap icon 11.9K



Suleiman Tawil

Qwen Just Quietly Became the Most Dangerous Open-Source AI Model

The most-downloaded AI model family on Earth was built by a small team with fewer resources than its competitors. Then Alibaba restructured...

Mar 31

A clap icon 1.8K

A response icon 51

See more recommendations