



Psychological Reality: Why We See the World the Way We Do

Video Transcript

The inverse problem of perception

Pattern recognition gives us a powerful starting point. By extracting features from the retinal image and combining them layer by layer, the visual system can detect edges, shapes, and objects with remarkable speed.

But feature extraction alone is not enough. Perhaps the most significant of all challenges for a perceptual system lies in the inherent ambiguity of natural stimuli. A challenge that no amount of feed-forward processing can resolve. This problem is called “the inverse problem of perception”.

To build some intuition, consider a simple example. Suppose I tell you that two numbers add up to 45. Can you tell me which two they were?

There is no single answer. Infinitely many pairs of numbers sum to 45: 1 and 44, 5 and 40, 10 and 35, and so on. Without further information, the problem has no unique solution.

In mathematics and science, we call this an ill-posed or underdetermined problem. The data do not constrain the answer to a single solution. Now, one way to narrow it down is to bring in extra knowledge. We could just make a guess. Maybe you know the person asking the question has a particular affinity for the number 23, being born on April 23rd, and so you guess 23 plus 22 is the solution that we’re looking for. But we don’t know the real solution.

The same problem goes for much of perception. Imagine seeing the projection of a 2D cube onto your retina. That image could have been produced by a perfect cube viewed straight on. But it could equally come from a trapezoid seen at a particular angle – or any number of other three-dimensional shapes in specific positions. In fact, an infinite amount of different 3D scenes could project to the same 2D retinal image.

This is the inverse problem of perception. Because, in the physical world, objects reflect light that reaches our eyes.



But perception works in the opposite direction. The brain must work backward from the retinal image to infer which object or scene produced it. Since many different world states can lead to the same image, this inversion has no unique solution. And that means the brain must go beyond the incoming sensory data. It must infer the most probable cause.

The good news is that perceptual systems are remarkably good at doing exactly that. In everyday life, our perception feels clear, immediate and unambiguous. But under the hood, unconsciously, the brain is constantly solving this ill-posed problem. It does so by drawing on prior knowledge – expectations about how the world typically works. These expectations constrain the space of possible interpretations and guide the brain toward the most likely solution.

Here is a classic example: the Adelson checkerboard illusion. A checkerboard with a cylinder casting a shadow. Look at the squares labeled "A" and "B". "A" clearly looks darker, right? But here's the twist: they're exactly the same shade of gray. You don't believe it?

Well, that's the illusion. Your brain automatically, and without you noticing it, adjusts for the shadow of the cylinder. So, what's happening is that your brain is considering the light source and subtracting that light source from your resulting perception. Your brain is performing an unconscious inference – a best guess, if you will – to produce a stable experience.

So, what actually happens during the process of perception? How does the brain arrive at its best guess of what's out there? Understanding perception requires us to explore how the visual system infers the hidden causes of sensory input. We will next turn to the analysis-by-synthesis framework. In this account, the brain uses internal generative models to predict what it expects to see – and tests those predictions against incoming sensory data.