UC Berkeley
Berkeley Scientific Journal

Title
Filling in the Blanks: The Neurons That Build Visual Illusions

Permalink
https://escholarship.org/uc/item/0sm5x53v

Journal
Berkeley Scientific Journal, 30(1)

ISSN
1097-0967

Authors
Sridhar, Sahana
Yongho, Adrianne

Publication Date
2026-10-30

DOI
10.5070/BS3.69980

Copyright Information
This work is made available under the terms of a Creative Commons Attribution-NonCommercial-NoDerivatives License, available at https://creativecommons.org/licenses/by-nc-nd/4.0/

eScholarship.org Powered by the California Digital Library
University of California

Filling In the Blanks: The Neurons That Build Visual Illusions

INTERVIEW WITH: PROFESSOR HILLEL ADESNIK

BY: SAHANA SRIDHAR AND ADRIANNE YONGHO

***Professor Hillel Adesnik** is a neuroscientist at UC Berkeley, where he is an Associate Professor of Neurobiology and a member of the Helen Wills Neuroscience Institute. His research centers on how neural circuits and synaptic mechanisms give rise to sensory perception, using techniques like optical imaging, electrophysiology, and optogenetics. He got his PhD at UC San Francisco and completed postdoctoral work at UC San Diego.*

BSJ: How did you become interested in circuit and system neuroscience? Was there any particular problem that motivated you to go down this path?

HA: When I did my PhD about 15-20 years ago in San Francisco, I was very excited about learning and memory. I knew it was how we encode information — the phone number, we remember where we grew up, what we ate for breakfast, and so forth. And this was 20 years ago, so there wasn't a lot of circuit neuroscience. The tools were not around, so we were studying synapses and how synapses change strength. Over time, I realized that as exciting as that was, it was a molecular-level question. We weren't actually getting at the question that I thought was a little more relevant: how is the circuit changing to represent the stored information? The tools were just coming online at the time. This was in 2007, so we had these tools called optogenetics that had just been engineered, and suddenly it was possible to start looking at those questions. When I started my postdoctoral fellowship at UCSD, I got caught up in understanding circuits — how the brain computes, how we see, how we sense touch — without focusing as much on plasticity. I did that for about 10 years. More recently, in my lab, we started studying learning again. Now that we have the tools to combine circuit understanding with how synapses change their strength to encode information, we're coming full circle.

BSJ: Transitioning to your most recent paper that identifies a specialized subset of neurons in the primary visual cortex that encode illusory contours (ICs), can you explain what functional features distinguish these IC encoders from other V1 neurons?

HA: Functionally, how we define an illusory contour coding neuron is a neuron that will respond as if there is a stimulus, but there is not. In the visual system, you represent the visual world. There is a map of the world, of what your eyes are seeing on your brain. Every neuron in your visual cortex responds to a certain region in space of where your eyes are looking. In the space of the retina, most neurons respond when something is in that little space. It is known as the receptive field — a little portion of the visual field that it responds

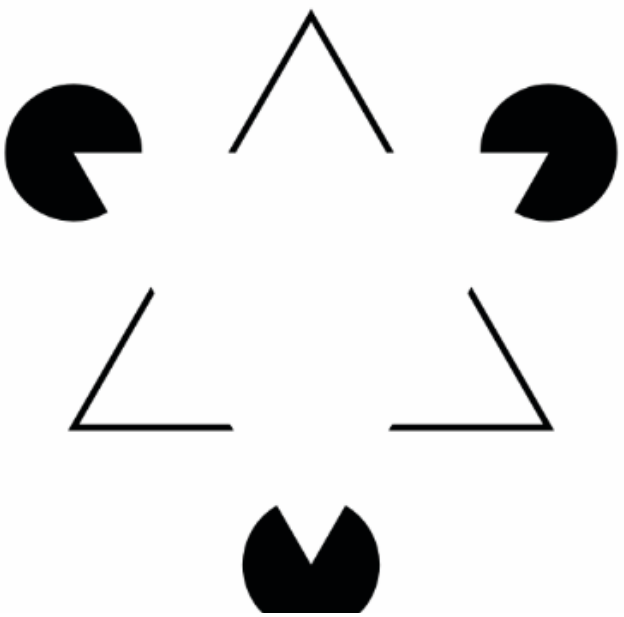

Figure 1: The Kanizsa Triangle. An example of an optical illusion in which strategically arranged "Pac-Man" inducers create the perception of a bright, equilateral triangle that is not physically drawn. Observers experience clear contour boundaries and surface shape despite the absence of actual defined edges.

to. Almost all neurons respond when something is in their receptive field, like a fly moving across your vision or a baseball coming at your face; most neurons will not respond when there is nothing there. That is by definition, but there is a small group of neurons that will respond even if there is nothing, as long as what is around you gives you the illusion of an object that goes through the receptive field. So if you think about the Kanizsa Triangle, this illusory triangle created by these three Pac-Man shapes, we see a border where there is literally nothing. There is a neuron that is taking in that gap, and there is nothing in that neuron's receptive field, and yet the brain is telling you there is a triangle. In other words, these neurons are encoding this illusory edge. These illusory edges together come to form the triangle. The IC encoding neuron is that neuron that responds as if there is a real edge, when there is only an implied or inferred edge.

BSJ: Your findings show that layers two and three in V1, but not layer four, carries the perceptual representation of illusory contours. What does this difference suggest about the relative contributions of feedforward versus recurrent processing pathways in the early visual cortex?

HA: Layer four is known as the input layer. The numbers just indicate the depth, but it turns out the information coming from the retina first hits layer four, which is in the middle, and then information gets broadcast into layer two or three. Layer four does not have these IC encoding neurons or representation so it makes sense; this is the first layer, and it is dominated by what is coming from the eyes. Layers two and three are these very interesting layers, which we study a lot. It integrates bottom-up, coming from the eyes via layer four, but also top-down from higher cortical areas that are deeper into your brain. So layer 2/3 sits at this nexus of bottom-up from layer four and top-down from higher visual areas. We could have seen something in layer four, but layer four doesn't get much top-down input. This is a clear one where a computation that depends on inference, the idea there is a triangle, even though there is not actually a triangle, is not occurring based on (top-down) or purely-bottom up. It requires context, and that context is being fed back down from higher visual areas.

BSJ: How do higher visual areas generate and deliver that contextual information back to earlier stages like V1?

HA: So the question is: how do higher visual areas have that context? The simple way of thinking about it, from a mechanistic circuit point of view, is that as you go deeper and deeper into the brain, the receptive fields that we previously talked about get bigger and bigger until the receptive fields can encompass the whole world. For example, there are neurons in the highest level of your visual system that are apt in humans, that are absolutely selective for certain things, and they will respond wherever that certain thing is. When you are in the primary visual cortex, it is very small. The higher visual areas get bigger, so they can integrate. We think about the illusion of the triangle, you need to integrate the little vertices of those Pac-Men to make sense of it as a triangle. Neurons in the higher visual areas can actually see all of that. Neurons deeper in the brain, in some way, have access to much larger regions of space. So they look at the scene, in some sense, they analyze and say, "Maybe there is a triangle here." They fire a little bit. Our theory is they then send that information back to the lower visual areas. Remember, lower visual areas are more sensitive to what's in little pieces of space. So, in that first pass of information, these IC neurons did not really fire, but they got some information in this kind of feedback loop. As that accumulates over time, they actually ramp up the activity. That amplification or ramping up eventually says there's a triangle. Usually it does take some time for these illusions to appear longer than just a normal feed-forward stimulus.

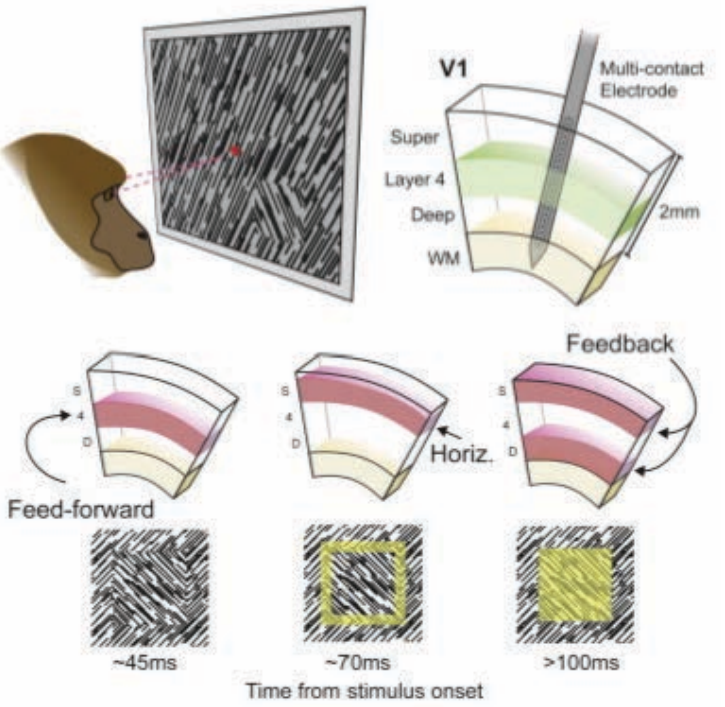

Figure 2: The layers of V1 optical feedback. As a primate views a textured wall, an electrode can be used to record activity across the layers of V1. Early feed-forward responses in Layer 4 capture initial visual signals, followed by interactions that start to strengthen the boundaries and outlines we visualize. Later feedback from higher visual areas amplifies these visual boundaries, supporting the emergence of illusory or context-driven contours.

BSJ: Using two-photon holographic optogenetics, you were able to selectively activate IC-encoding neurons to artificially recreate IC representations in V1. How does this pattern completion mechanism enhance the brain's ability to interpret incomplete sensory inputs?

HA: The theory is that you start with missing information that's coming from the retina. It is just bits and pieces; you can not really make sense of it. That information propagates through the visual cortex. These looser contour neurons are not active. It goes to higher visual areas and bounces back. What we showed is why it is bouncing back. The idea is that looser contour neurons are sensitive to this feedback. They ramp up activity, and they actually locally activate other neurons. As you ramp up that activity, that information bounces up again. So you have, you know, retina to cortex, to higher visual areas, back to low cortex amplification, back to higher visual cortex. What we think is going on based on the experiments where we stimulated these IC encoder neurons is they were able to selectively amplify, when that information made sense and looked like a triangle. For argument's sake, if it did not look like a triangle, imagine, there are those three PAC

men. If you rotate one outwards, you lose the triangle. And so we think only when those inducing elements that give you this illusion are in the right configuration, do the IC neurons fire and ramp up activity to give you the illusory person.

BSJ: You differentiate IC-encoders from segment responders, the latter encoding bottom-up sensory signals conveyed to the higher cortex. How do these two neuron groups functionally interact to allow perceptual inference across the visual processing hierarchy?

HA: It encompasses the whole circuit. More work needs to be done, but based on our experiments, what we think is happening again is when we stimulate the IC encoding neurons, they are sending axons locally, right or laterally within a few 100 microns in the primary visual cortex, and they are ramping up the activity of the segment responders which then ramps up activity further in the higher visual areas, and then it goes back. You can think of it as like a loop: it comes from the retina, and it goes forward, and then it does this a few times, and then goes into perception. Almost like when they have satellites that go around the earth, and they shoot them off with gravity, the idea is you need to ramp up the activity. It is the illusory contour coding neurons there. One hypothesis is that they make extra connections, maybe the looser content neurons are like hub cells that send out locks of axons. What we do know is they are sufficient to drive this representation, and that has to be through some divergent axonal output. But admittedly, in our study, it was purely functional, so we have no anatomical explosion. That is a hypothesis you could imagine testing it by tracing these cells, which we have not done, by doing electron microscopic reconstructions, which is not trivial, so we have not done those things, but in the future, it would be nice to try.

BSJ:The recurrent loop circuit model you propose implies active reinforcement of top-down signals within V1. To what extent might this circuit architecture generalize to sensory inference mechanisms in other cortical areas or sensory modalities?

HA: I could only speculate. The general process we are talking about is known as perceptual fill-in where there is missing information. In this circuit, the brain uses sort of prior information or experience to fill in like the best example is, actually, is the blind spot. Have you heard of the blind spot, which we do not experience right? Unless you cover one eye, and so we have no awareness of the blind spot. We're humans; we are very visual. Are there corresponding illusions? I think there are tactile illusions that you can get. There must be auditory illusions. To be honest, I am not sure. It is not nearly as well studied as vision, which is just so highly studied. I think it is very possible that the same mechanisms we are talking about are in play in audition and somatosensation. In fact, olfaction is an interesting one, even though humans tend to be less olfactory even than rodents. Oftentimes when you think about odor mixtures, most things that we eat are mixtures of many odors. Oftentimes you take out one or two of the odors and it is kind of the same, because your brain can fill it in. And so there is probably a lot of pattern completion going on in olfaction. And this has been studied in rats, for example, which are super olfactory animals. On the one hand, they will easily generalize even when you are missing a few odors. You can also train a rat to distinguish two mixtures of odors where only one is different. So it goes both ways. On the one hand, you can generalize if you need to. On the other hand, you can be very specific, if necessary. So, it depends on the demands that you are putting on the animal or that the animal has in its own natural environment. So, I think they are probably general mechanisms. They are not necessarily identical. The olfactory system, in fact, is a very different circuit system than the rest of the brain, because it evolved much earlier than vision and audition and somatic sensation. The basic idea is, the computational processes are probably similar.

BSJ: How do your findings provide experimental evidence supporting theoretical predictions from artificial neural networks regarding the necessity of recurrent loop connectivity for perceptual inference? How might these insights shape future directions in computational neuroscience and AI?

HA: I am not an expert in the field of AI, but there is a large body of work. Most artificial networks that are still used are mostly feed-forward, like computational neural networks that do really good image identification. But there's definitely a large field that is asking, what happens if we have feedback connections, local connections, and so there is some evidence in the AI literature that reconstructing these networks with more biologically realistic feedback connectivity can improve things. There are recurrent neural networks, which are not exactly the same as recurrent neural networks in the brain, but they have features that allow for missing information to be inferred through pattern completion. This has definitely been shown. Historically, my understanding is that illusions that trick humans did not normally trick ANN (artificial neural networks) in the same way, and that gave us insight. There is at least one study that argues that if you structure it more like the brain, it can be more fooled. But it is unclear how similar these systems are ever going to be to the human brain. They have been inspired by the human brain. I think the real similarity between the artificial networks and brain networks, other than some very general similarities that were by design, is that they had the same training objective, in the sense that an artificial network is trying to recognize faces, and the primate brain has spent millions of years trying to recognize faces. And I think that at a very high level, you can think that they are both being trained on the same objective, or objective function, which is to identify faces wherever they are, despite looking very different under lighting, under distance and so forth. Because they have the same objective, they might end up using very similar mechanisms, and one of them might be feedback connections to deal with uncertainty. But of course, it is such a different thing, right? An AI network from the brain that you know, it may not be instantiated in any similar way, but because the objective is the same, it might have something more similar.

BSJ: You were talking about how ICs fill in our information, like if you see something, your brain is tempted to make a pattern or connect the edges. Is this also a similar mechanism that is in play when we visualize something in our head, like a memory?

HA: Intriguing question. So here is the problem, right? Mice, which we study, and we can use these techniques on, might be hard to get them to do visual imagery. You have to read out visual imagery. And in any animal, reading out visual imagery is hard because they have to tell you what they are imagining. There are maybe some

clever behavioral ways to get at this. There's quite a large literature about humans, right? People put humans in fMRI scanners, where they do brain imaging, and ask people to visualize things, and there are dedicated regions of the human brains that will light up when you are doing visual imagery. Usually that does not propagate down to the lowest level of the visual cortex. That is my understanding. I think it stays in the higher regions. Now, mice are very different from humans. They have a much larger, relatively speaking, lower visual cortex. Of course, it is tiny compared to ours, but you can think that over evolution, what happened in humans is that the higher visual areas became increasingly complex, large and elaborate, and the lowest visual area V1, even though it is still big, became a much smaller component of the whole process. Whereas in mice, it is mostly the primary visual cortex and a few little visual areas. And so far, there is no evidence they can see faces or anything like that. There is a recent paper that says they can detect objects. I think it would be fascinating. I am not sure that mice or rodents are going to be the appropriate model to answer your questions, but there are other primates that we can do experiments on, that you might be able to infer what their visual imagery is, at least, by doing neural recordings.

BSJ: What inspired you to focus on the network influence of cortical ensembles rather than just the properties of each individual neuron during stimulus detection?

HA: In fact, that was not our initial focus. The initial hypothesis was that if we are trying to understand how a mouse senses a stimulus on a computer screen — they are not the most visual animals, but they are still very good, actually, almost as good as humans — it is not just about detecting a stimulus. In nature, they need to avoid a cat or something like that. So, we thought that if we went in and artificially stimulated the neurons that best responded to or represented the stimulus, those would have the biggest effect on behavior. The student at the time, Haley Bounds, who was working on this project, found that this was not the case. After doing many, many experiments, Haley realized that this clearly was not the best explanation for the results. The data looked noisy at first, but when Haley did an analysis correlating not just how responsive the neurons were, but the actual impact those neurons had on the network, because the neurons we were stimulating project to other parts of the brain but also project locally in the cortex, she found that the best explanation for how the animal actually improved its behavior was whether those cells amplified the activity locally. That is what we mean by "network influence". So, it was not our initial hypothesis at all. It was counter to the dogma in the field, which is why it is such an impactful study. The idea is that rather than what the neurons are encoding being the most important factor, it is their overall effect on the network–at least in the context of stimulus detection.

BSJ: Can you explain the significance of discovering that downstream visual areas sum overall cortical activity, without necessarily weighing the most visually selective neurons more heavily?

HA: In Haley's study, we did not record from downstream areas directly. That part is inferred. We were using the mouse's behavior as a proxy for those downstream responses. We were studying the primary visual cortex (V1), which does not control behavior itself — it represents information. It connects to other parts of the brain that actually drive behavior. In this experiment, the mouse's behavior was to lick for a little juice reward when it detected something on the screen. By inference, if the mouse's behavior improves, it must be because downstream areas are responding better. The logic is that the animals' improved performance could be explained not by what those neurons encoded, but by their net impact on the network. So even though we did not record from the downstream areas, we can assume that they responded more strongly when the population of neurons in V1 amplified the activity locally, rather than just encoding specific information.

BSJ: How do you think your lab's discoveries reshape our understanding of sensory perception?

HA: I would say that going into this, you would expect that if an animal is trying to perform optimally, to sense the weakest possible stimulus and still get its reward, it would rely on the most informative neurons. Those would be the neurons that have the biggest response to the stimulus. But that is not what we saw. It is important to realize that the experiment we were doing focused on one of the simplest aspects of perception: detection. You can divide perception into detection and discrimination, or identification. Detection is just asking, "Is there something there?", like when a mouse is trying to tell if there is a cat nearby. Once it detects something, it can then turn its head or move its eyes a little and ask the next question: "What is it?", that is discrimination. So, for detection, what we found is that it does not make sense for the brain to pay attention only to the neurons that respond best to a specific stimulus in a specific location. Because at that stage, it is not trying to identify whether it is a cat or a bird, it is just trying to see if anything is out there. If the brain only relied on neurons that responded to one type of cat in one position, it would not detect cats in other positions or other kinds of animals. We think this means that detection depends on generalization. The brain is essentially asking, "Is there something out there? Yes or No?" And it does this by amplifying overall activity rather than focusing on the most selective neurons. If we had trained the mouse to do identification instead–asking whether it is stimulus A or B, like a cat or a bird — we would expect different results. In fact, that is something we are doing now. So, what Haley found in this study is that detection works in a way that is not what we would expect, it prioritizes overall sensitivity and generalizability over specific stimulus information.

BSJ: As a follow-up, you mentioned that this paper focuses more on detection. Do you think it at least lays the groundwork for understanding discrimination/identification, or are those processes more separate?

HA: I would say that this paper really focused on detection, just asking whether there is something there or not. What we found is that the representation that matters for that is very simple: all the activity matters, and if it passes a threshold, the brain says, "There was something." It is trying to generalize, not to respond specifically to stimulus A or B. I do think it lays the groundwork for understanding discrimination. This was about detection; future studies, which we and others are doing now, are about discrimination. That is actually a very well-studied field, but not with the tools we are using. When we move into discrimination, if we activate groups of neurons that represent

different categories–say "cat neurons" versus "bird neurons" — we find different results. So yes, this paper is really detection, but it sets the stage for the next step: understanding how the brain distinguishes one thing from another.

BSJ: The study used two-photon imaging coupled with holographic optogenetic photostimulation. Could you describe how these techniques helped you establish a causal link between neural activity and perception?

HA: Two-photon imaging is a technique that uses lasers to generate dynamic pictures of brain activity. It is powerful because you can image hundreds or even thousands of neurons at a time, and when a neuron fires, it flashes. You can literally see these movies of brain activity, showing which neurons respond to specific stimuli,like different contrasts or shapes or different images like cats versus dogs. The unique part of our study was combining that imaging with holographic optogenetics, which lets us recreate those activity patterns using light. Optogenetics allows us to turn neurons on or off with light, and in this study, we turned them on. Haley's technology allowed her to identify neurons that responded to specific stimuli, or did not, and then reactivate about 20 or 30 of them out of a population of over a thousand. That is incredibly powerful because it lets you test causality. If you just shine light across the brain with a big laser, all the neurons would activate randomly, and it would be hard to interpret the results. But with holographic optogenetics, you can use holography to shape the light in three dimensions and target specific neurons — those defined by their functional responses. That allowed us to directly test how stimulating particular neural ensembles affects perception.

BSJ: Your results suggest that activating highly selective neural ensembles does not necessarily have a larger behavioral impact than activating random ones. Why do you think that is?

HA: It comes back to that trade-off between precision and generalization. In our task, the animal's goal was to detect the presence or absence of a stimulus. If the brain only focused on the most selective neurons, it might do great with that exact stimulus,but what about all the others in the world? Like, all the other cats or birds? So, it is a trade-off. We think that for detection, it is not about reading out the most informative neurons; it is about amplifying the overall signal. It is like turning up the dial on your speaker, what matters is the total output. Haley found that some groups of neurons just amplify activity better than others. We can hypothesize why that might be the case. In another paper, we actually found supporting evidence that certain local systems in the brain are designed to amplify weak or noisy input, helping the brain detect faint signals. So, in this study, we think what matters is the strength of amplification, not how precisely the neurons encode specific details.

BSJ: The study emphasizes the role of network interactions in perception. What was the key challenge you faced in isolating and interpreting these network effects?

HA: There were definitely several challenges, but one of the biggest is that the neocortex that we are studying contains thousands of neurons, all connected through relatively weak synapses. For example, the little window we are imaging through might show around 1,000 neurons, but those are just a tiny fraction of the 100,000 or so in that region. Because the connections are weak, it is hard to see the effect of a single neuron. You can see effects when you stimulate small groups, but there is a trade-off. If you shine a laser on the whole brain, you get a big effect, but it is hard to interpret. If you only stimulate 20 or 30 neurons, the effects are small and subtle. It took a long time for Haley to realize the net outcome of the experiment, because it required collecting and analyzing data across many sessions. The effects were not huge, but after around 40 different sessions, over several years, she could clearly see the correlation. So, the challenge was that the effects of small perturbations are subtle, but with enough data, you can see they are significant and consistent.

BSJ: Looking ahead, what future experiments or technological advances do you think will be crucial for further unraveling how cortical ensembles influence behavior and perception?

HA: I think one big direction is moving from detection, which we have talked a lot about here, to discrimination and identification–really understanding how the brain identifies, say, birds versus cats versus dogs. The technologies we are using now can definitely be leveraged in slightly different ways to answer those questions, and I think we will get different results when we do. Another important step is increasing the scale of our perturbations. The human brain has billions of neurons, and even the mouse brain has around 100 million. Right now, we are only looking at a tiny subset. There is always a trade-off between scale and specificity–if you stimulate more neurons, you risk losing precision. But we have another paper currently under review where we have increased the scale while maintaining most of the specificity. It is not trivial, but it is doable. As technology continues to advance, allowing us to stimulate larger groups of neurons while keeping that precision, I think we will start to unravel the brain's coding strategies in much greater detail.

REFERENCES

1. Shin, H., Ogando, M. B., Abdeladim, L., Jagadisan, U. K., Durand, S., Hardcastle, B., Belski, H., Cabasco, H., Loefler, H., Bawany, A., Wilkes, J., Nguyen, K., Suarez, L., Johnson, T., Han, W., Ouellette, B., Grasso, C., Swapp, J., Ha, V., … Adesnik, H. (2025). Recurrent pattern completion drives the neocortical representation of sensory inference. Nature Neuroscience, 28(11), 2319–2329. https://doi.org/10.1038/s41593-025-02055-5
2. [2] Bounds, H. A., & Adesnik, H. (2025). Network influence determines the impact of cortical ensembles on stimulus detection. Neuron, 113(14), 2358-2369.e5. https://doi.org/10.1016/j.neuron.2025.04.023

IMAGE REFERENCES

1. [Figure 1] Fibonacci. (2007). English: Kanizsa triangle. [Graphic]. Own work. https://commons.wikimedia.org/w/index.php?curid=1788215
2. [Figure 2] Self, M. W., van Kerkoerle, T., Supèr, H., & Roelfsema, P. R. (2013). Distinct Roles of the Cortical Layers of Area V1 in Figure-Ground Segregation. Current Biology, 23(21), 2121–2129. https://doi.org/10.1016/j.cub.2013.09.013