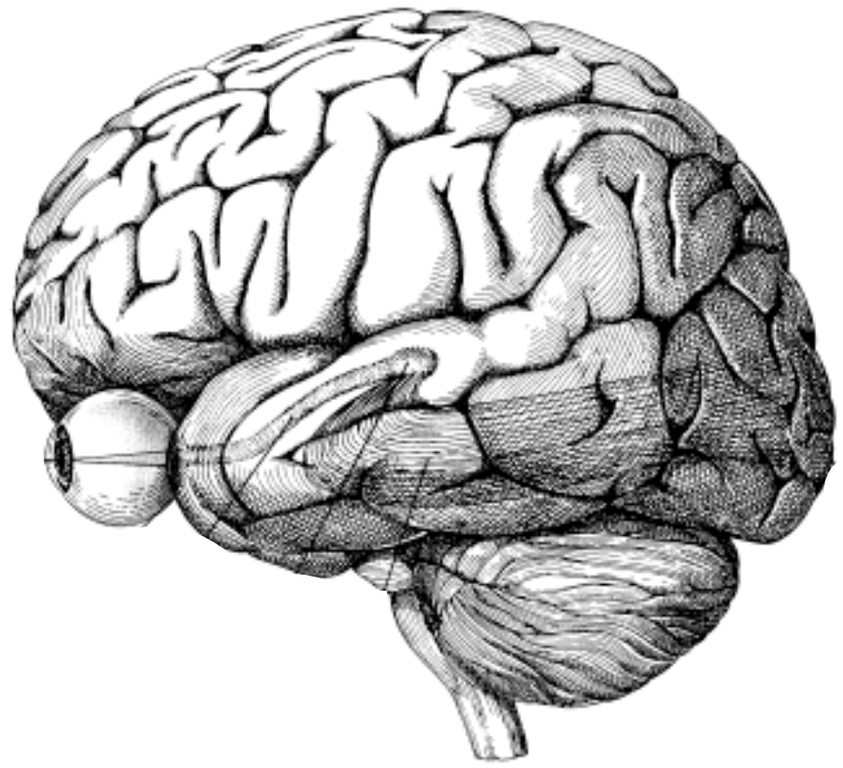
Probing sensory representations with metameric stimuli

Eero Simoncelli

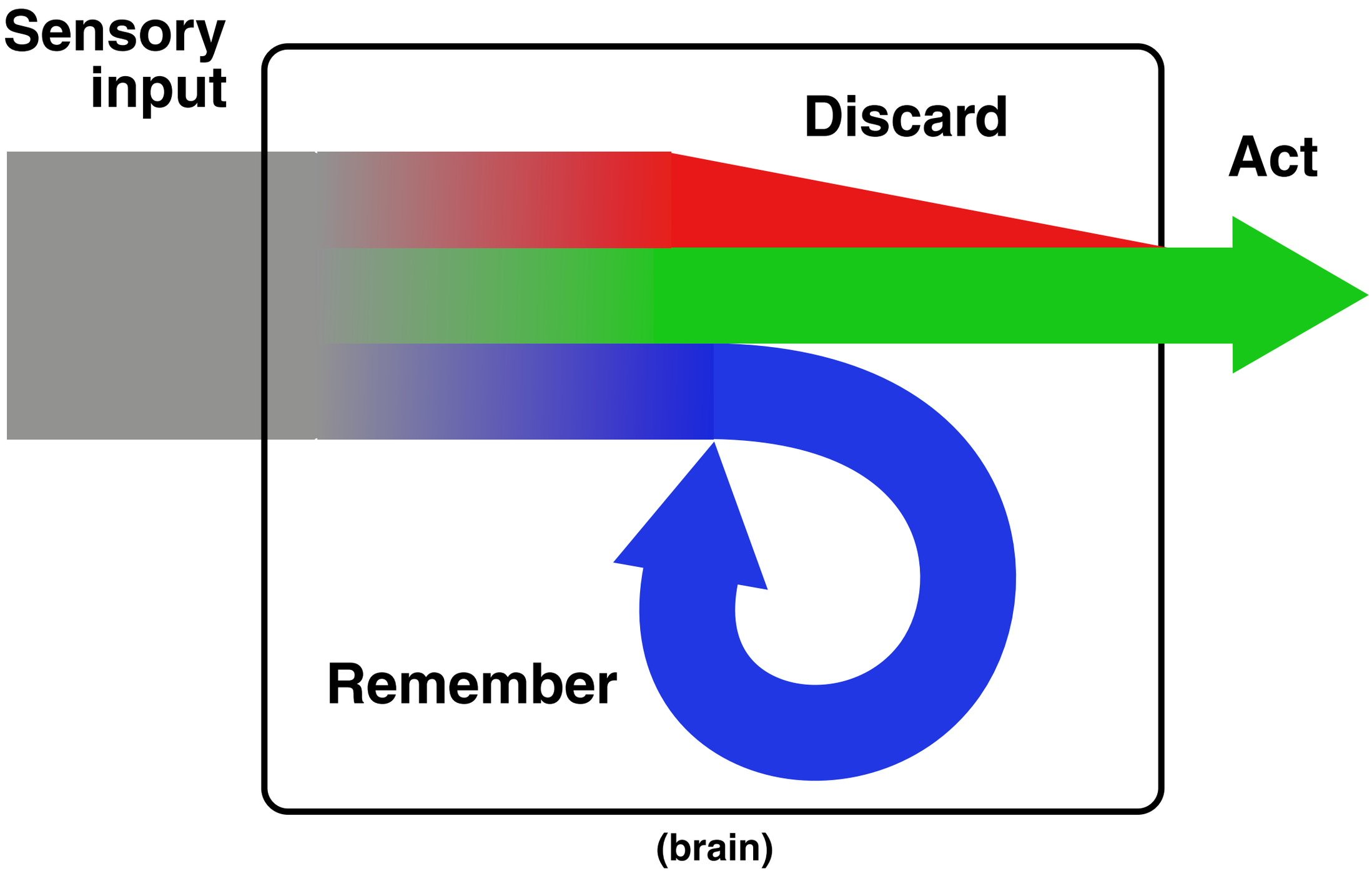
HHMI / New York University



Where does all that visual information go?

[figure: Hubel ‘95]

Destiny of sensory information



Metamers

- Two stimuli that are physically different, but appear the same to a human observer

- Classic example: trichromatic color perception

- Another example: texture perception

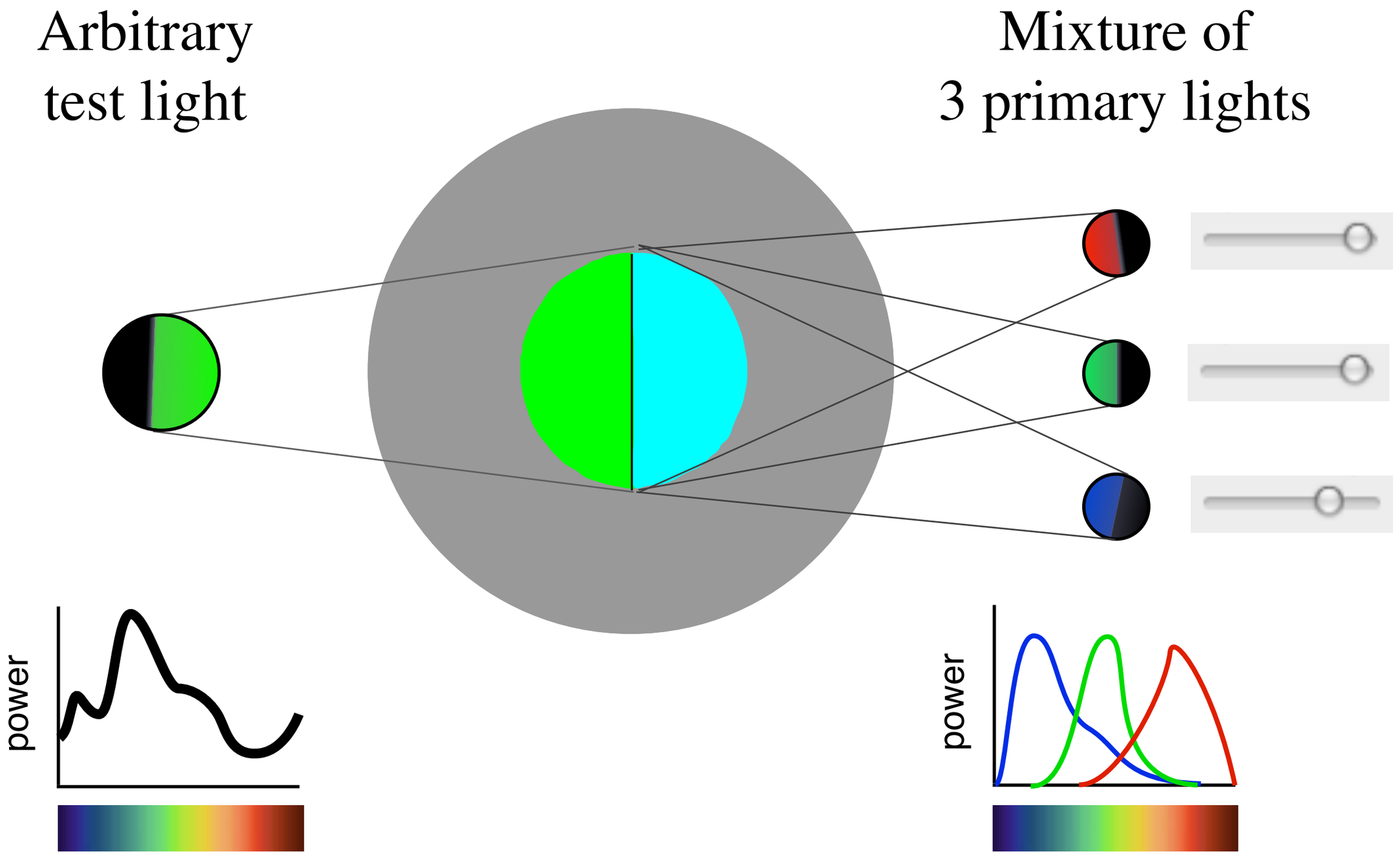
Spectral nature of light

This image is in the public domain.

Diagram of a prism removed due to copyright restrictions. Please see the video.

[Newton, 1665]

Perceptual color matching experiment



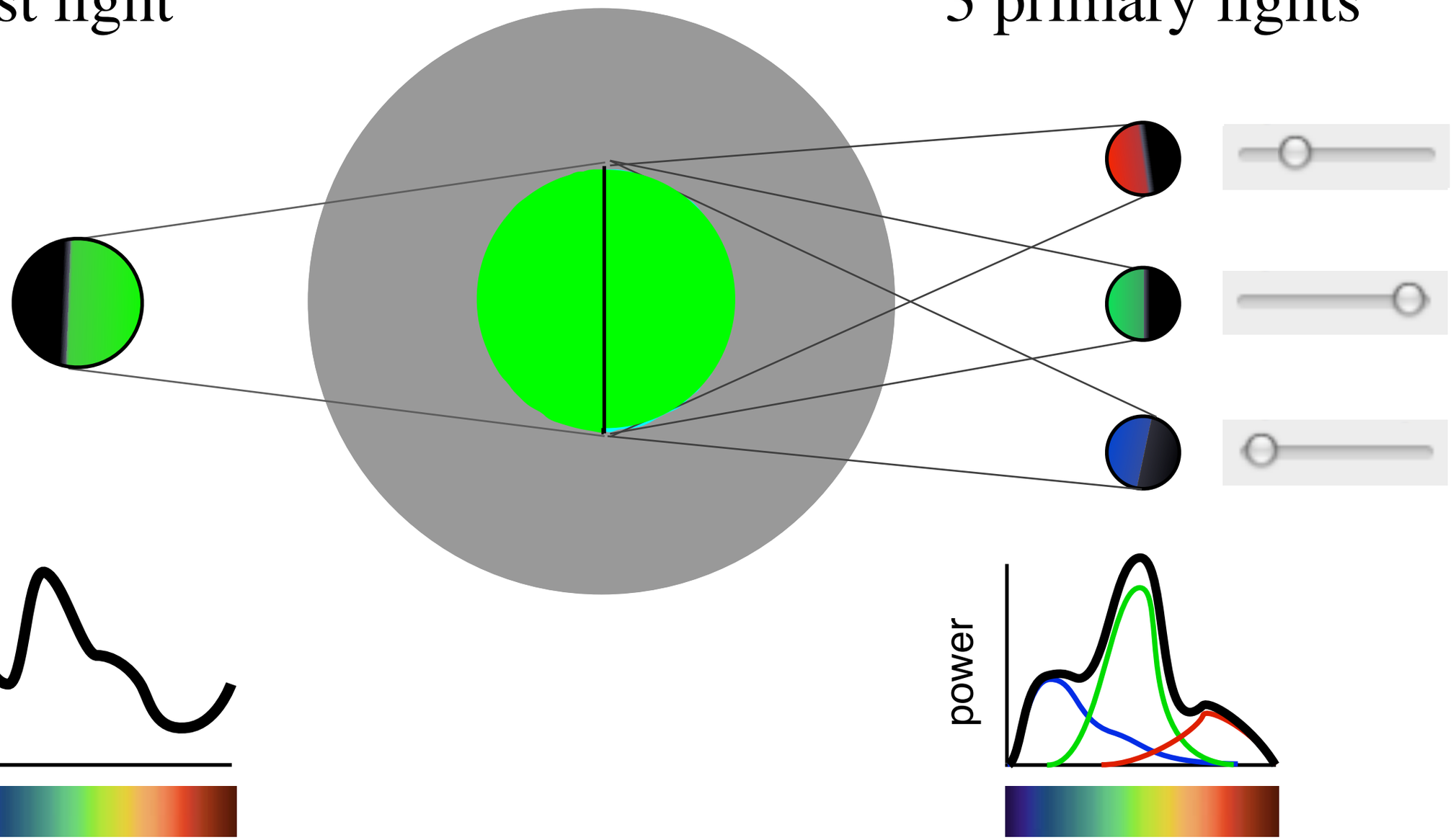
Courtesy of David Brainard. Used with permission.

[Young, Helmholtz, Grassman, etc, 1800‘s; slide c/o D. Brainard]

Perceptual color matching experiment

Arbitrary
test light

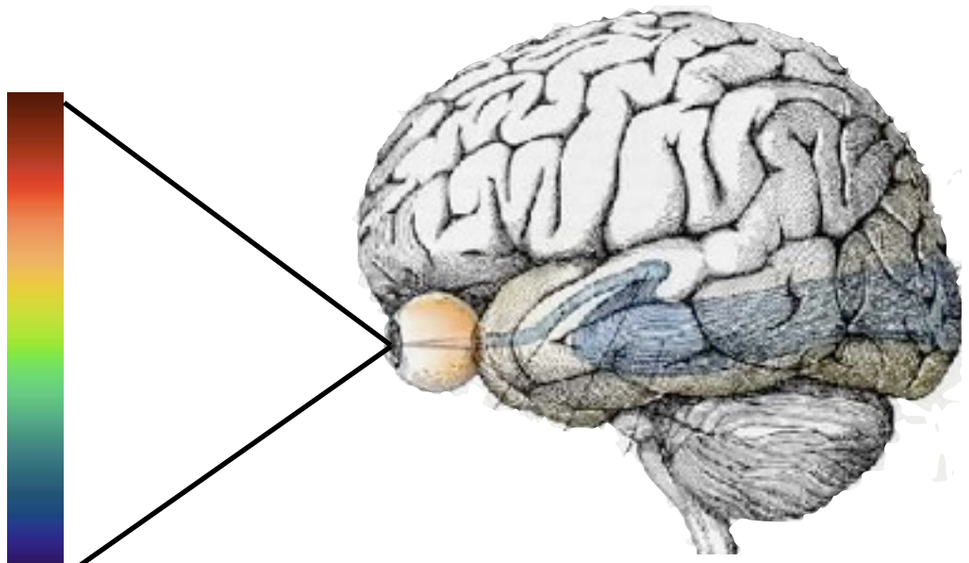
Mixture of
3 primary lights

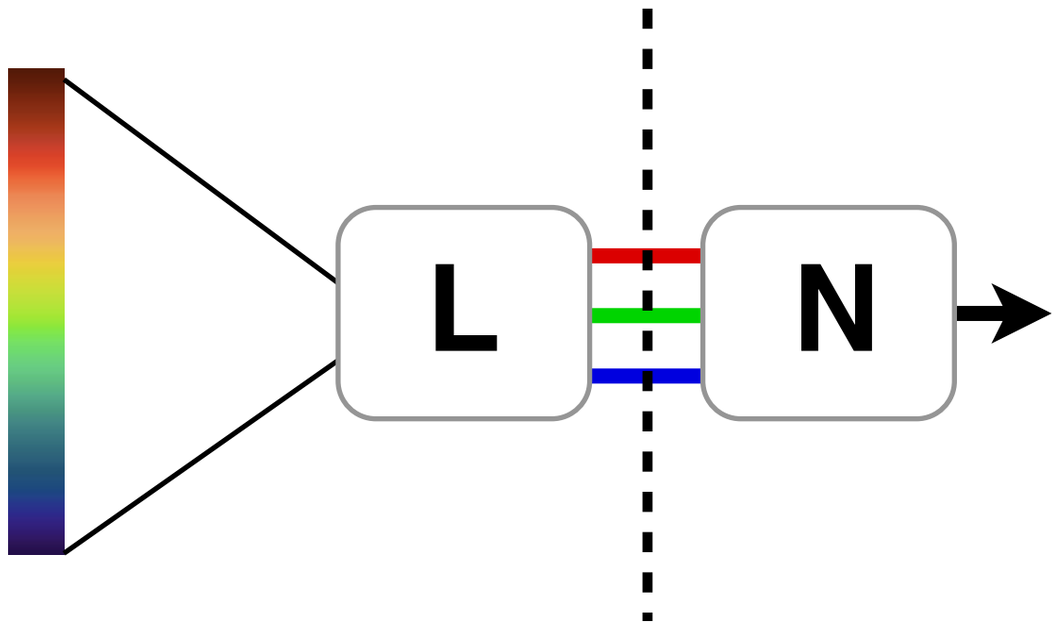


Courtesy of David Brainard. Used with permission.

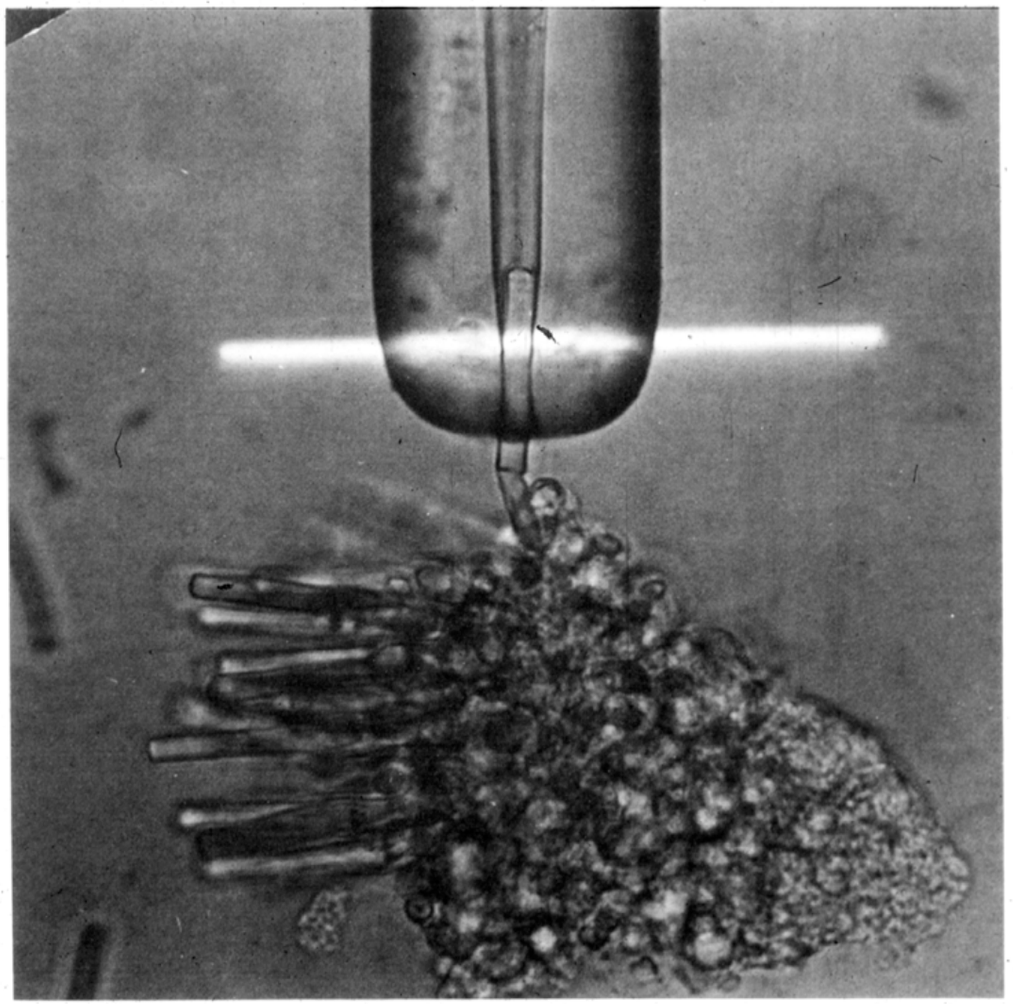
[Young, Helmholtz, Grassman, etc, 1800‘s; slide c/o D. Brainard]

Theory (Grassman, 1853): the visual system performs a **linear projection** of the wavelength spectrum onto a three-dimensional response space





- Predicts/explains perceptual “metamers” - lights that appear identical, but have physically distinct wavelength spectra (1800’s)
- Codified in CIE standards for color representation (1931)
- Underlying mechanism (cone photoreceptors) verified (1987)

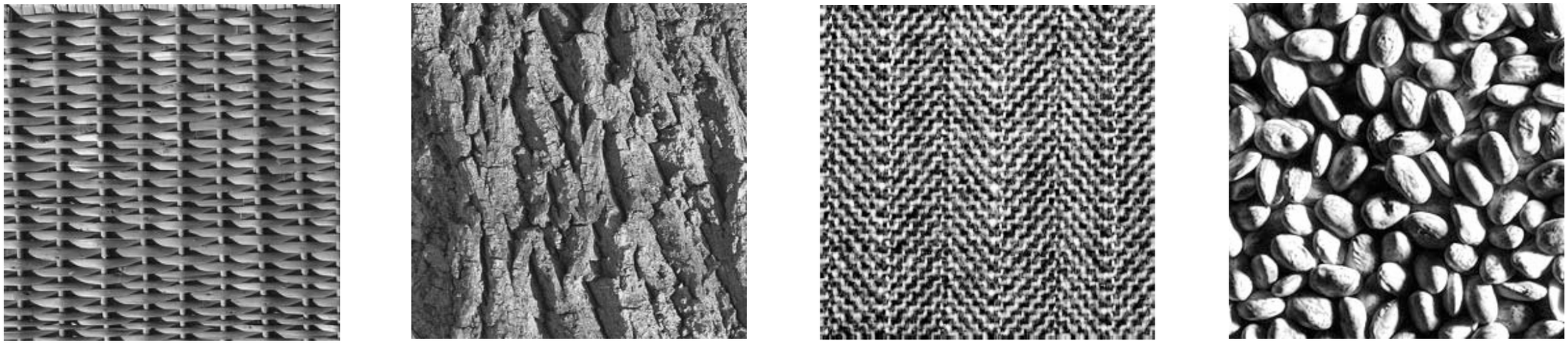


Courtesy of Denis Baylor. Used with permission.

Figure removed due to copyright restrictions. Please see the video.
Source: Figure 3 from Baylor, D. A., B. J. Nunn, and J. L. Schnapf.
"Spectral sensitivity of cones of the monkey Macaca fascicularis."
The Journal of Physiology 390, no. 1 (1987): 145-160.

[Baylor, Nunn & Schnapf, 1987]

Visual texture



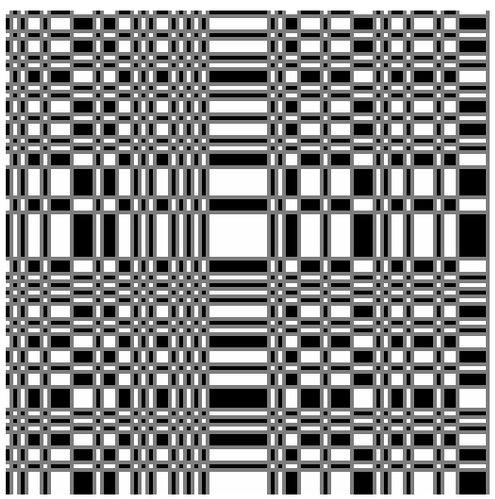
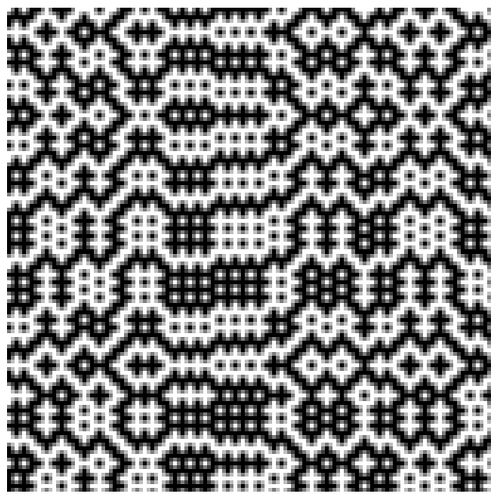
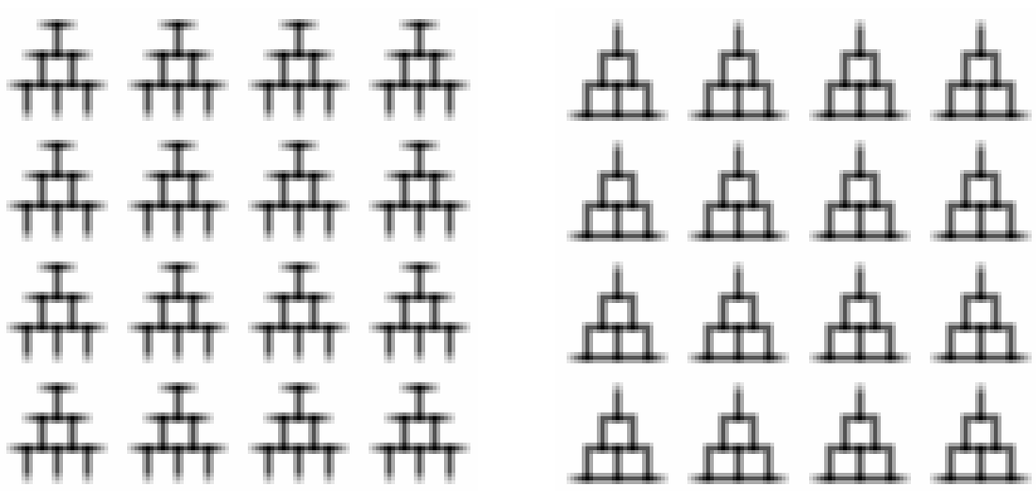
Source: Portilla, Javier, and Eero P. Simoncelli. "A parametric texture model based on joint statistics of complex wavelet coefficients." International journal of computer vision 40, no. 1 (2000): 49-70.

Homogeneous, with repeated structures

Julesz (1962)

- Hypothesis: Two textures with identical Nth-order pixel statistics will appear the same (for some N).

- Hand-constructed counter-examples (N=3):

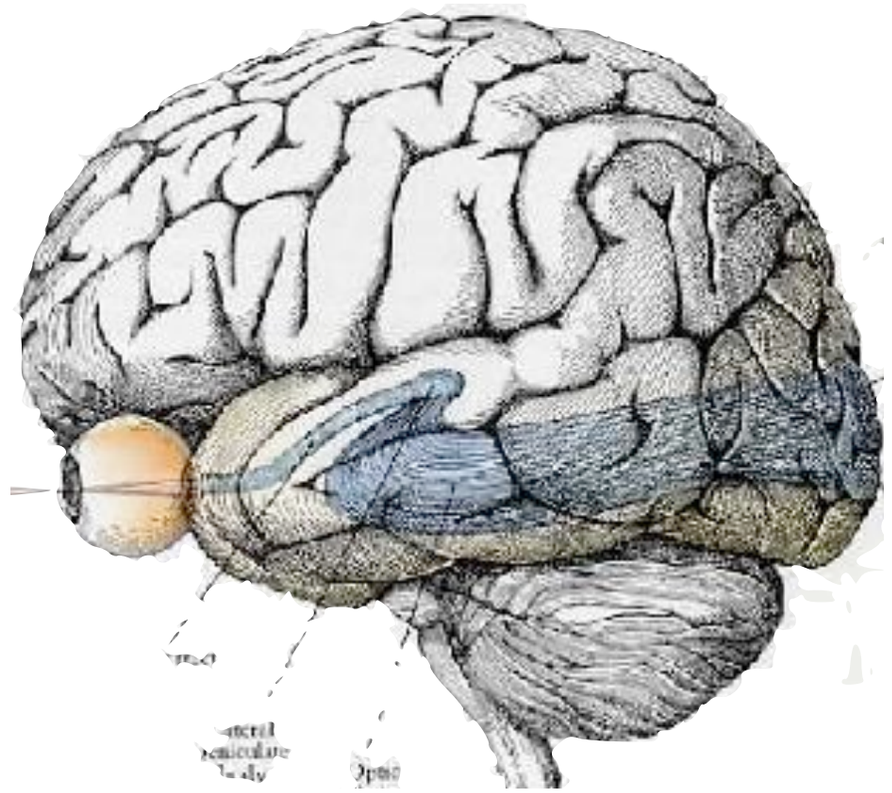
Source: Portilla, Javier, and Eero P. Simoncelli. "A parametric texture model based on joint statistics of complex wavelet coefficients." International journal of computer vision 40, no. 1 (2000): 49-70.

Julesz ‘78

Yellott ‘93

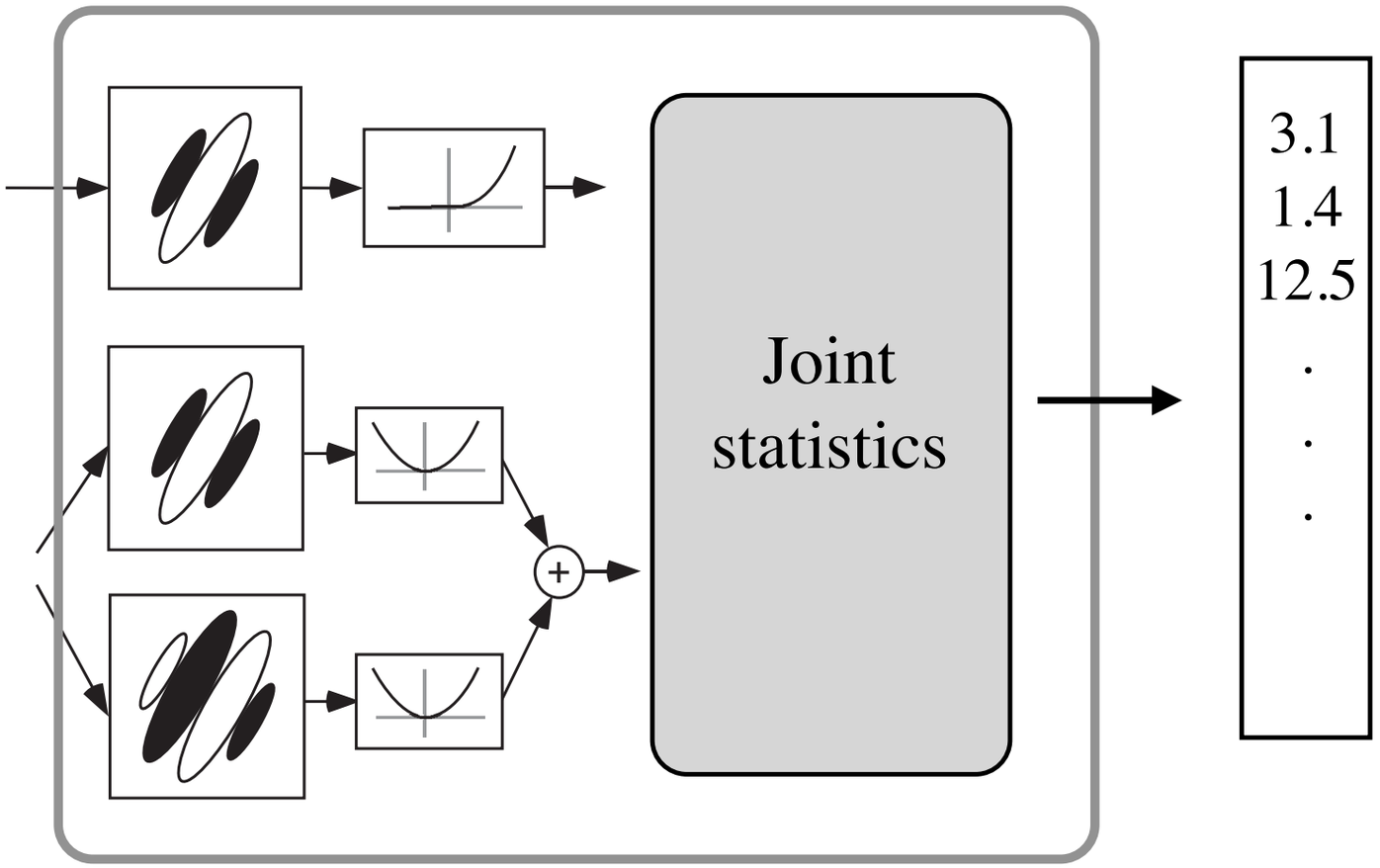
Physiologically-inspired Julesz-style texture model

Source: Portilla, Javier, and Eero P. Simoncelli. "A parametric texture model based on joint statistics of complex wavelet coefficients." International journal of computer vision 40, no.1 (2000): 49-70.



[Portilla & Simoncelli, 2000]

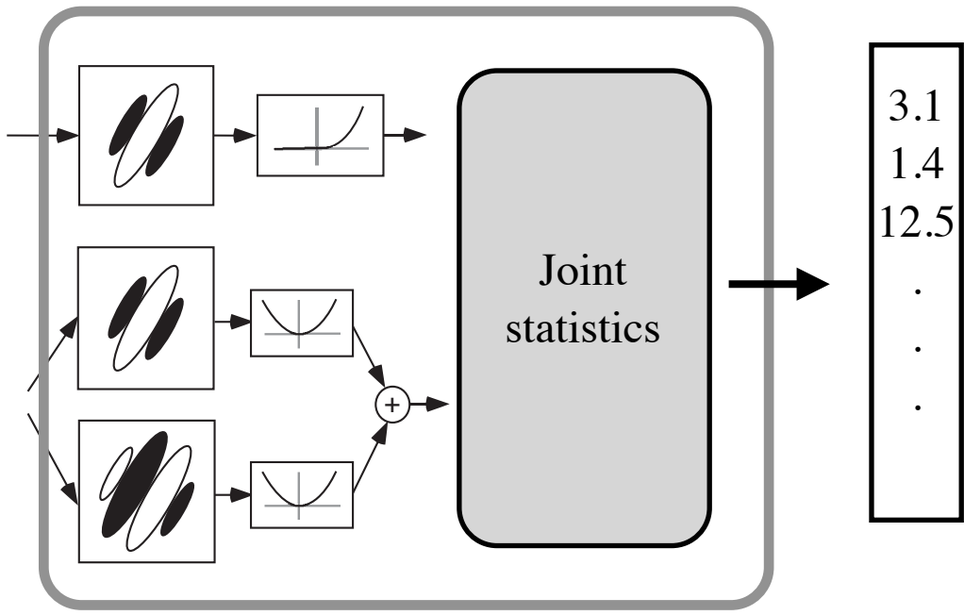
Physiologically-inspired Julesz-style texture model



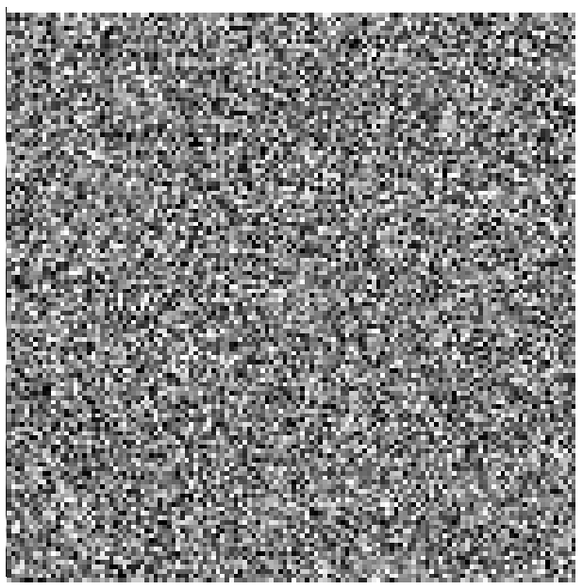
Source: Portilla, Javier, and Eero P. Simoncelli. "A parametric texture model based on joint statistics of complex wavelet coefficients." International journal of computer vision 40, no.1 (2000): 49-70.

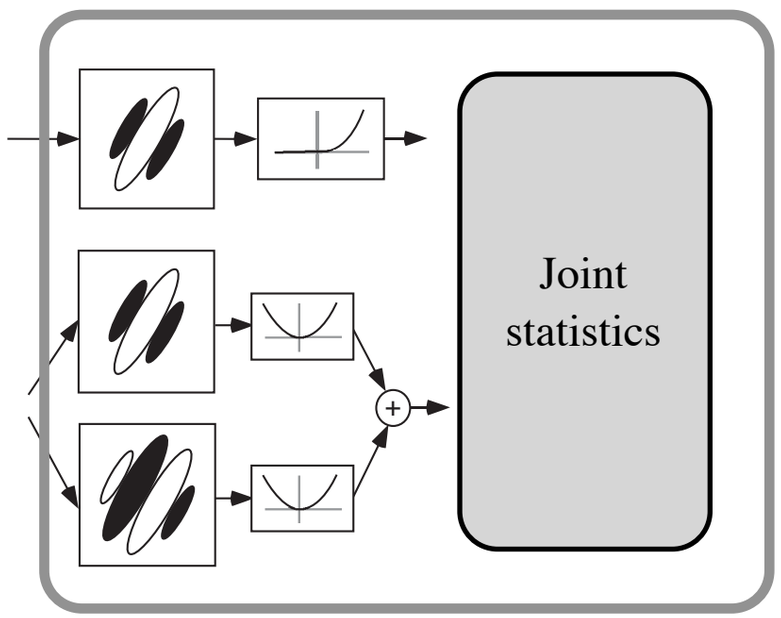
[Portilla & Simoncelli, 2000]

Texture synthesis



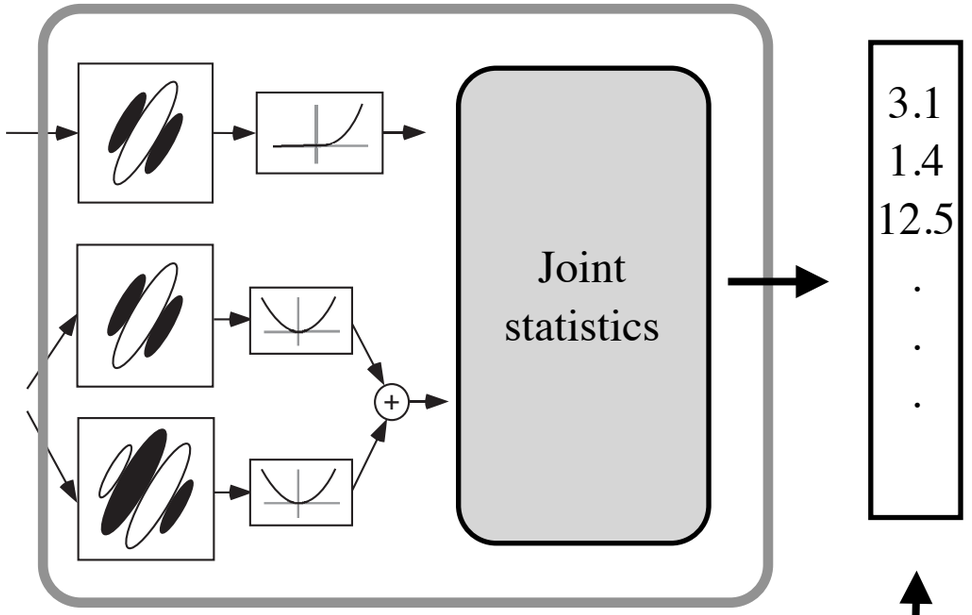
Source: Portilla, Javier, and Eero P. Simoncelli. "A parametric texture model based on joint statistics of complex wavelet coefficients." International journal of computer vision 40, no.1 (2000): 49-70.

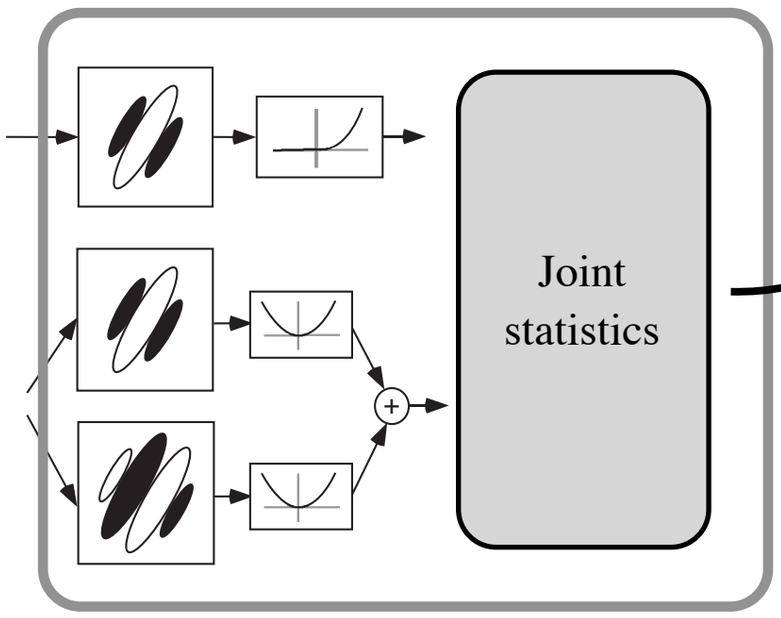




[Portilla & Simoncelli, 2000]

Texture synthesis





Source: Portilla, Javier, and Eero P. Simoncelli. "A parametric texture model based on joint statistics of complex wavelet coefficients." International journal of computer vision 40, no.1 (2000): 49-70.

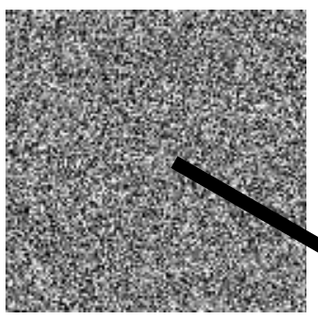
[Portilla & Simoncelli, 2000]

Images

original image

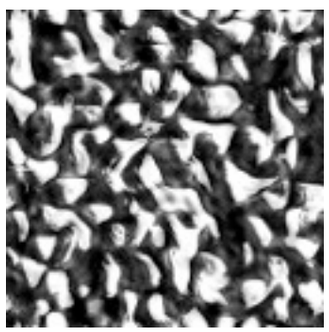
Source: Portilla, Javier, and Eero P. Simoncelli. "A parametric texture model based on joint statistics of complex wavelet coefficients." International journal of computer vision 40, no. 1 (2000): 49-70.

noise seed



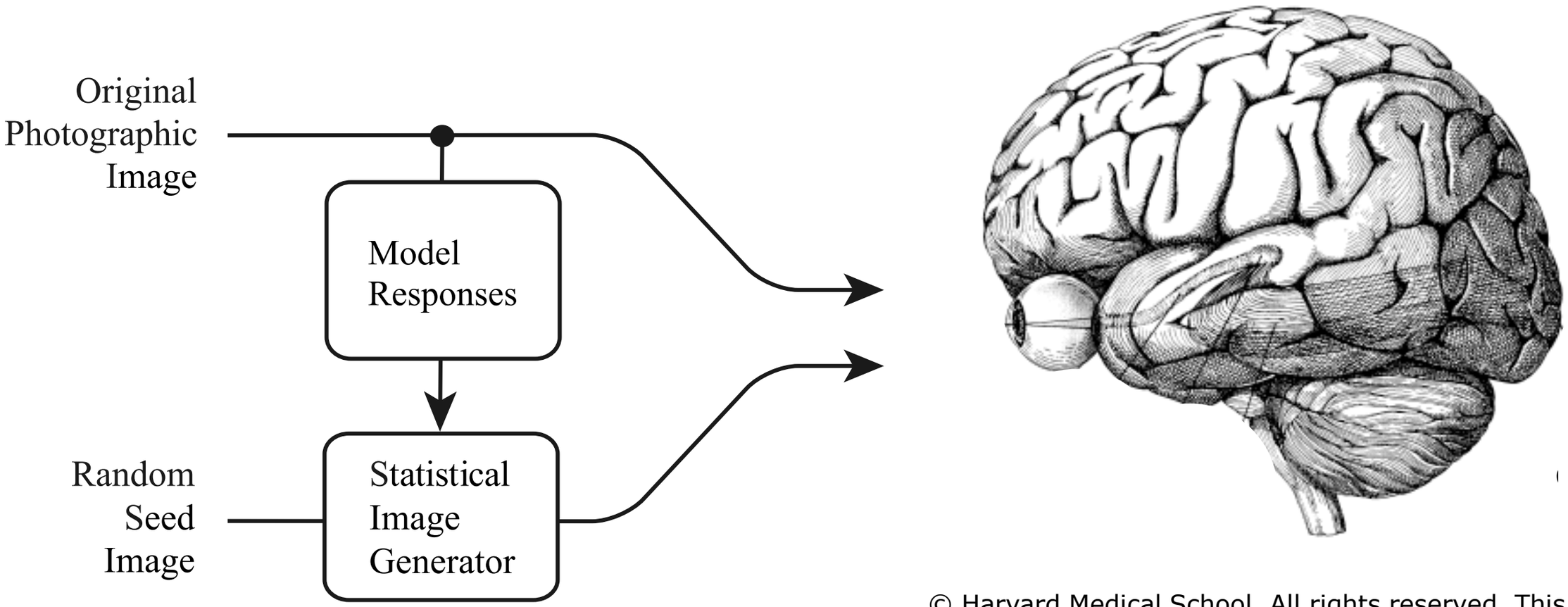
Images with identical model responses

synthesized



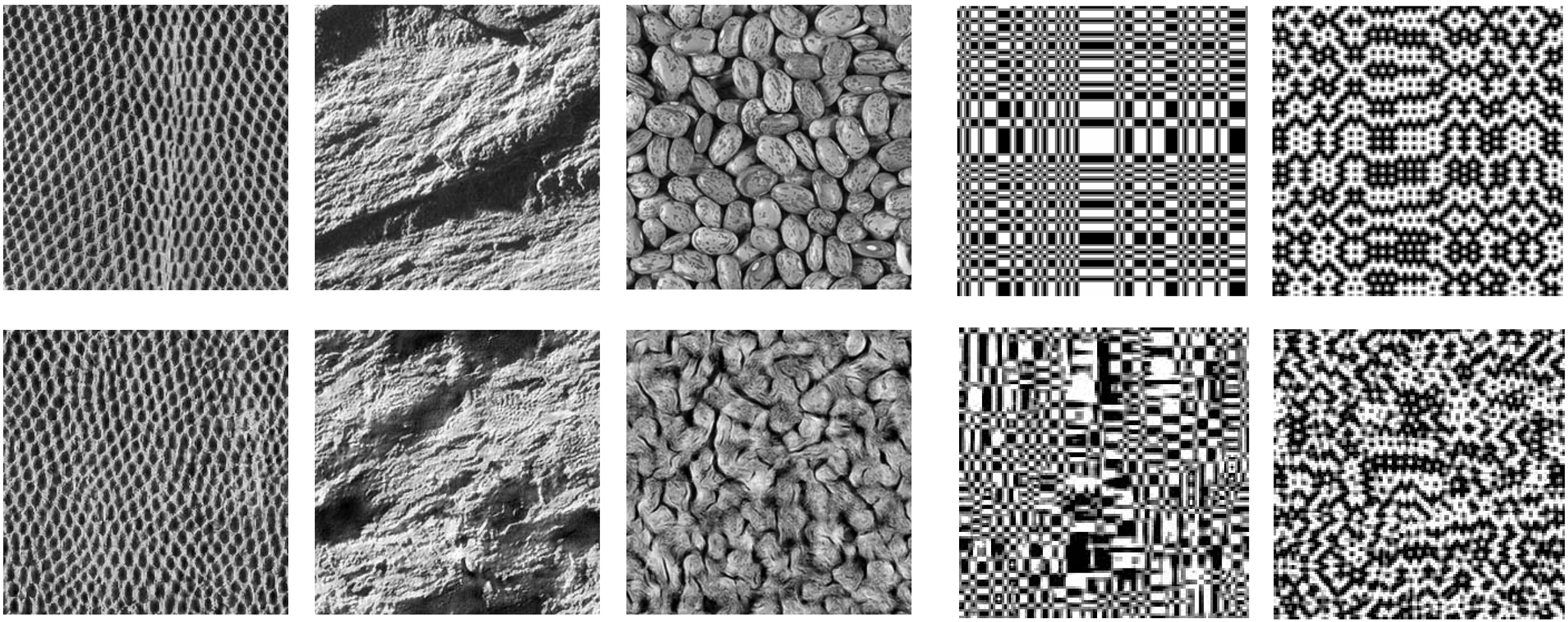
Model responses

Experimental logic



If model captures the same properties as the visual system, images with identical model responses should appear identical to a human.

Pairs of images with identical model responses:



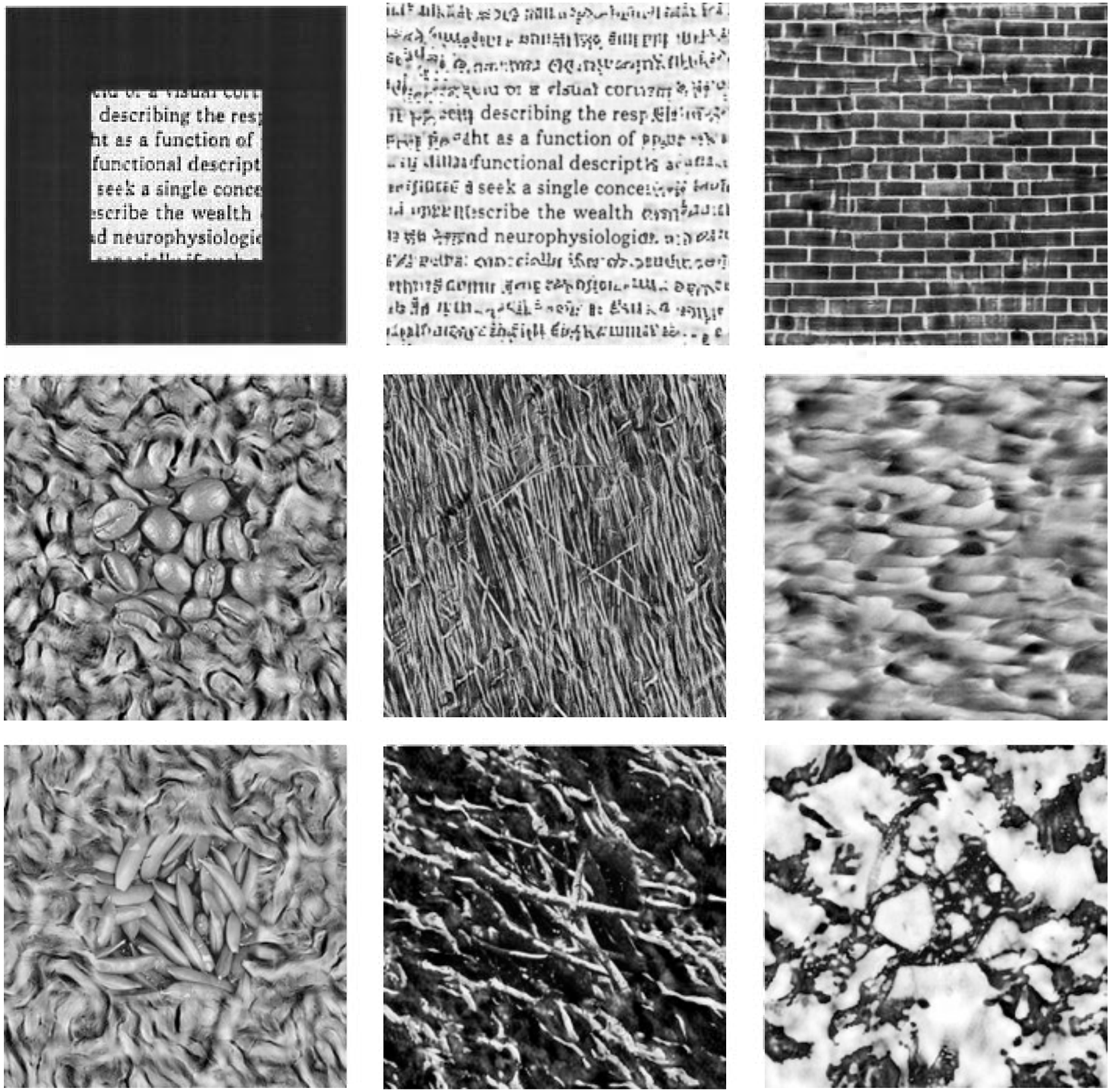
Source: Portilla, Javier, and Eero P. Simoncelli. "A parametric texture model based on joint statistics of complex wavelet coefficients." International journal of computer vision 40, no. 1 (2000): 49-70.

Top: original, Bottom: synthesized

[Portilla & Simoncelli 2000]

“outpainting”

Central square of each image is original texture. Surround is synthesized.

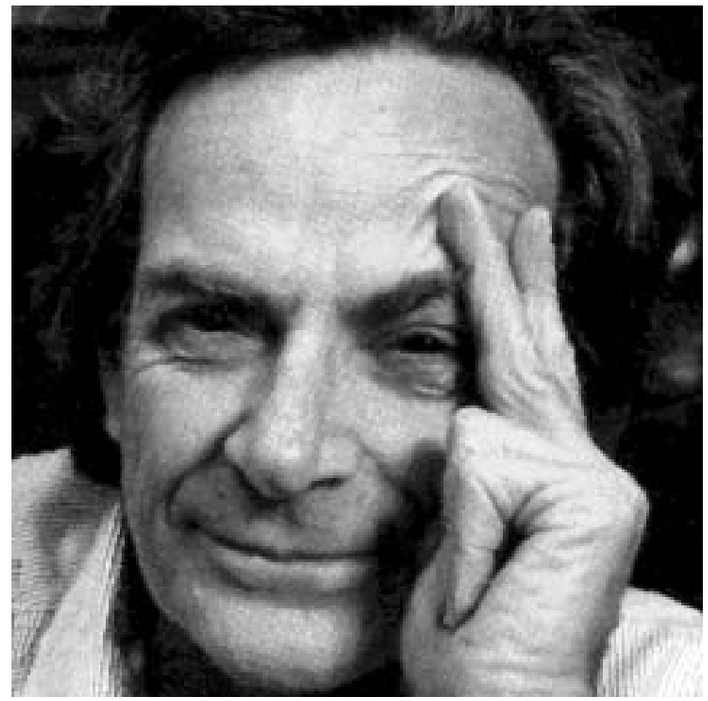


Source: Portilla, Javier, and Eero P. Simoncelli. "A parametric texture model based on joint statistics of complex wavelet coefficients." International journal of computer vision 40, no. 1 (2000): 49-70.

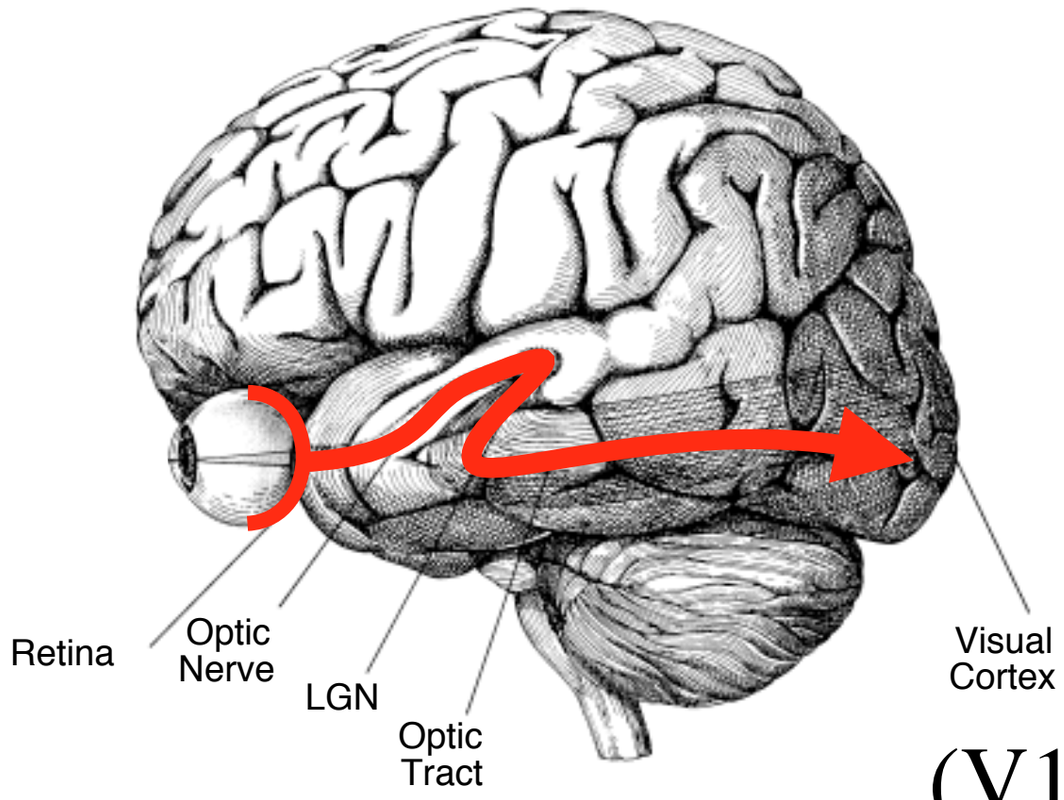
Structural seeding [cf. “adversarial examples” - Szegedy et. al. 2014]

Can we generalize to inhomogeneous stimuli?

Can we make the model more physiological?

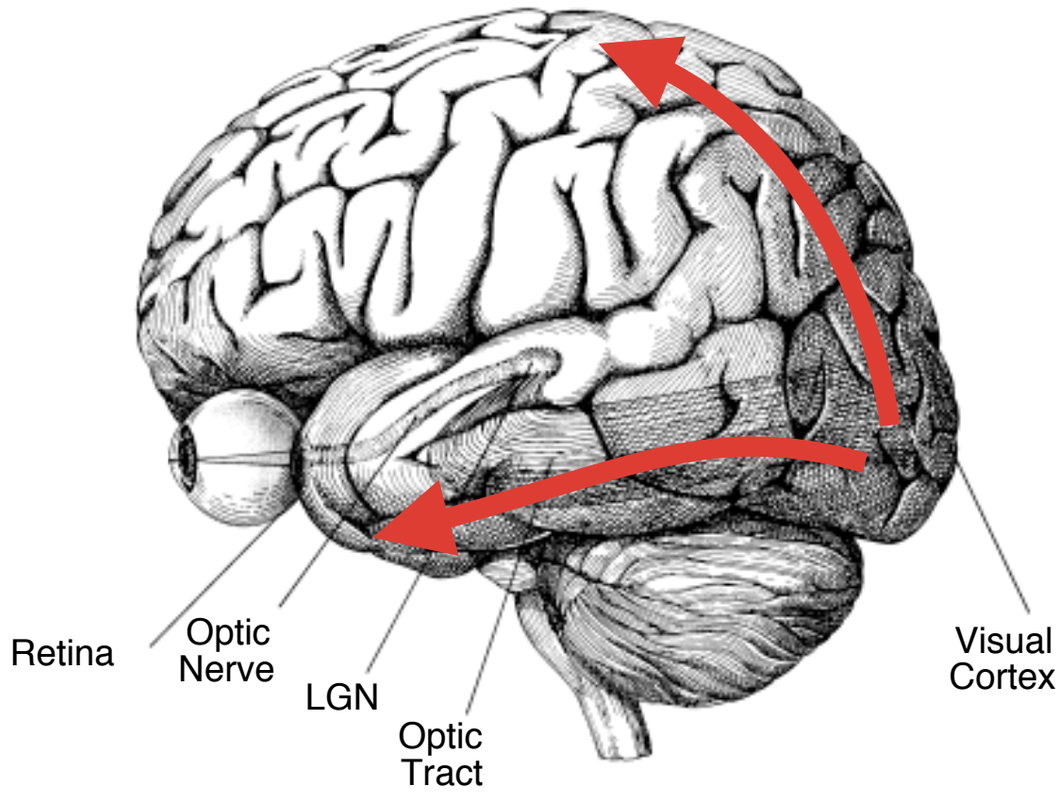


Source: Portilla, Javier, and Eero P. Simoncelli. "A parametric texture model based on joint statistics of complex wavelet coefficients." International journal of computer vision 40, no.1 (2000): 49-70.



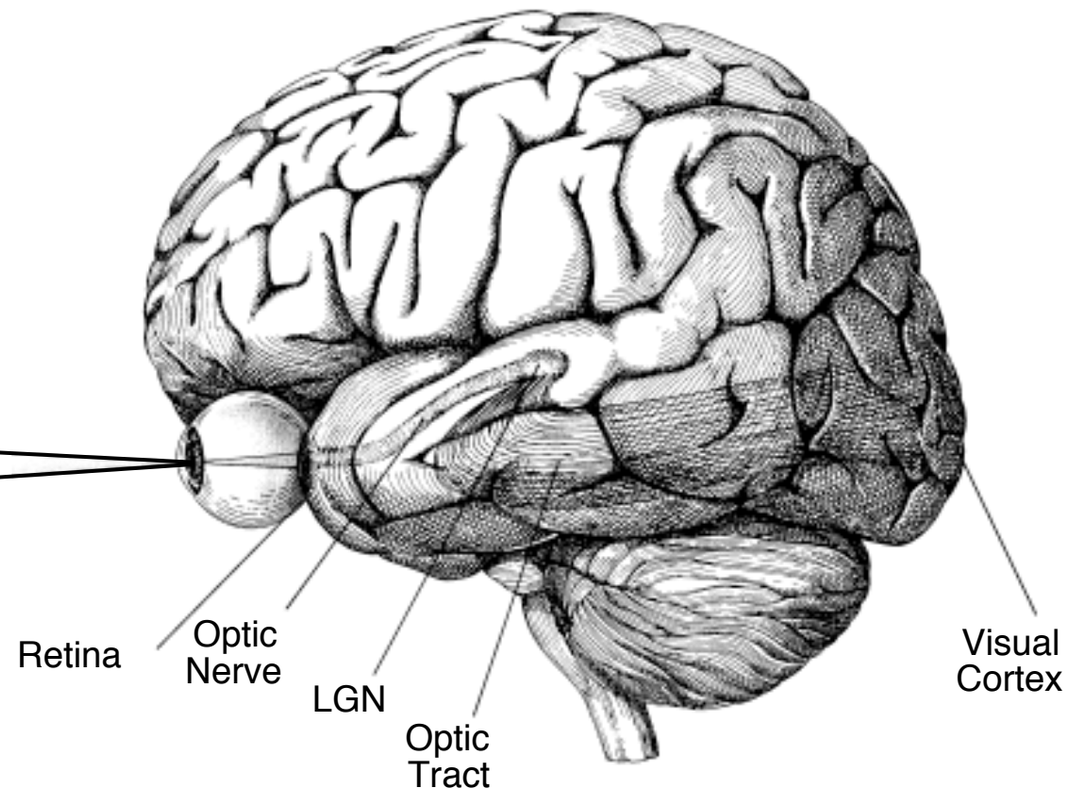
[figure: Hubel ‘95]

Dorsal pathway: V1->V3->V5
position, motion, action



Ventral pathway: V1->V2->V4-> IT
spatial form, recognition, memory

[Ungerleider & Mishkin, 1982]



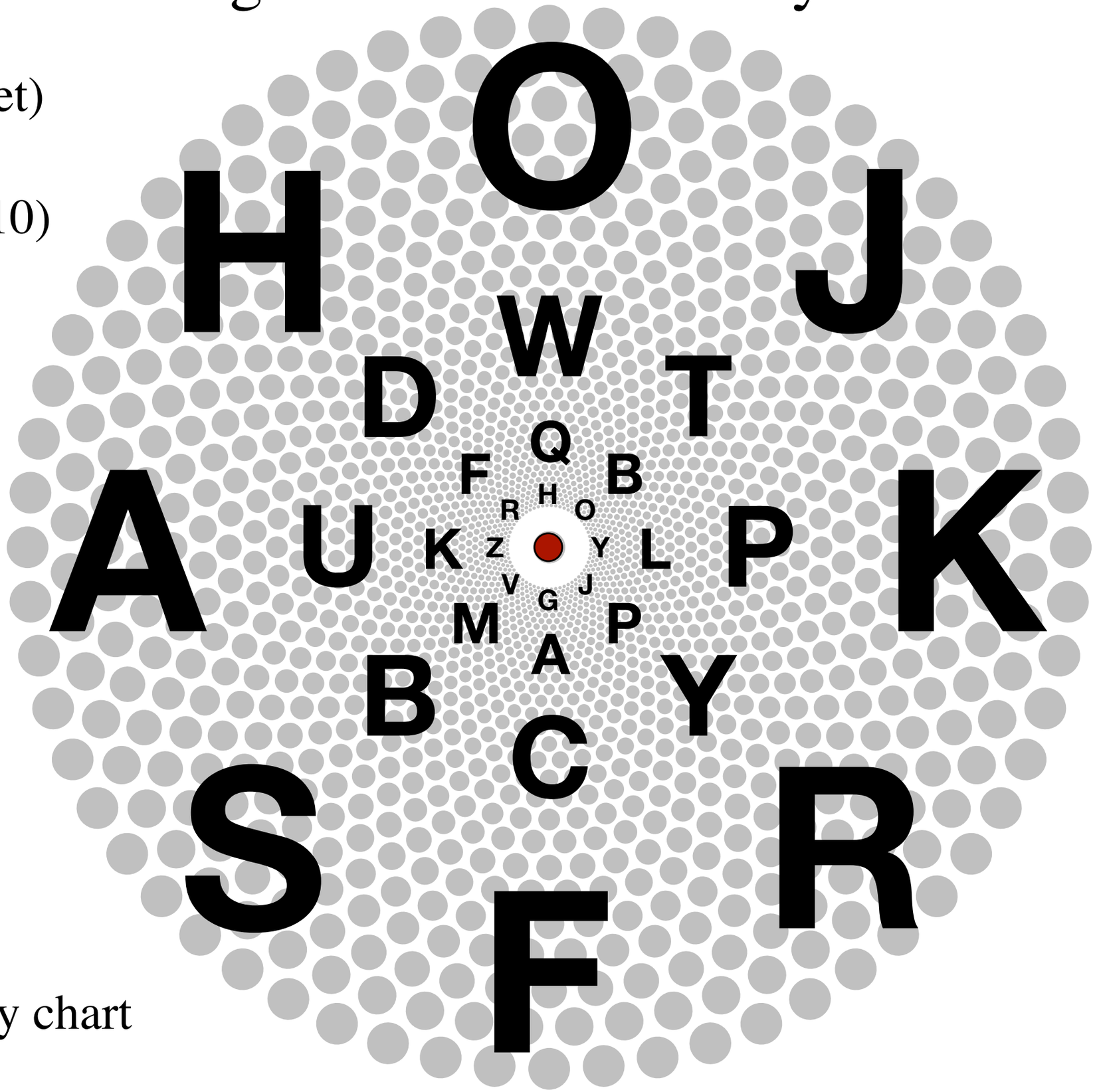
- Visual neurons responds to content within a small region of the visual input known as the **Receptive Field (RF)**

- In each visual area, we assume RFs cover the entire visual field

Inhomogeneity - RF sizes grow with eccentricity

Retinal ganglion (midget)
cell receptive fields
(macaque, magnified x10)
[Perry et.al., 1984;
Watanabe & Rodiek, 1989]

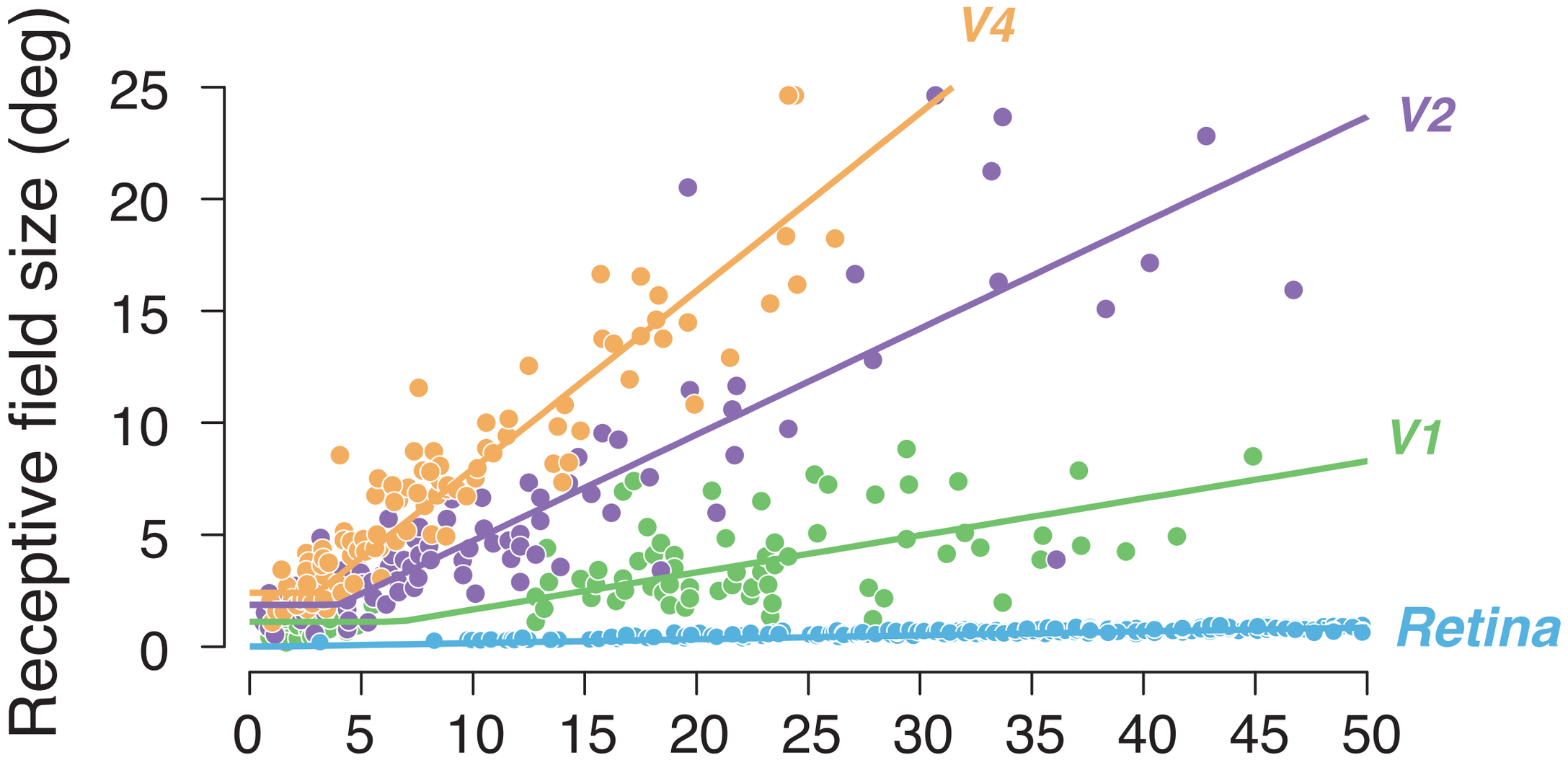
loss of resolution



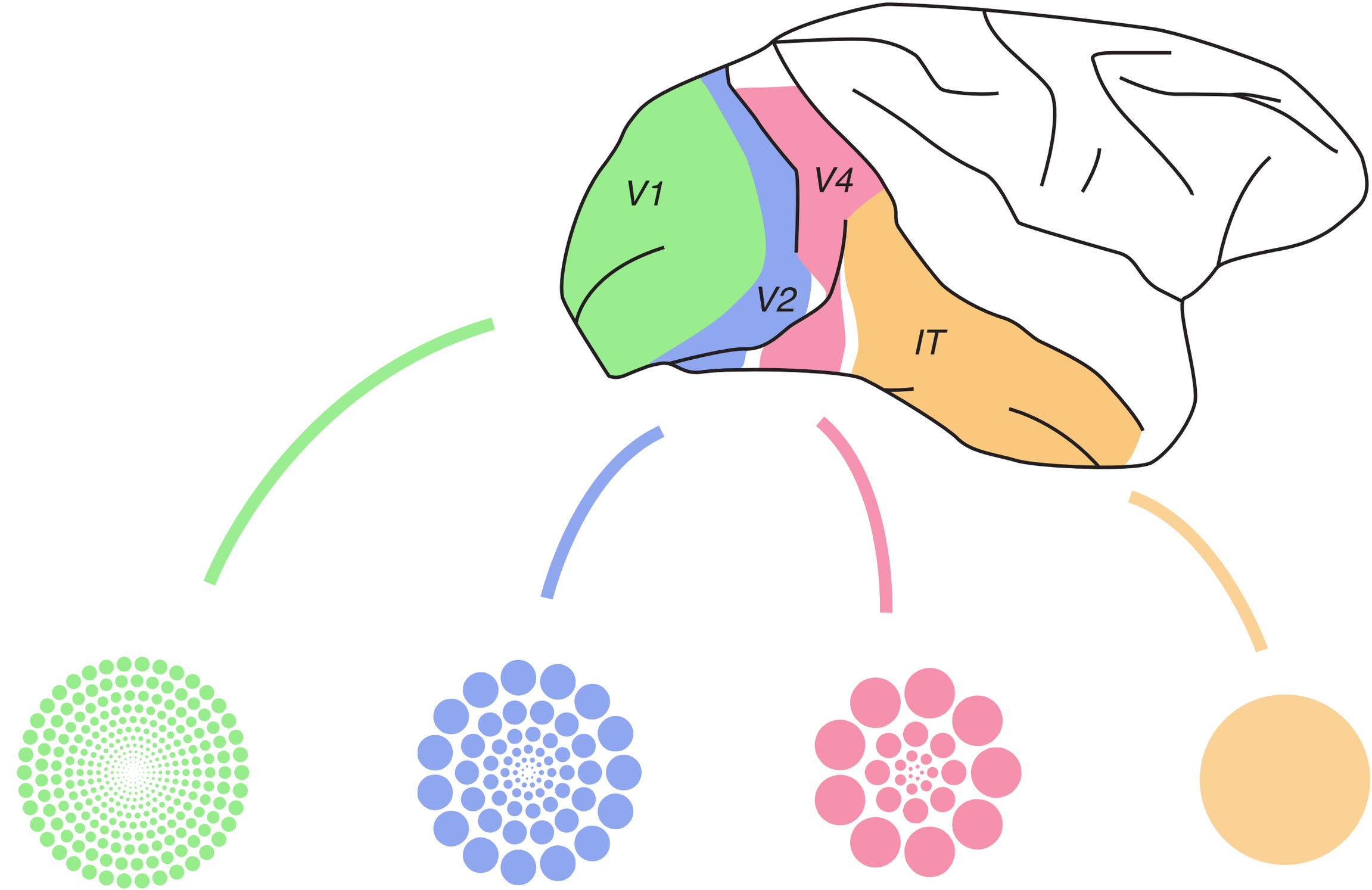
Modified Snellen acuity chart
(threshold, x10)
[after Anstis, 1973]

[after Geisler et al., 1999]

RF sizes grow with eccentricity



[Freeman & Simoncelli 2011,
data from Gattass et. al., 1981; Gattass et. al., 1988; Perry et. al., 1984]



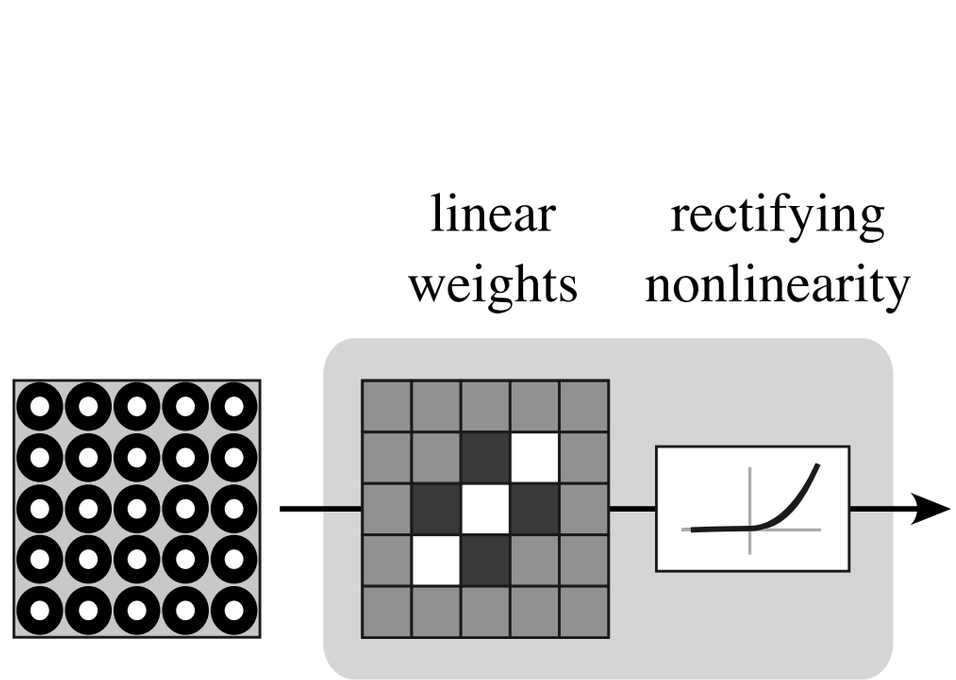
[Freeman & Simoncelli, 2011]

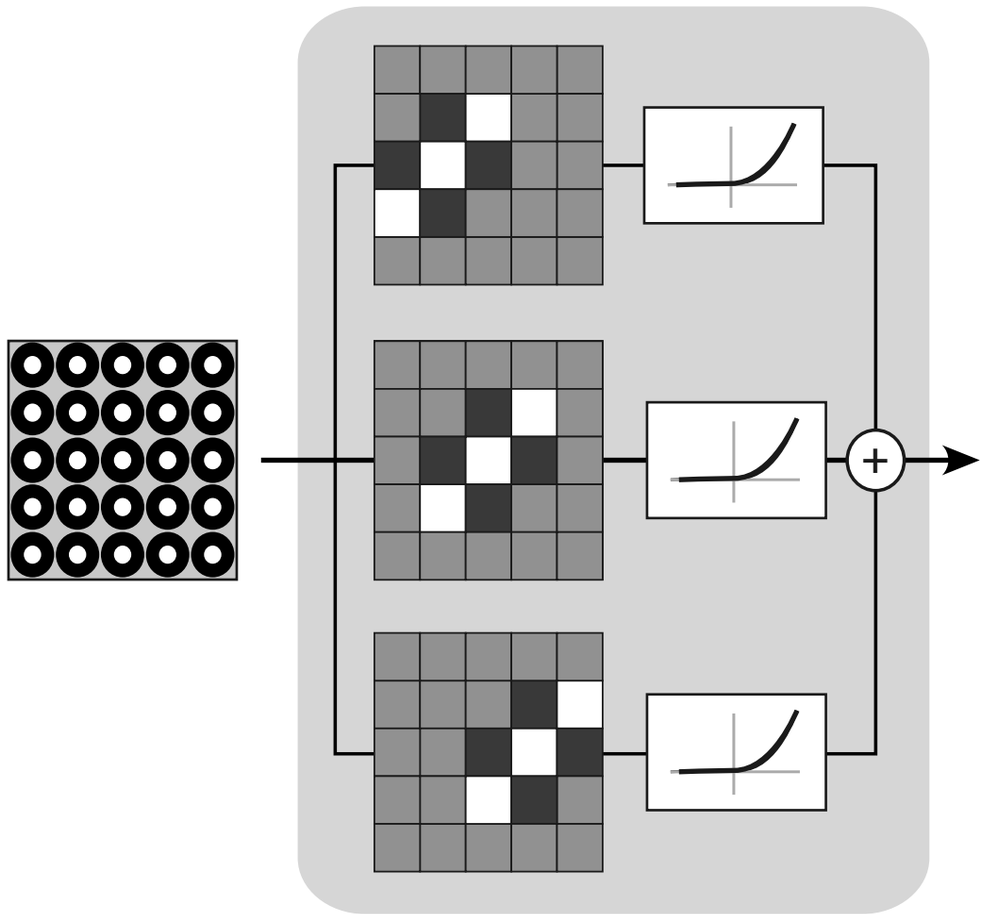
V1 simple cell

V1 complex cell

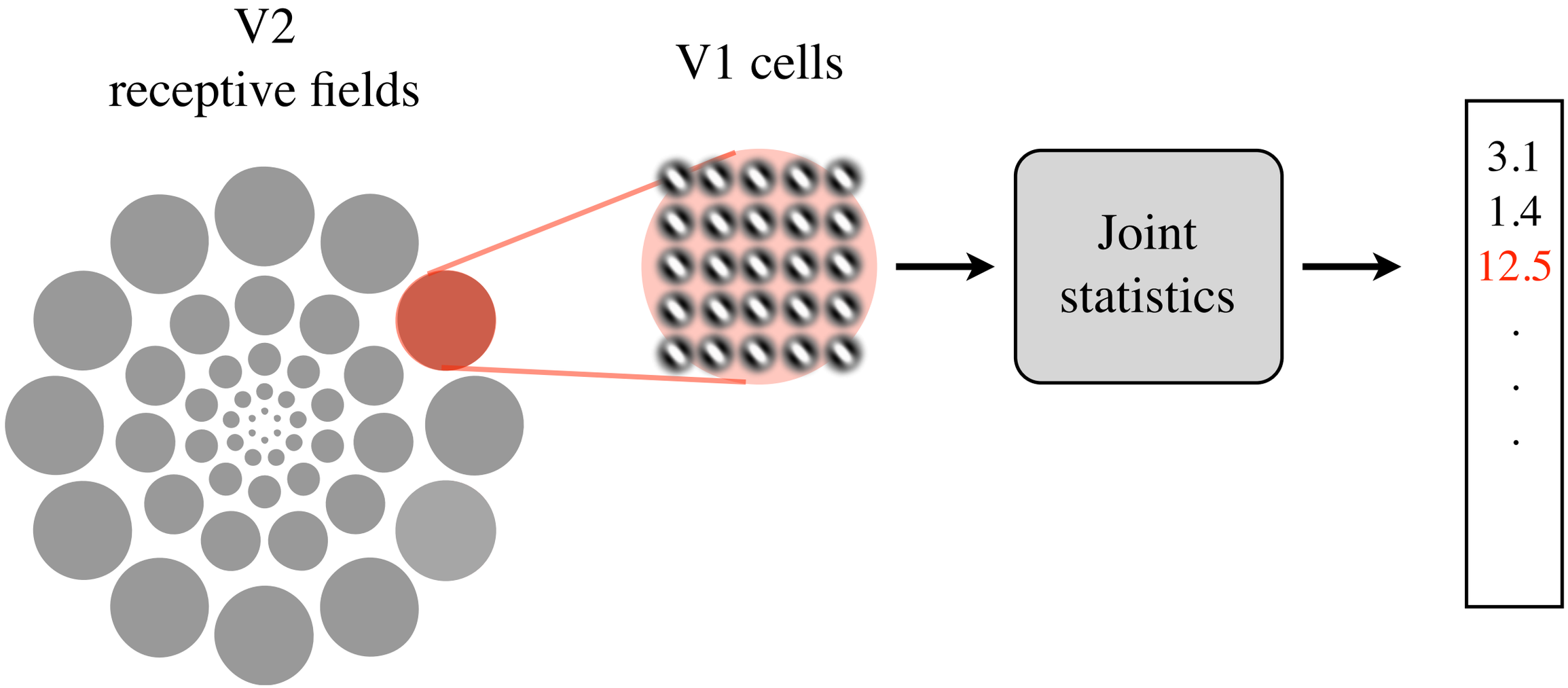
Figure removed due to copyright restrictions. Please see the video.
Source: Hubel, David H., and Torsten N. Wiesel. "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex." The Journal of physiology 160, no. 1 (1962): 106-154.

[Hubel & Wiesel, 1962]



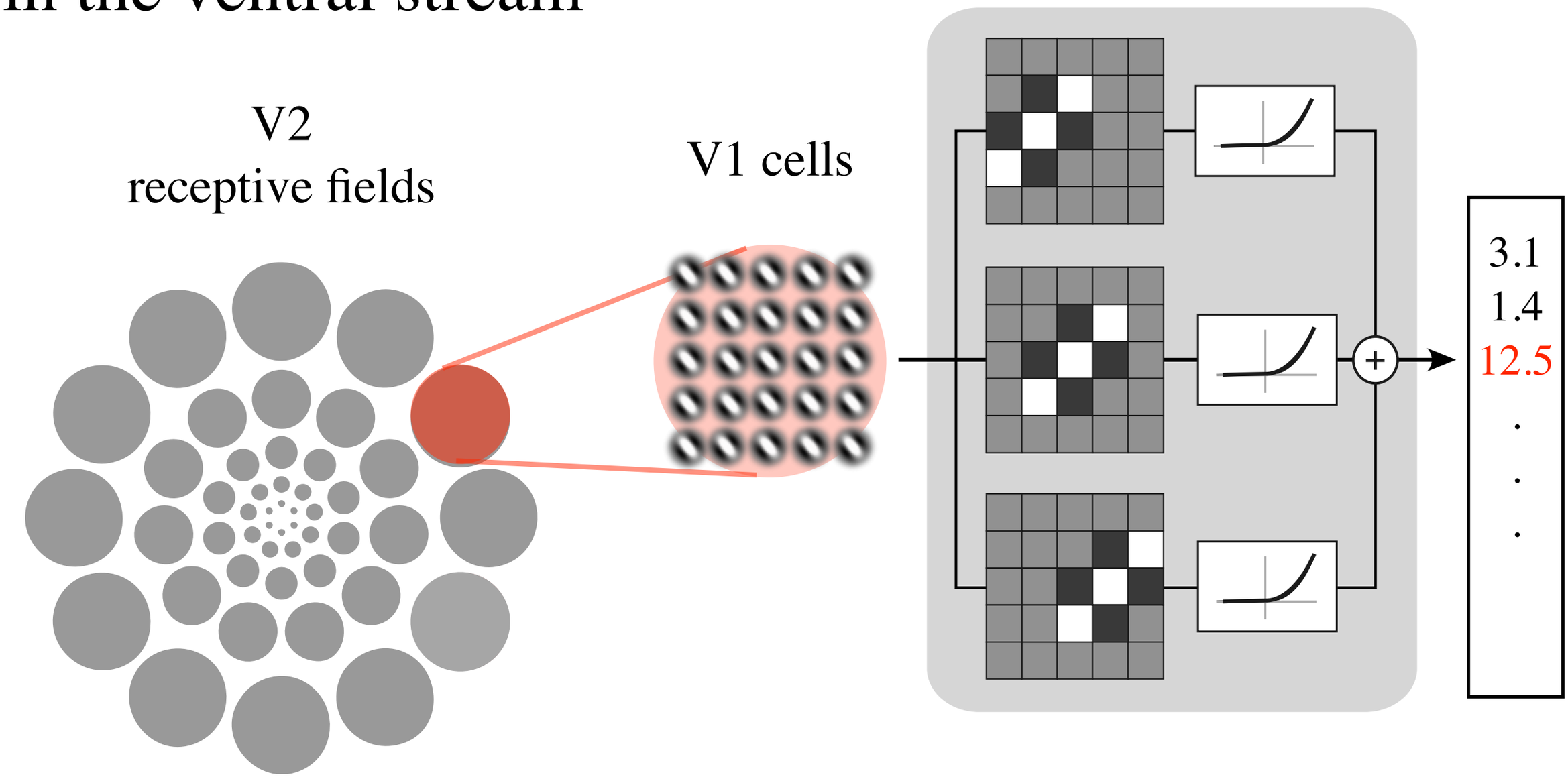


Local texture representation in the ventral stream



Local correlational statistics can be re-expressed as a “subunit” model...

Canonical computation in the ventral stream



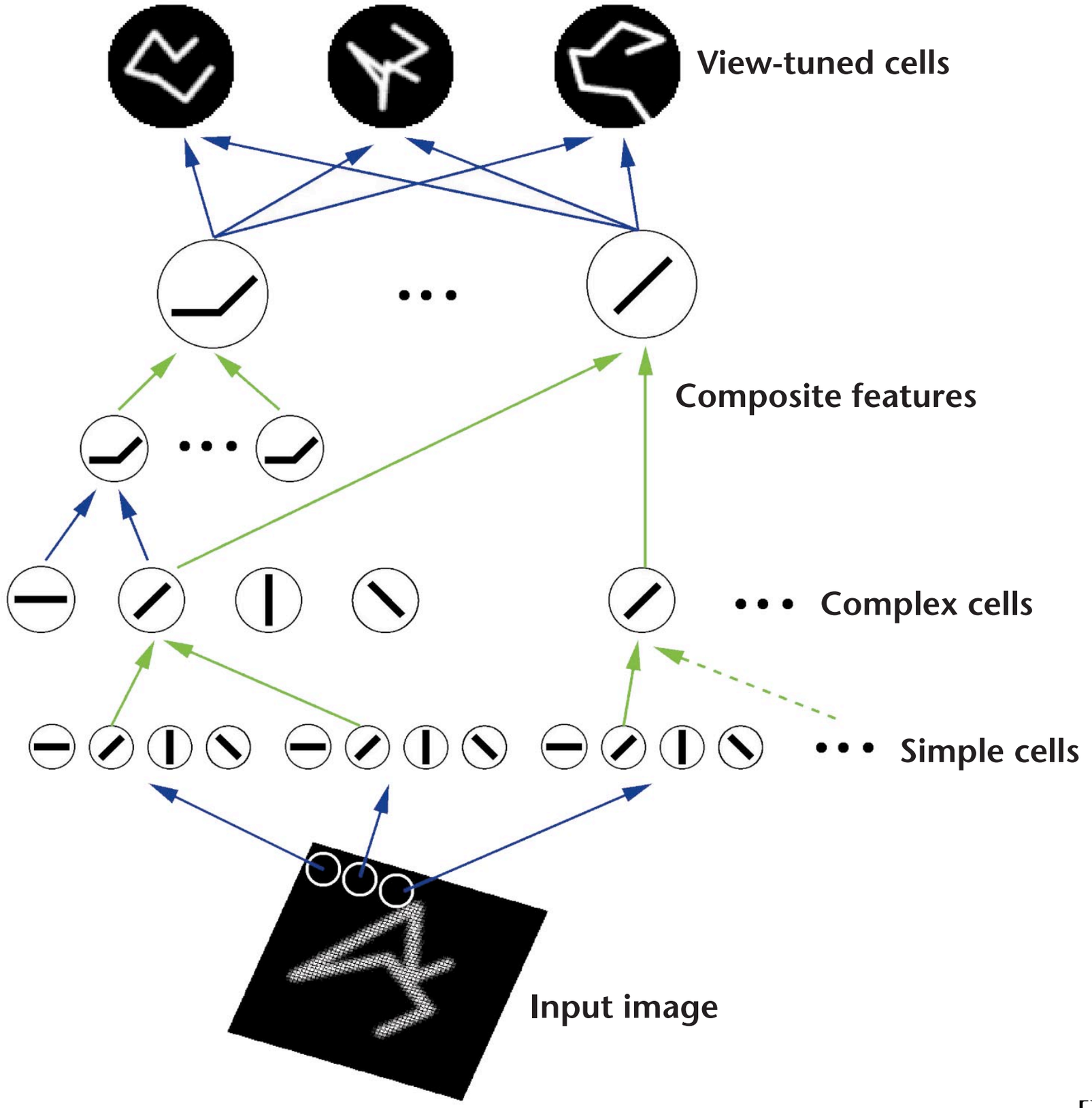
Substantial information loss => model predicts **metamers**

Canonical sensory computation

- Linear filter (determines pattern selectivity)
- Rectifying nonlinearity
- Local pooling (e.g., average, max)
- Local gain control
- Noise

Cascaded ...

[eg. Douglas, 1989;
Heeger, Simoncelli & Movshon 1996;
Heeger & Carandini 2014]



Reprinted by permission from Macmillan Publishers Ltd: Nature.
Source: Koch, Christof, and Tomaso Poggio. "Predicting the visual world: Silence is golden." Nature neuroscience 2, no. 1 (1999): 9-10. © 1999.

[Koch & Poggio, 1999;
cf. Fukishima, 1980;
Serre, Oliva, Poggio 2007; etc]

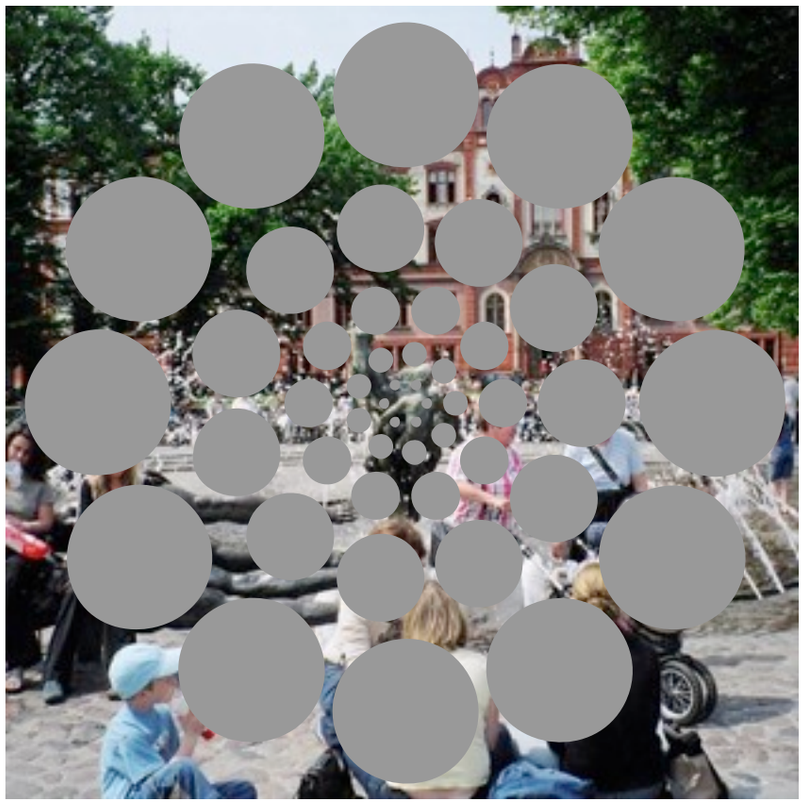
Synthesizing Ventral Stream Metamers

Original image

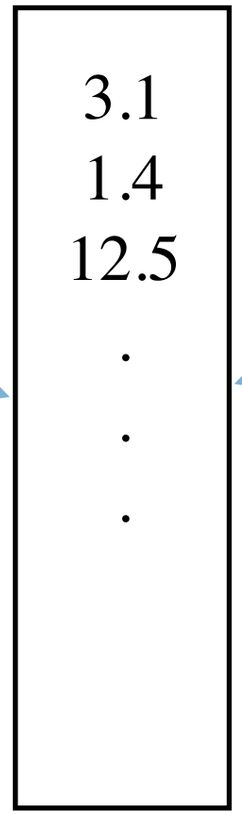
[Freeman & Simoncelli, 2011]

Synthesizing Ventral Stream Metamers

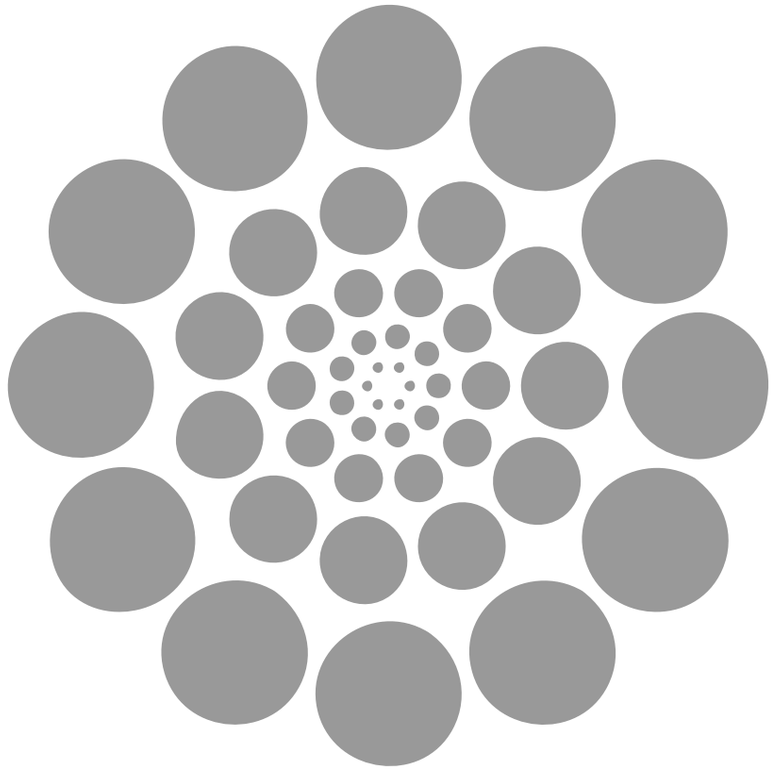
Original image



Model responses



Synthesized image



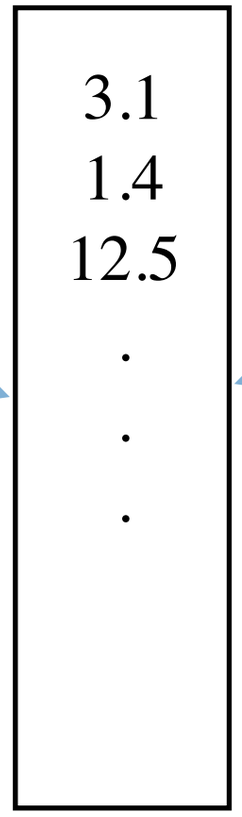
[Freeman & Simoncelli, 2011]

Synthesizing Ventral Stream Metamers

Original image

Model responses

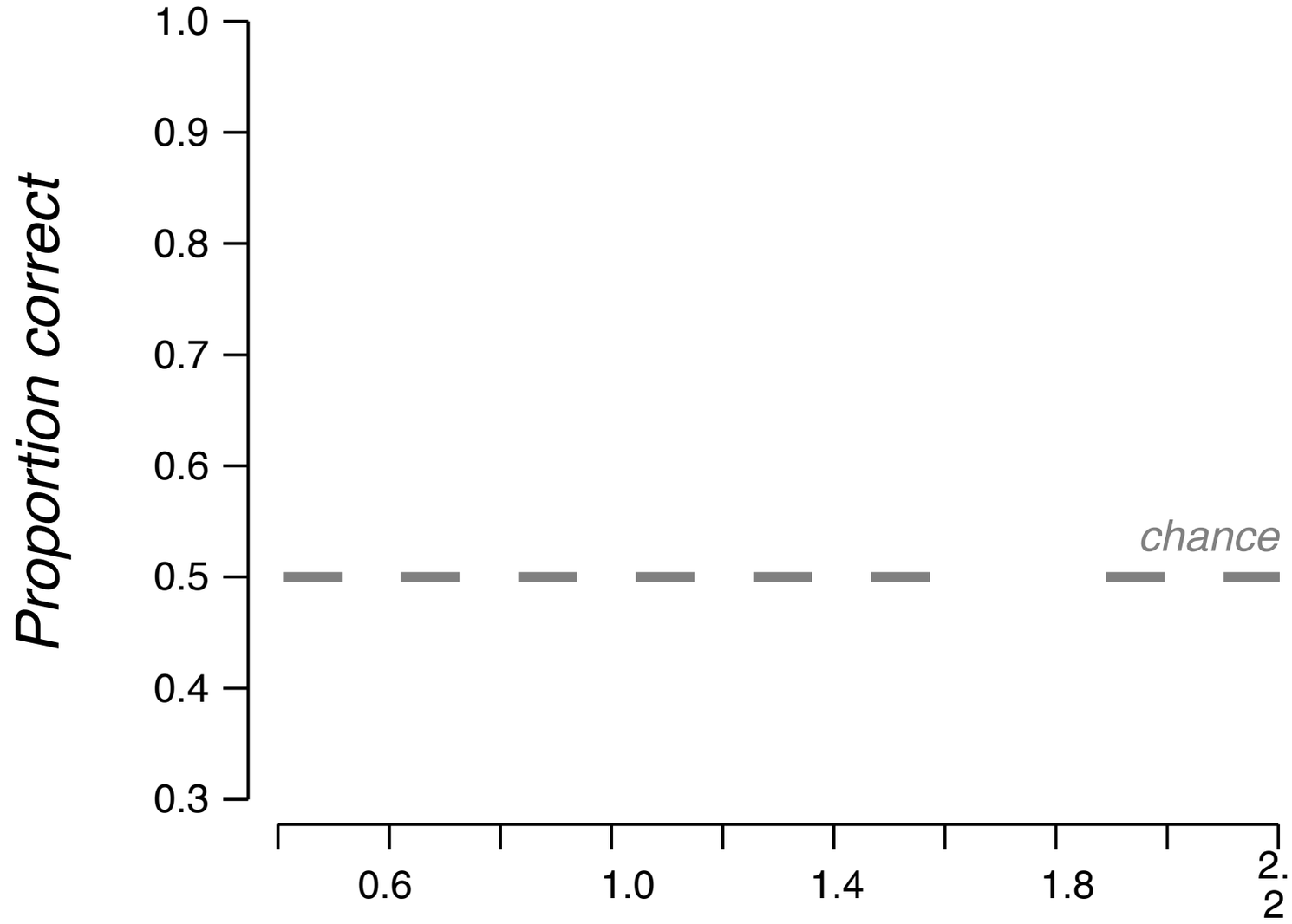
Synthesized image

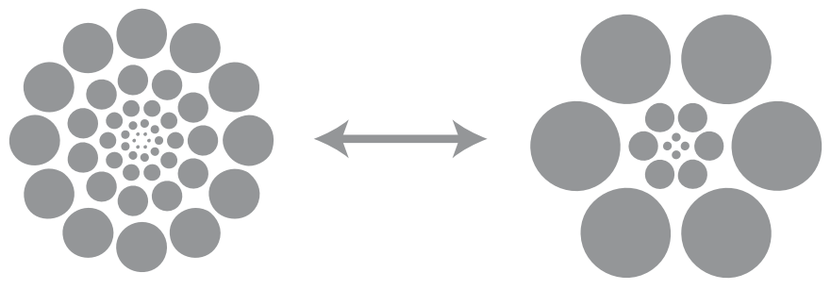


[Freeman & Simoncelli, 2011]

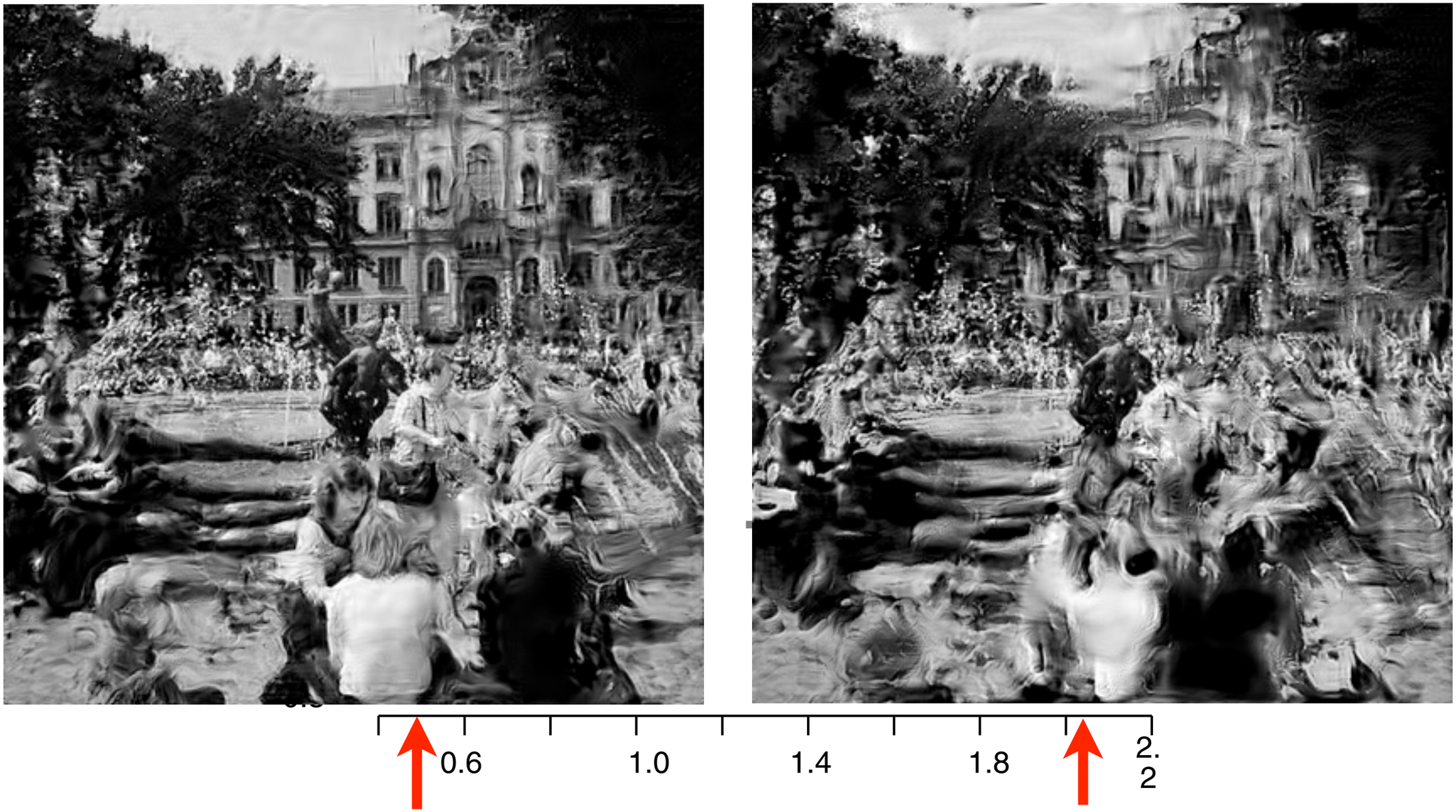
Courtesy of Elsevier, Inc., http://www.sciencedirect.com. Used with permission.
Source: Cohen, Michael A., Daniel C. Dennett, and Nancy Kanwisher. "What is the bandwidth of perceptual experience?" Trends in Cognitive Sciences 20, no. 5 (2016): 324-335.

Courtesy of Elsevier, Inc., http://www.sciencedirect.com. Used with permission.
Source: Cohen, Michael A., Daniel C. Dennett, and Nancy Kanwisher. "What is the bandwidth of perceptual experience?" Trends in Cognitive Sciences 20, no. 5 (2016): 324-335.

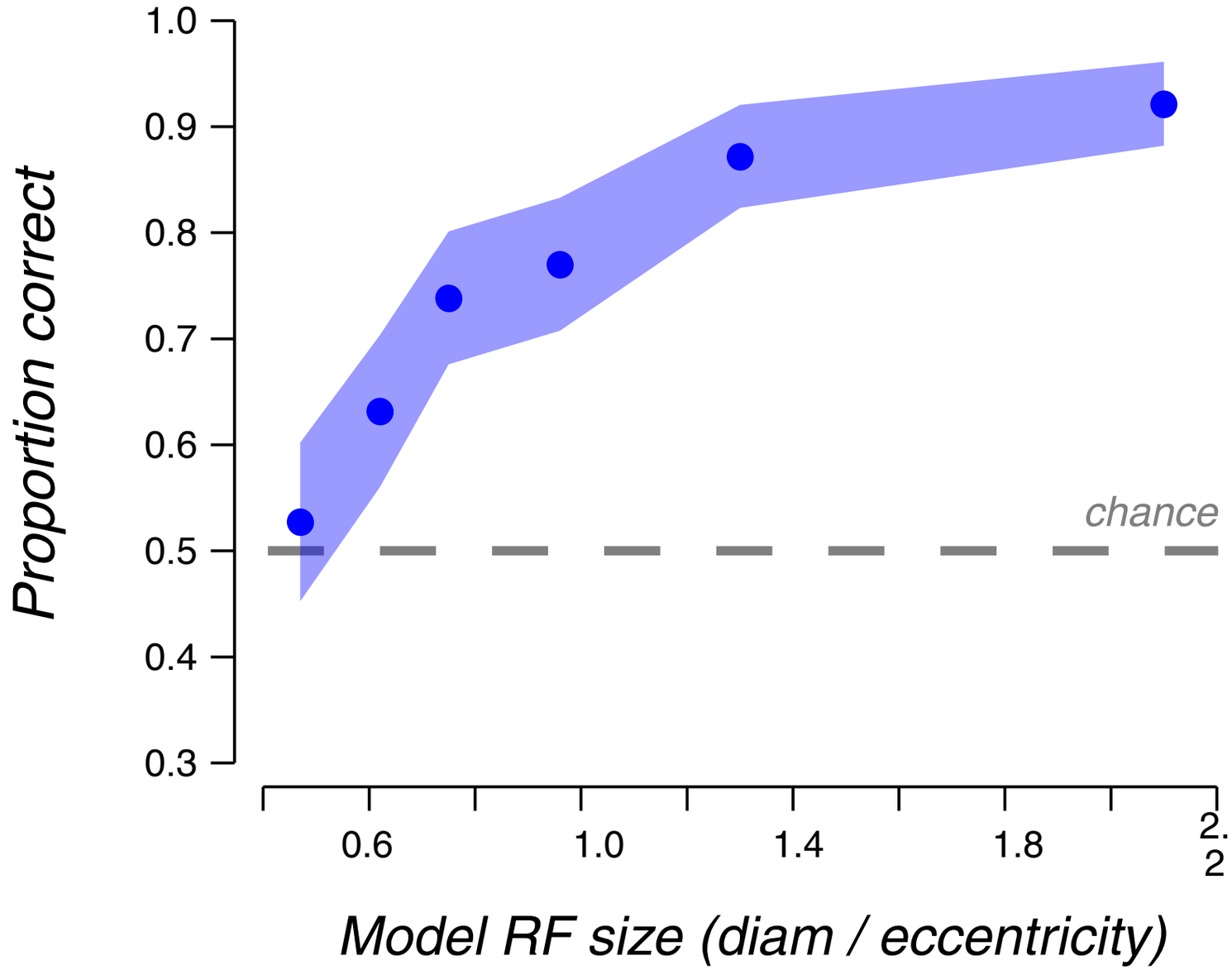


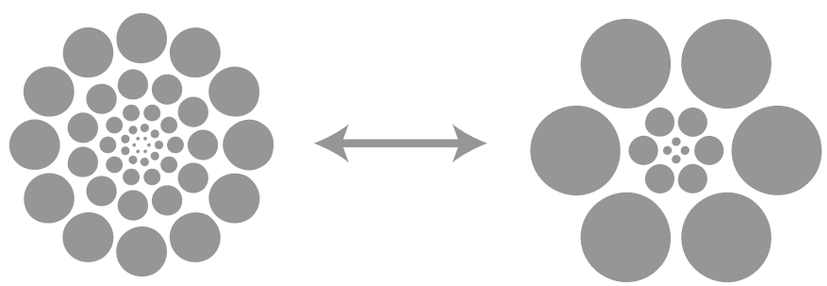


[Freeman & Simoncelli, 2011]

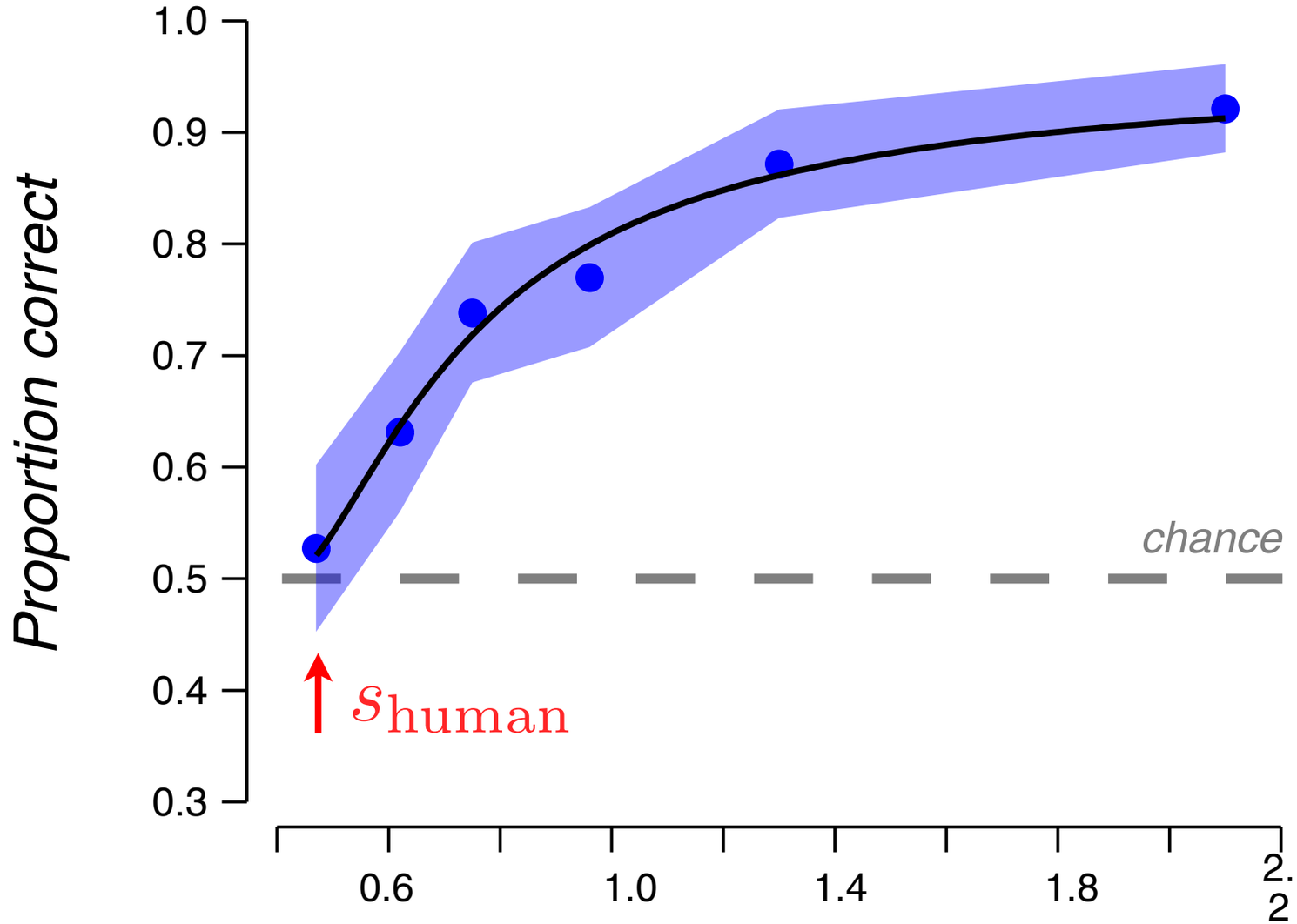


[Freeman & Simoncelli, 2011]



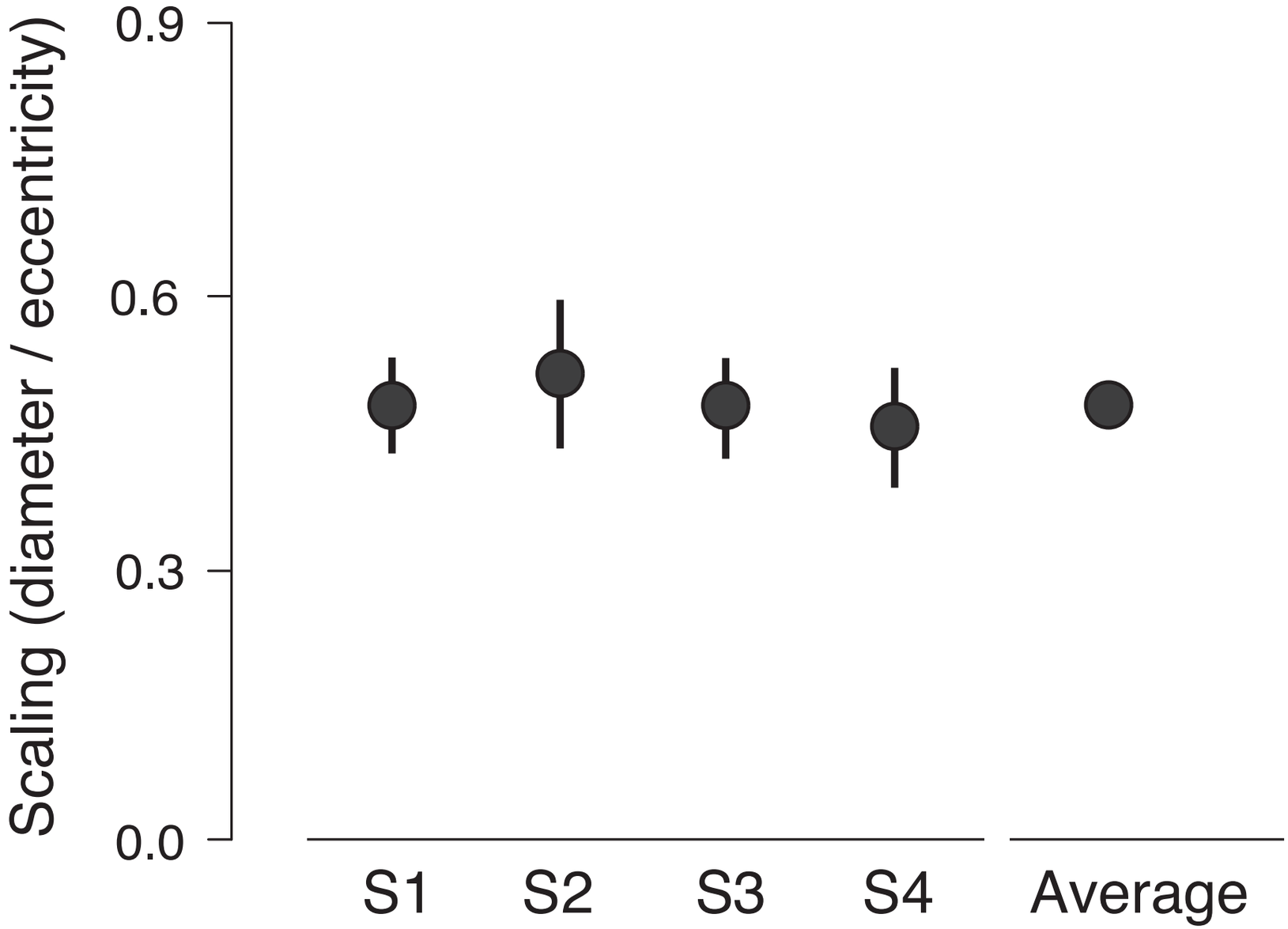


[Freeman & Simoncelli, 2011]



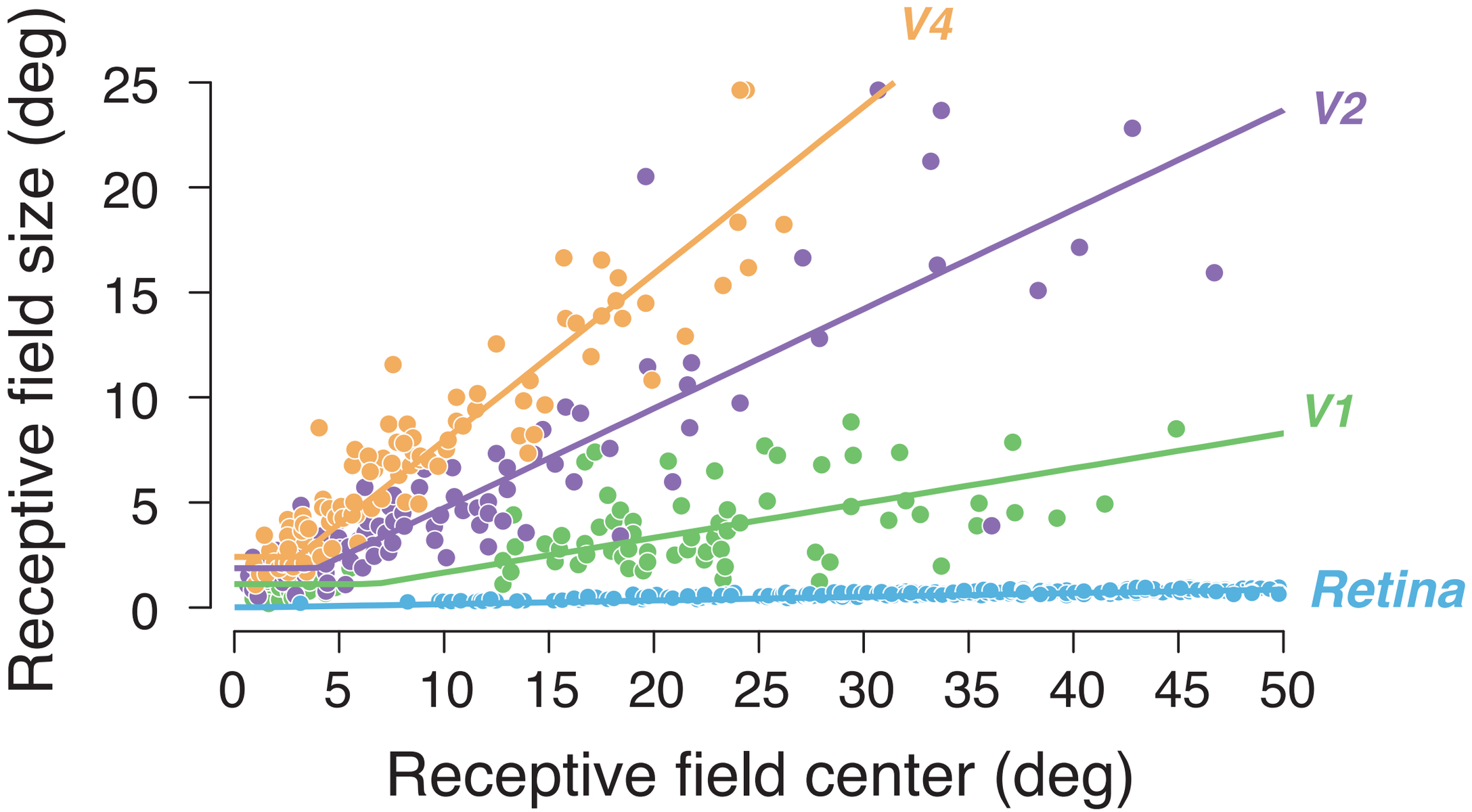
$$D = \Phi\left(\left\lfloor 1 - \frac{s^2_{\text{human}}}{s^2_{\text{model}}} \right\rfloor\right)$$

[Freeman & Simoncelli, 2011]

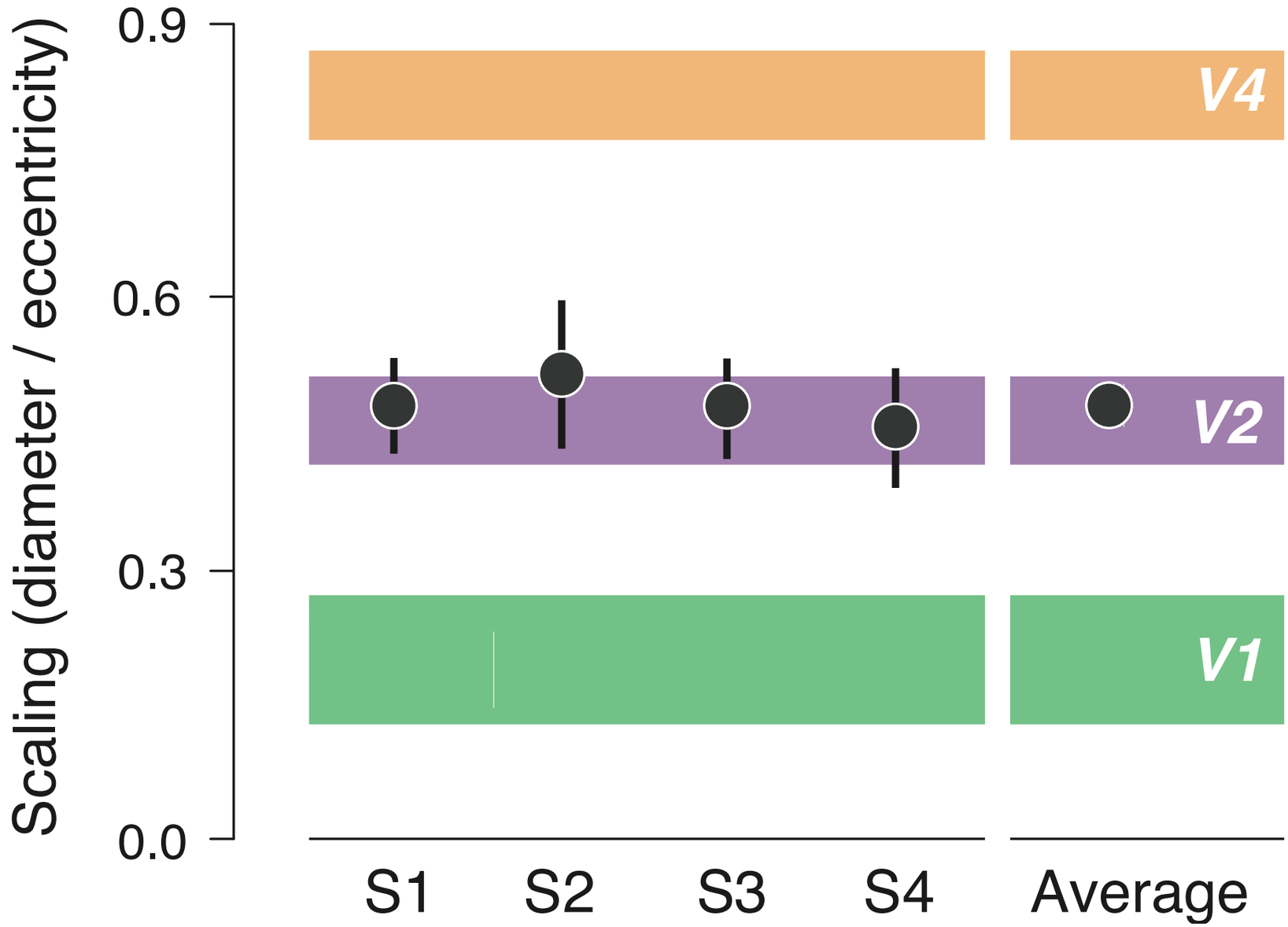


[Freeman & Simoncelli, 2011]

RF sizes grow with eccentricity



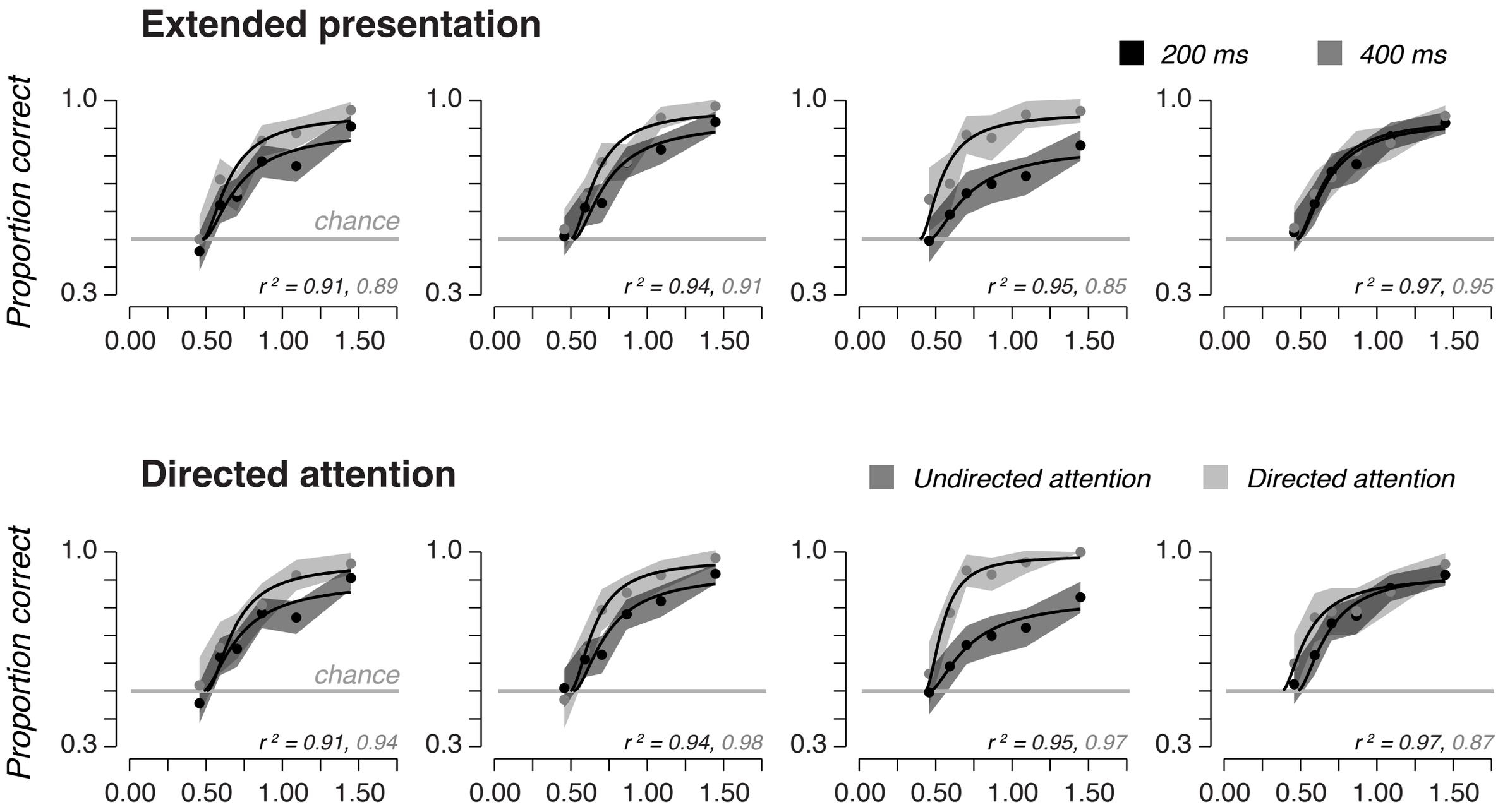
[Freeman & Simoncelli 2011,
from Gattass et. al., 1981; Gattass et. al., 1988; Perry et. al., 1984]



Macaque
Physiology

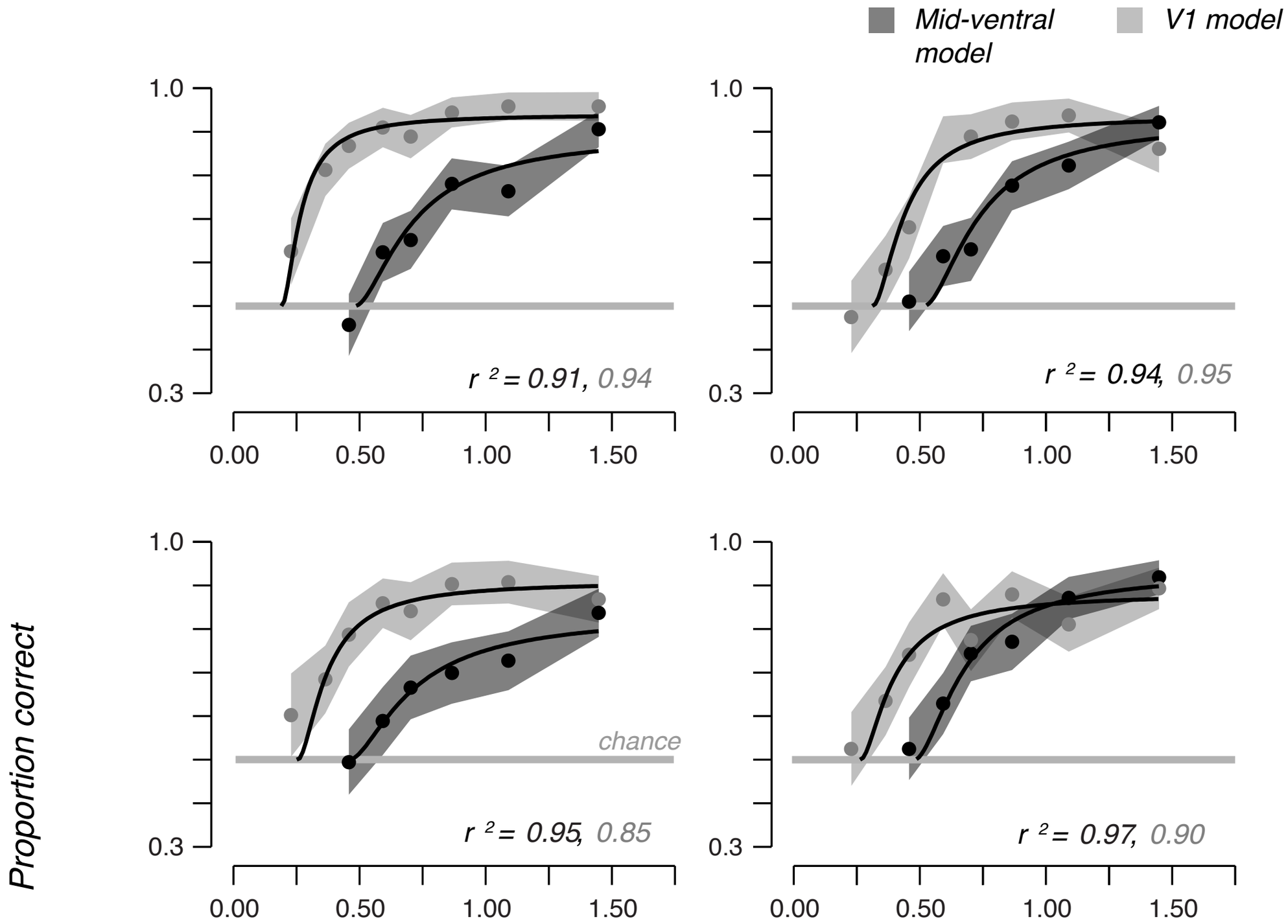
[Allman & Kaas, 1971; Allman & Kaas, 1974; Gattass et.al., 1981; van Essen et.al., 1984; Maguire & Baizer, 1984; Burkhalter & van Essen, 1986; Gattass et.al., 1987; Desimone & Schein, 1987; Gattass et.al., 1988; Cavanaugh et. al., 2002]

[Freeman & Simoncelli, 2011]



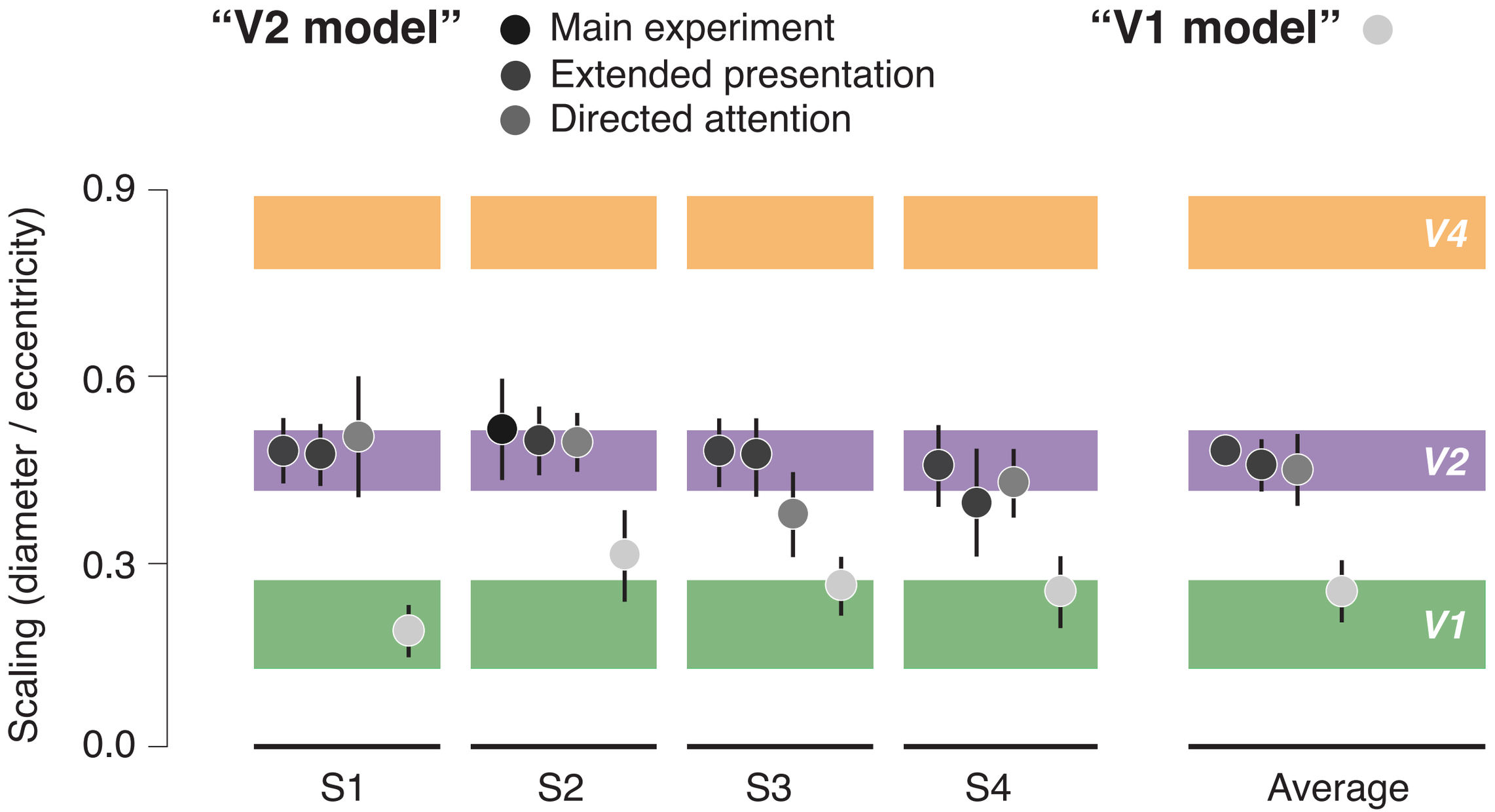
Scaling (diameter / eccentricity) of receptive fields in synthesis model

[Freeman & Simoncelli, 2011]



Scaling (diameter / eccentricity) of receptive fields in synthesis model

[Freeman & Simoncelli, 2011]

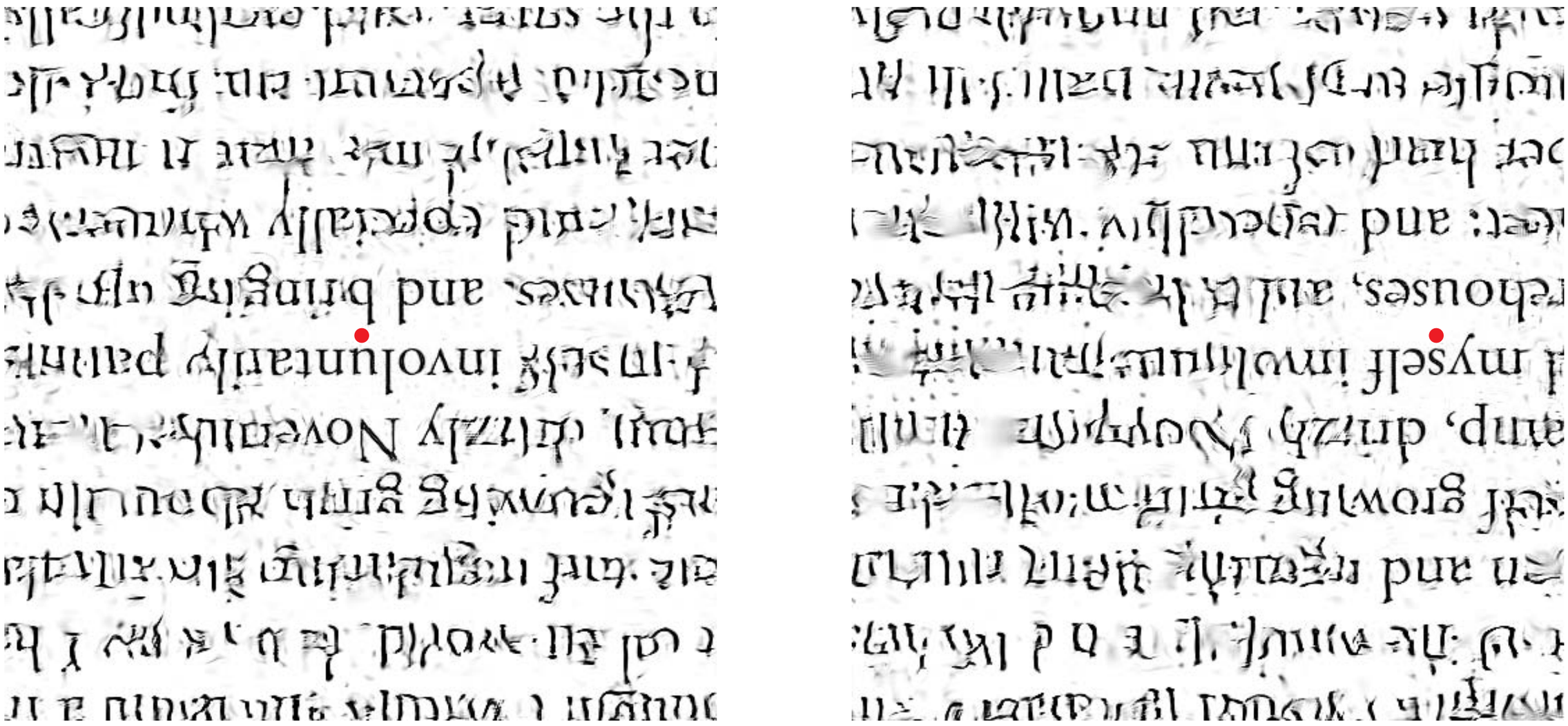


Reprinted by permission from Macmillan Publishers Ltd: Nature.
Source: Freeman, Jeremy, and Eero P. Simoncelli. "Metamers of the ventral stream." Nature neuroscience 14, no. 9 (2011): 1195 -1201. © 2011.

Macaque
Physiology

[Allman & Kaas, 1971; Allman & Kaas, 1974; Gattass et.al., 1981; van Essen et.al., 1984; Maguire & Baizer, 1984; Burkhalter & van Essen, 1986; Gattass et.al., 1987; Desimone & Schein, 1987; Gattass et.al., 1988; Cavanaugh et. al., 2002]

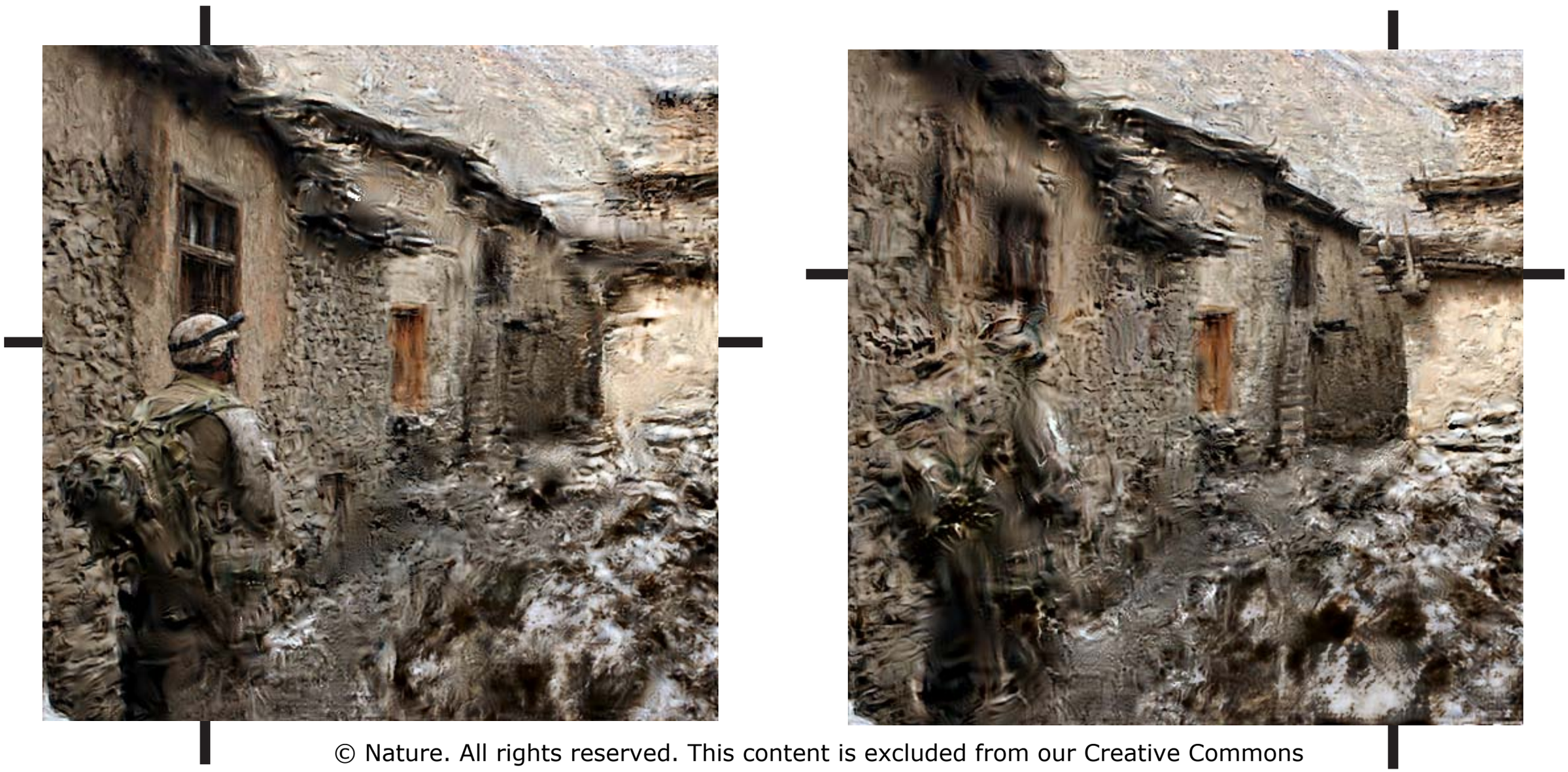
[Freeman & Simoncelli, 2011]



Reading

[Freeman & Simoncelli, 2011]

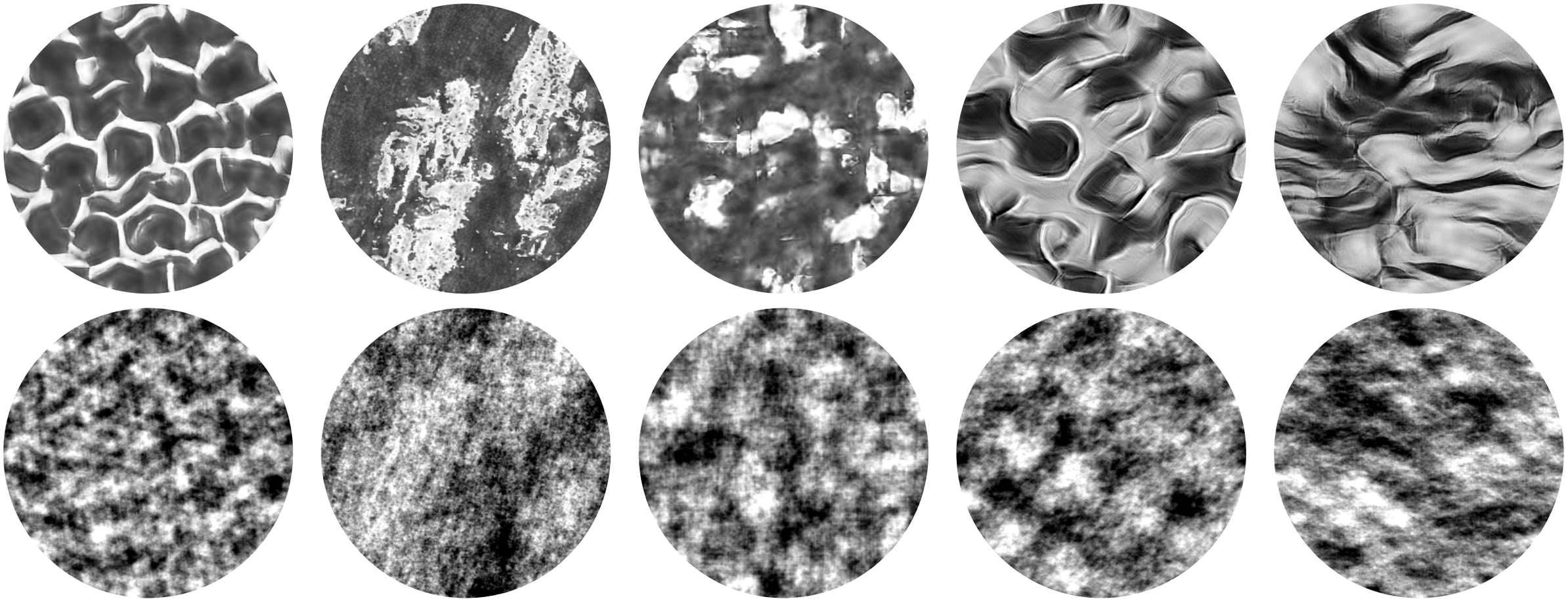
Camouflage



Source: Freeman, Jeremy, and Eero P. Simoncelli. "Metamers of the ventral stream." Nature neuroscience 14, no. 9 (2011): 1195-1201. © 2011.

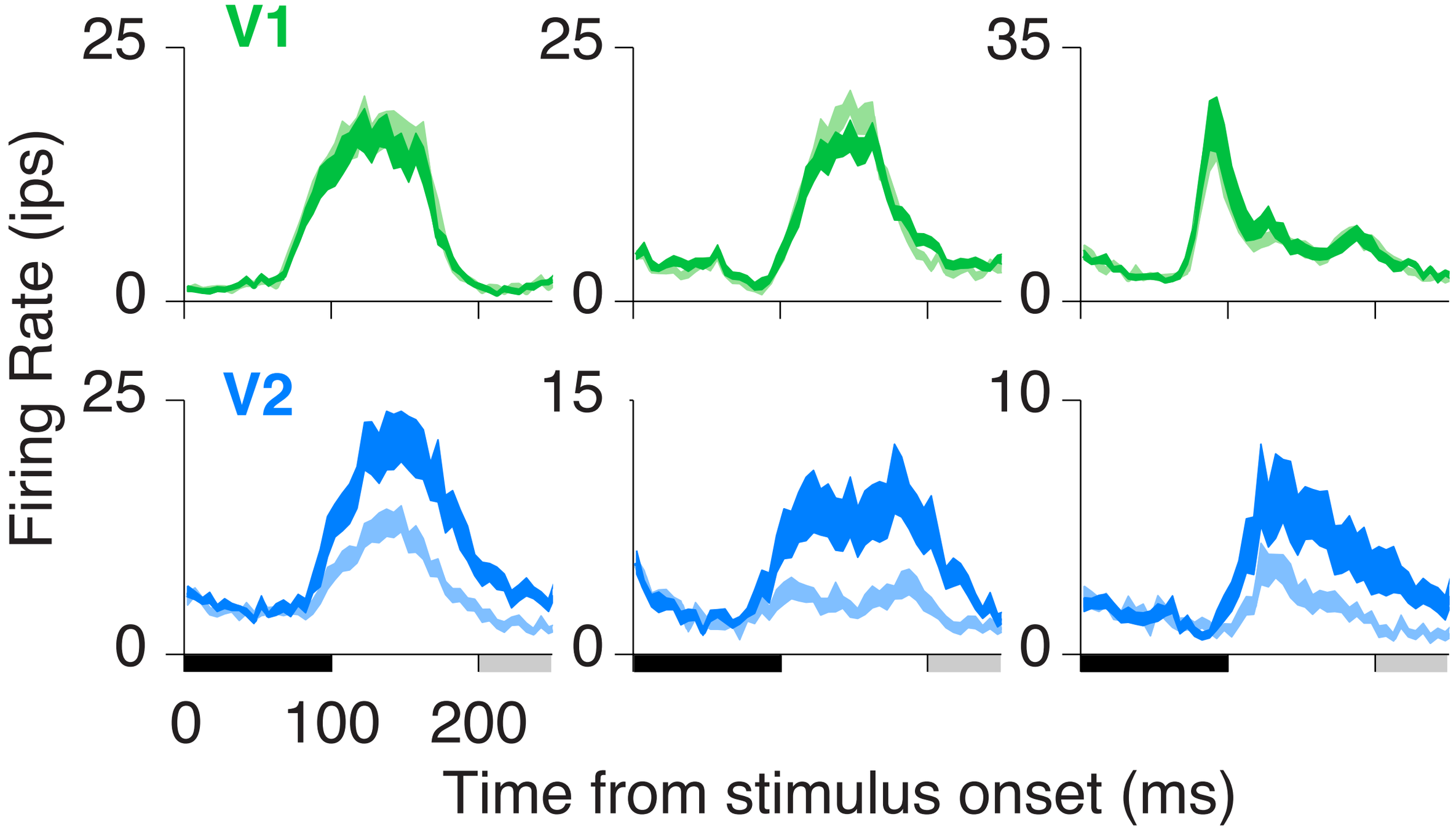
[Freeman & Simoncelli, 2011]

Can we drive individual V2 neurons using local texture stimuli?



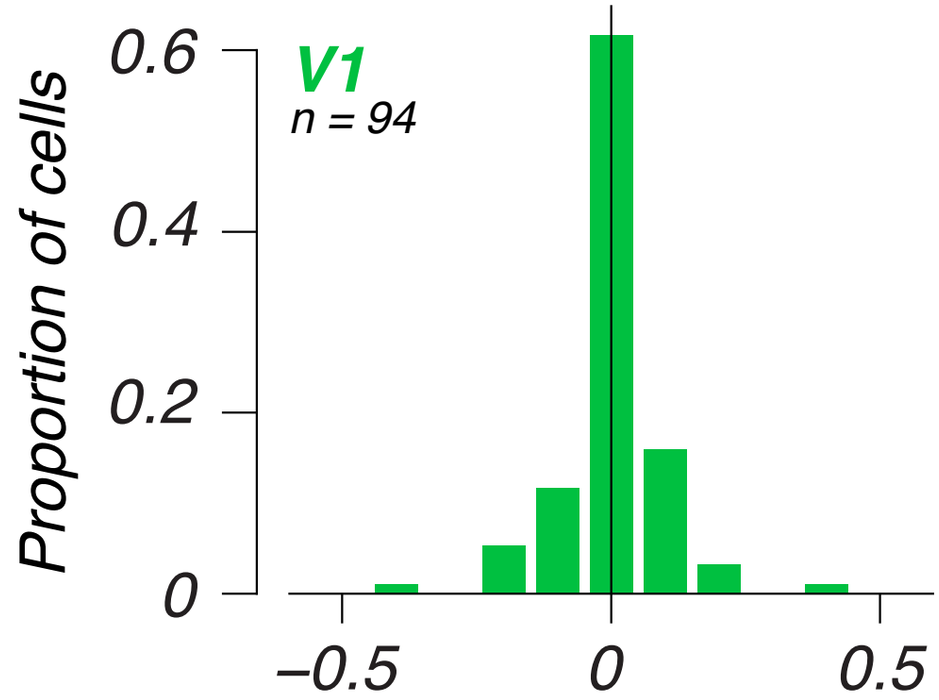
Top: synthetic textures, full model
Bottom: “spectral noise” (matched only for “V1” statistics)

[Freeman, et. al. 2013]

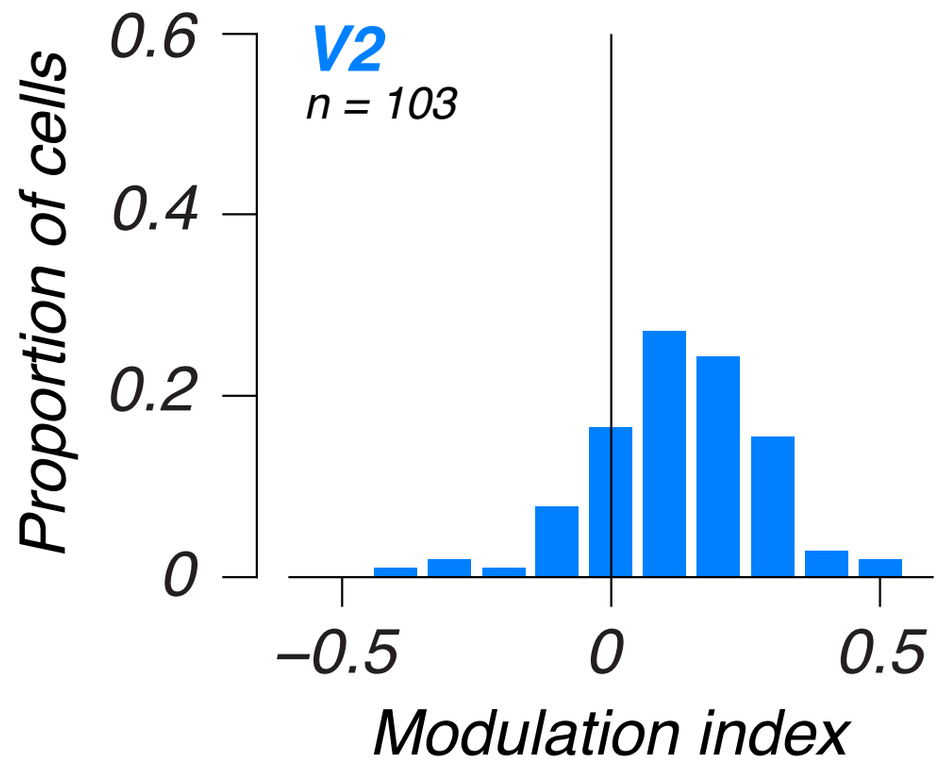


[Freeman, et. al. 2013]

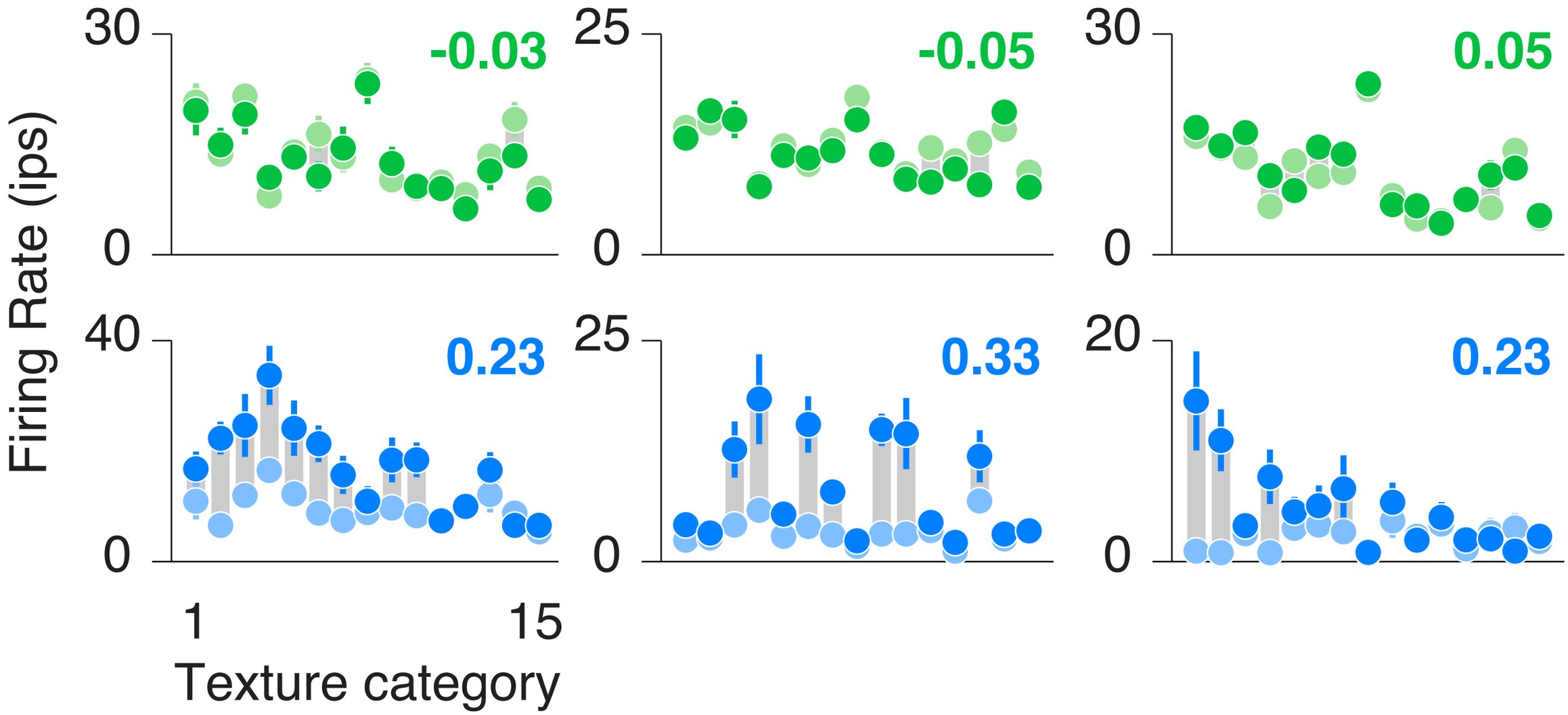
15% of V1 neurons significantly positively modulated



63% of V2 neurons significantly positively modulated

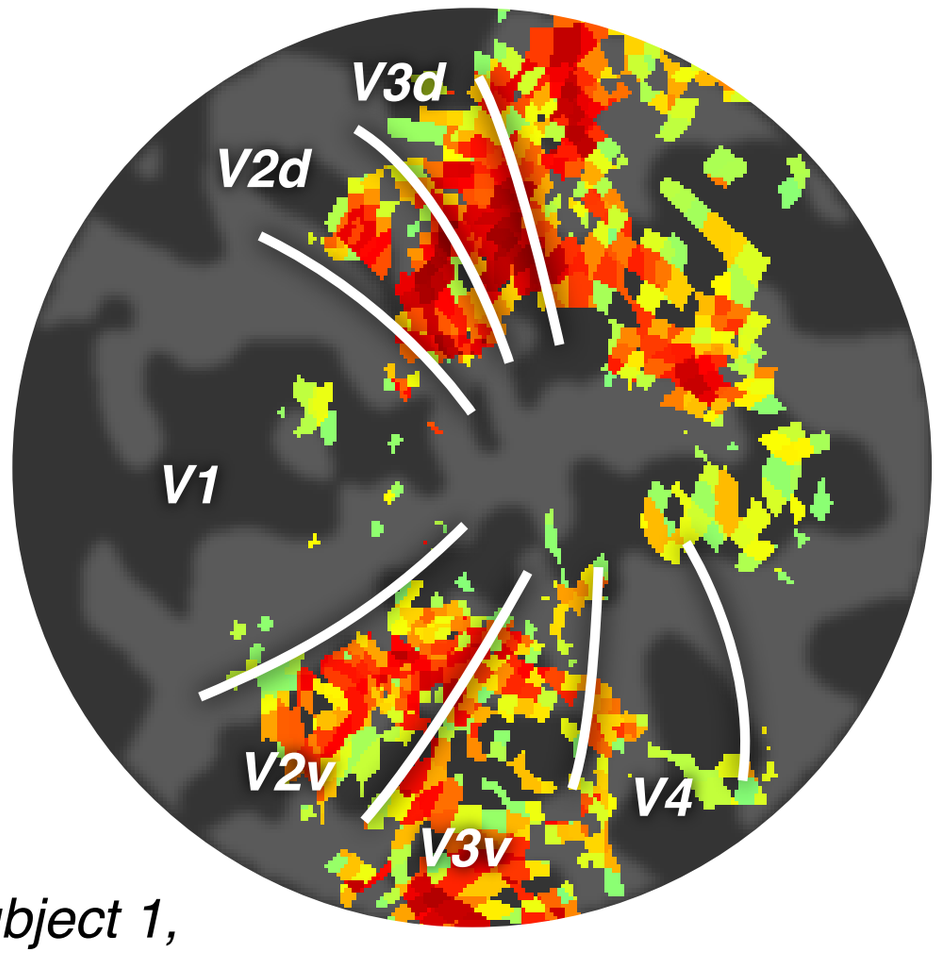


[Freeman, et. al. 2013]

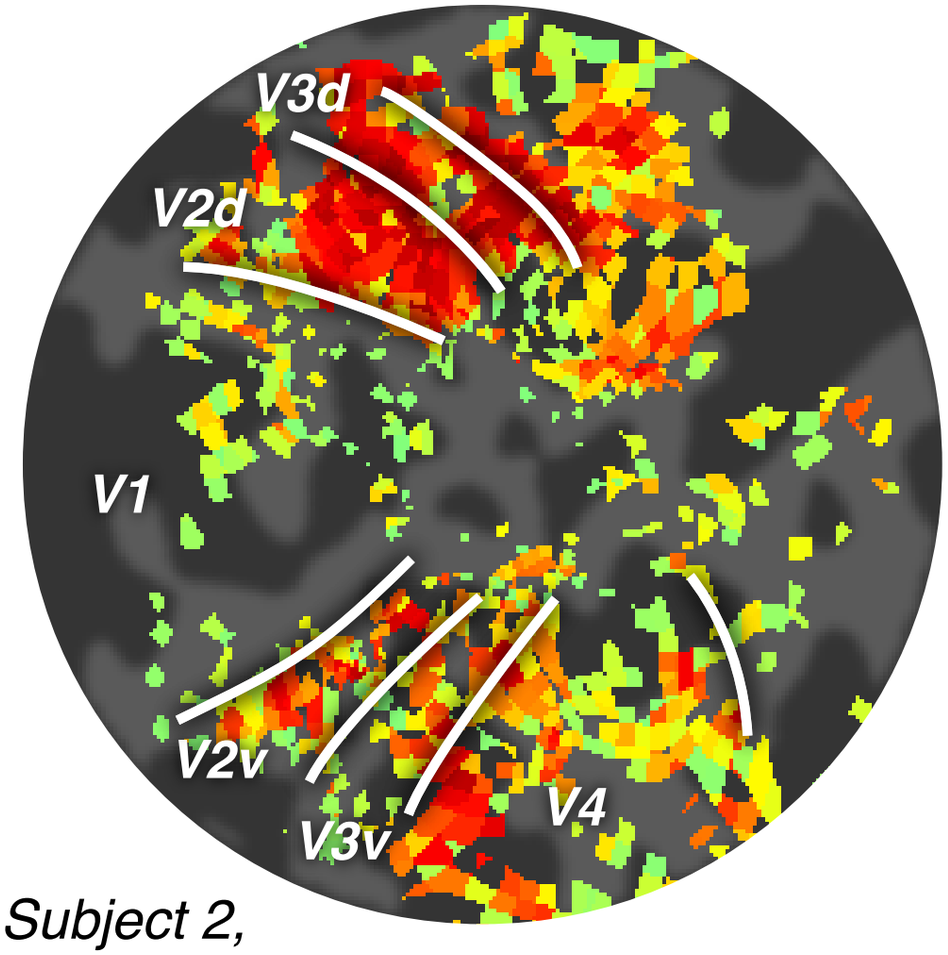


Reprinted by permission from Macmillan Publishers Ltd: Nature.
Source: Freeman, Jeremy, Corey M. Ziemba, David J. Heeger, Eero P. Simoncelli, and J. Anthony Movshon. "A functional and perceptual signature of the second visual area in primates." Nature neuroscience 16, no. 7 (2013): 974-981. © 2013.

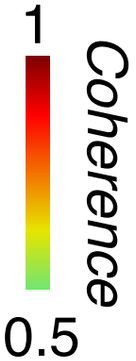
[Freeman, et. al. 2013]



Subject 1,
Right hemisphere

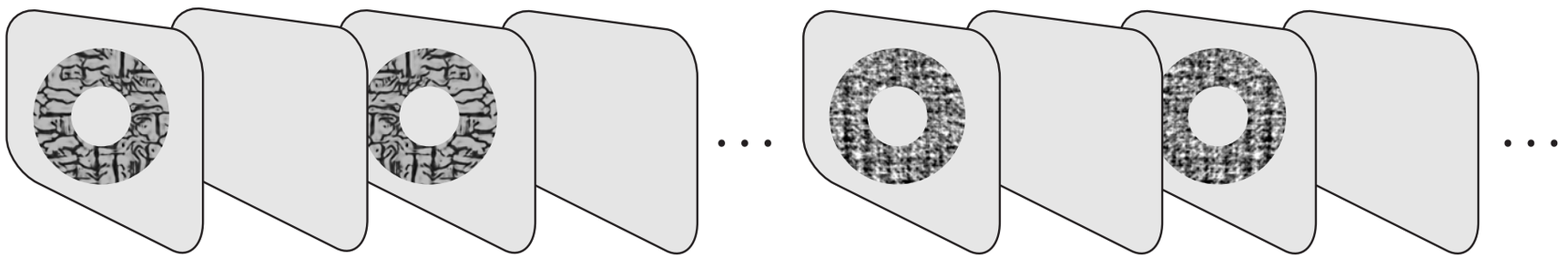


Subject 2,
Right hemisphere



Texture (9 sec)

Spectral Noise (9 sec)



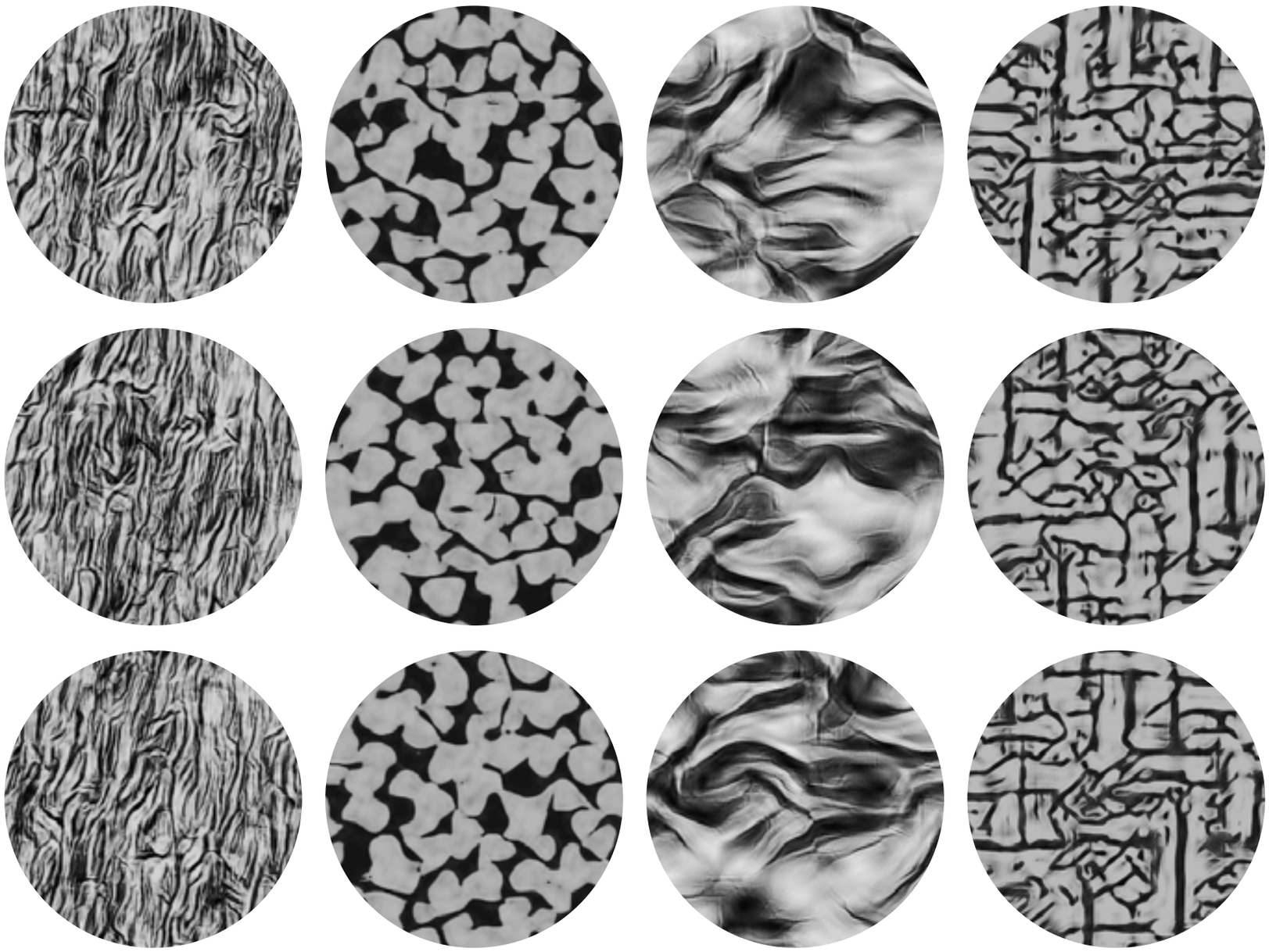
Time

[Freeman, et. al. 2013]

Predicting discriminability

Different families

Different exemplars
(same statistics)



Courtesy of Proceedings of the National Academy of Science. Used with permission.
Source: Ziemba, Corey M., Jeremy Freeman, J. Anthony Movshon, and Eero P. Simoncelli. "Selectivity and tolerance for visual texture in macaque V2." Proceedings of the National Academy of Sciences 113, no. 22 (2016): E3140-E3149.

[Ziemba, Freeman, Movshon, Simoncelli - unpublishd]

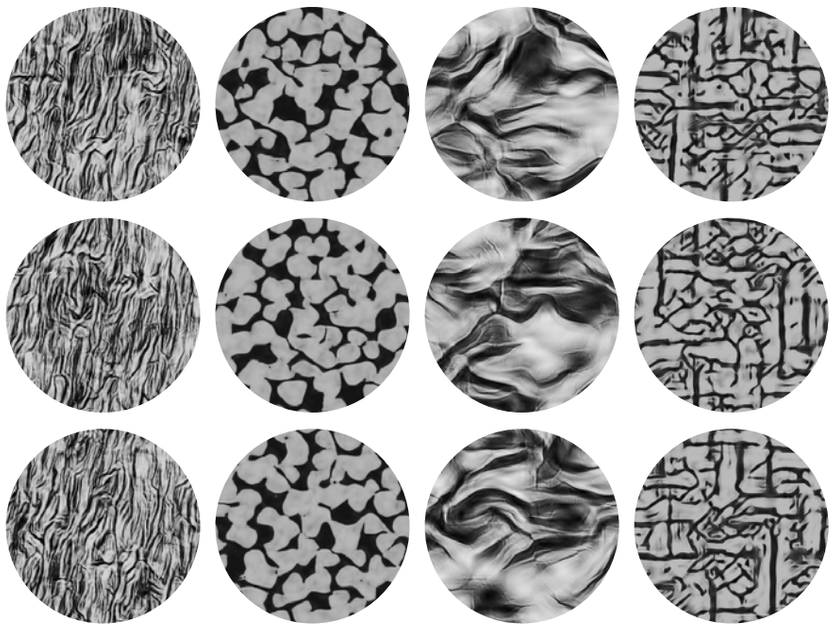
Example V1 neuron

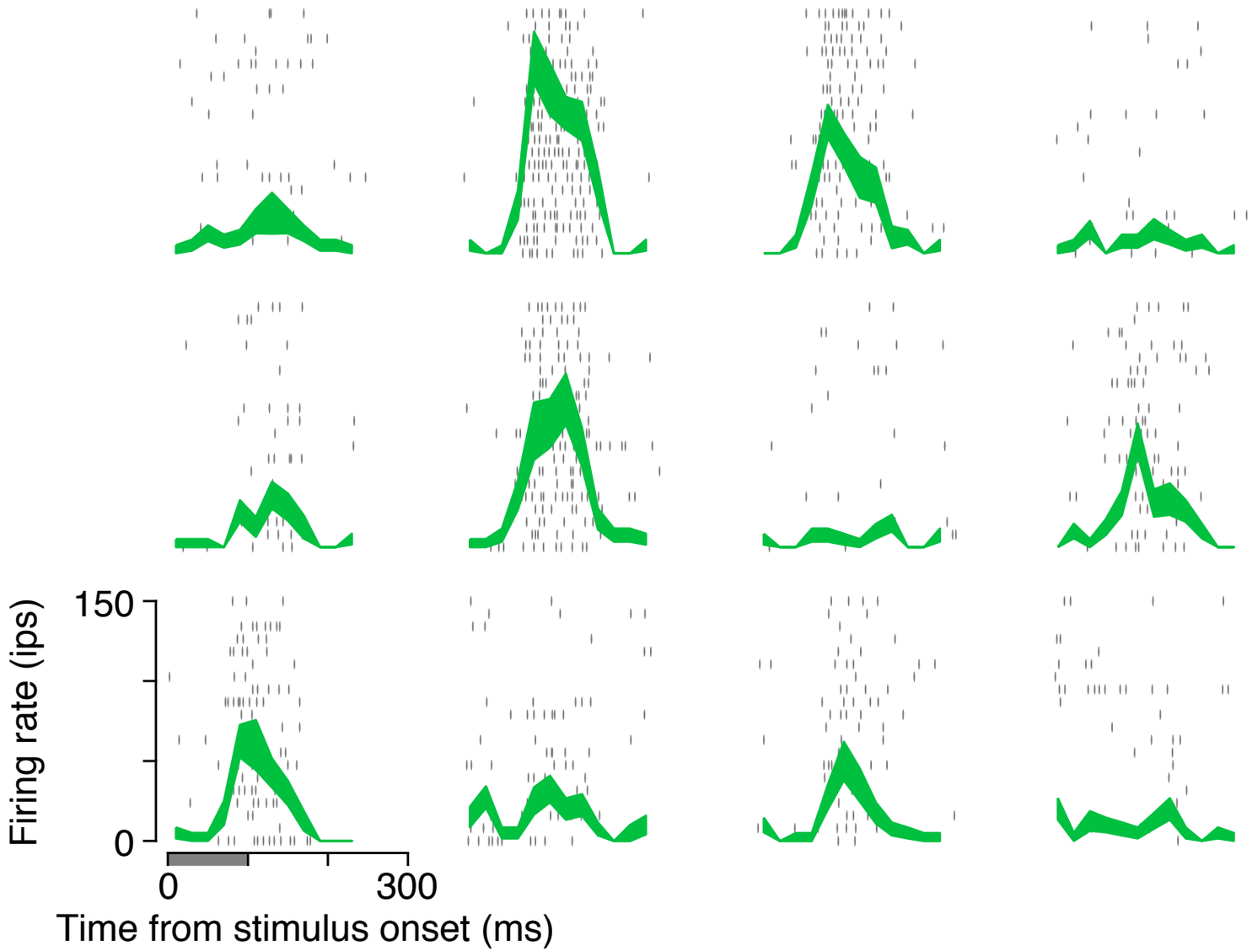
Anesthetized macaque

- V1: 102 neurons
- V2: 103 neurons

Stimuli presented for 100ms within a 4° aperture

20 repetitions each





Courtesy of Proceedings of the National Academy of Science. Used with permission.
Source: Ziemba, Corey M., Jeremy Freeman, J. Anthony Movshon, and Eero P. Simoncelli. "Selectivity and tolerance for visual texture in macaque V2." Proceedings of the National Academy of Sciences 113, no. 22 (2016): E3140-E3149.

[Ziemba, Freeman, Movshon, Simoncelli - unpublishd]

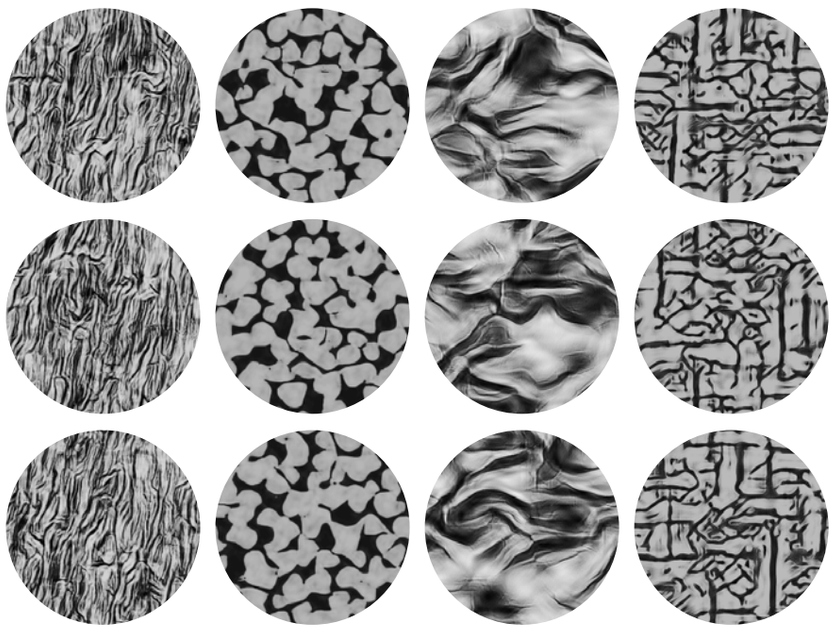
Example V2 neuron

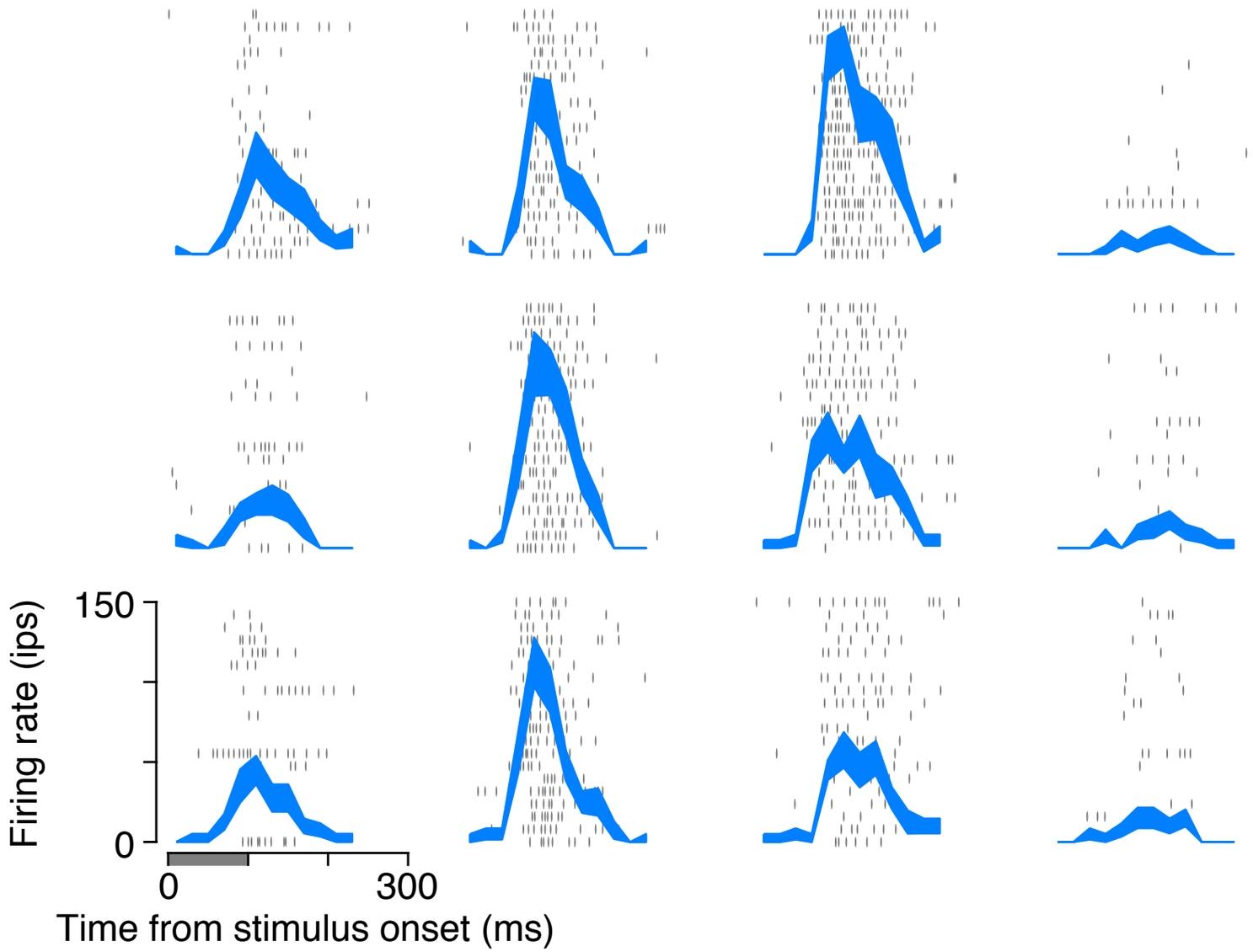
Anesthetized macaque

- V1: 102 neurons
- V2: 103 neurons

Stimuli presented for 100ms within a 4° aperture

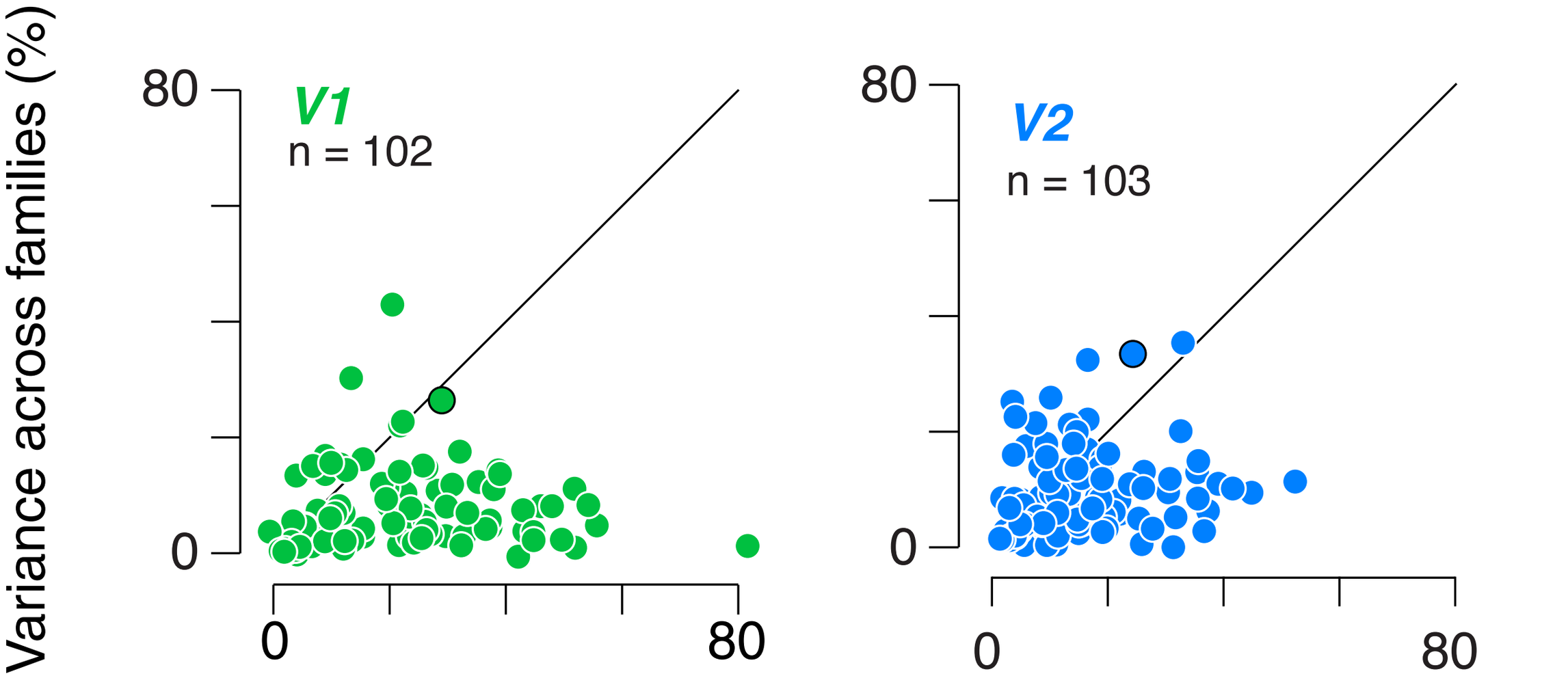
20 repetitions each





Source: Ziemba, Corey M., Jeremy Freeman, J. Anthony Movshon, and Eero P. Simoncelli. "Selectivity and tolerance for visual texture in macaque V2." Proceedings of the National Academy of Sciences 113, no. 22 (2016): E3140-E3149.

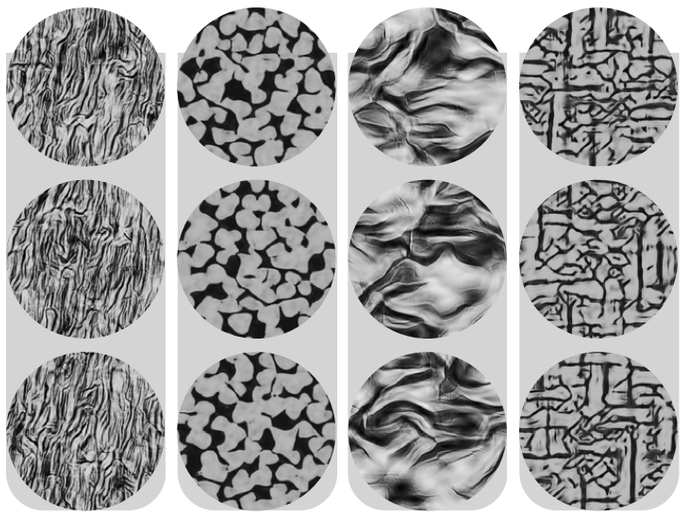
[Ziemba, Freeman, Movshon, Simoncelli - unpublishd]



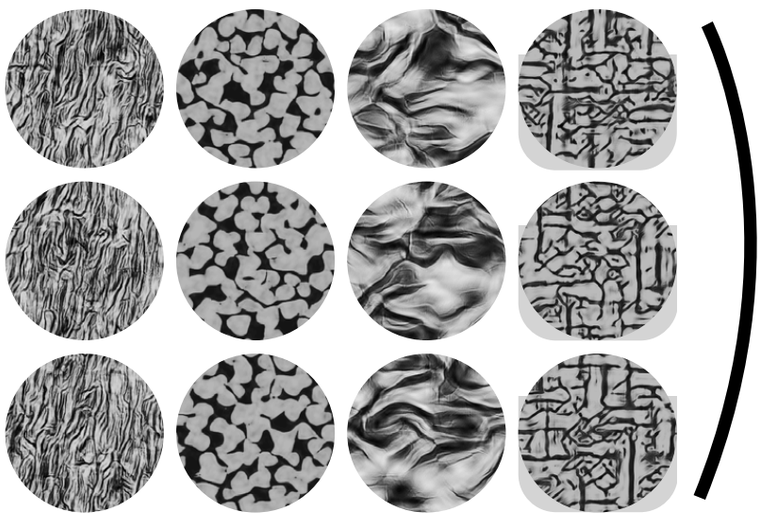
Courtesy of Proceedings of the National Academy of Science. Used with permission.
Source: Ziemba, Corey M., Jeremy Freeman, J. Anthony Movshon, and Eero P. Simoncelli. "Selectivity and tolerance for visual texture in macaque V2." Proceedings of the National Academy of Sciences 113, no. 22 (2016): E3140-E3149.

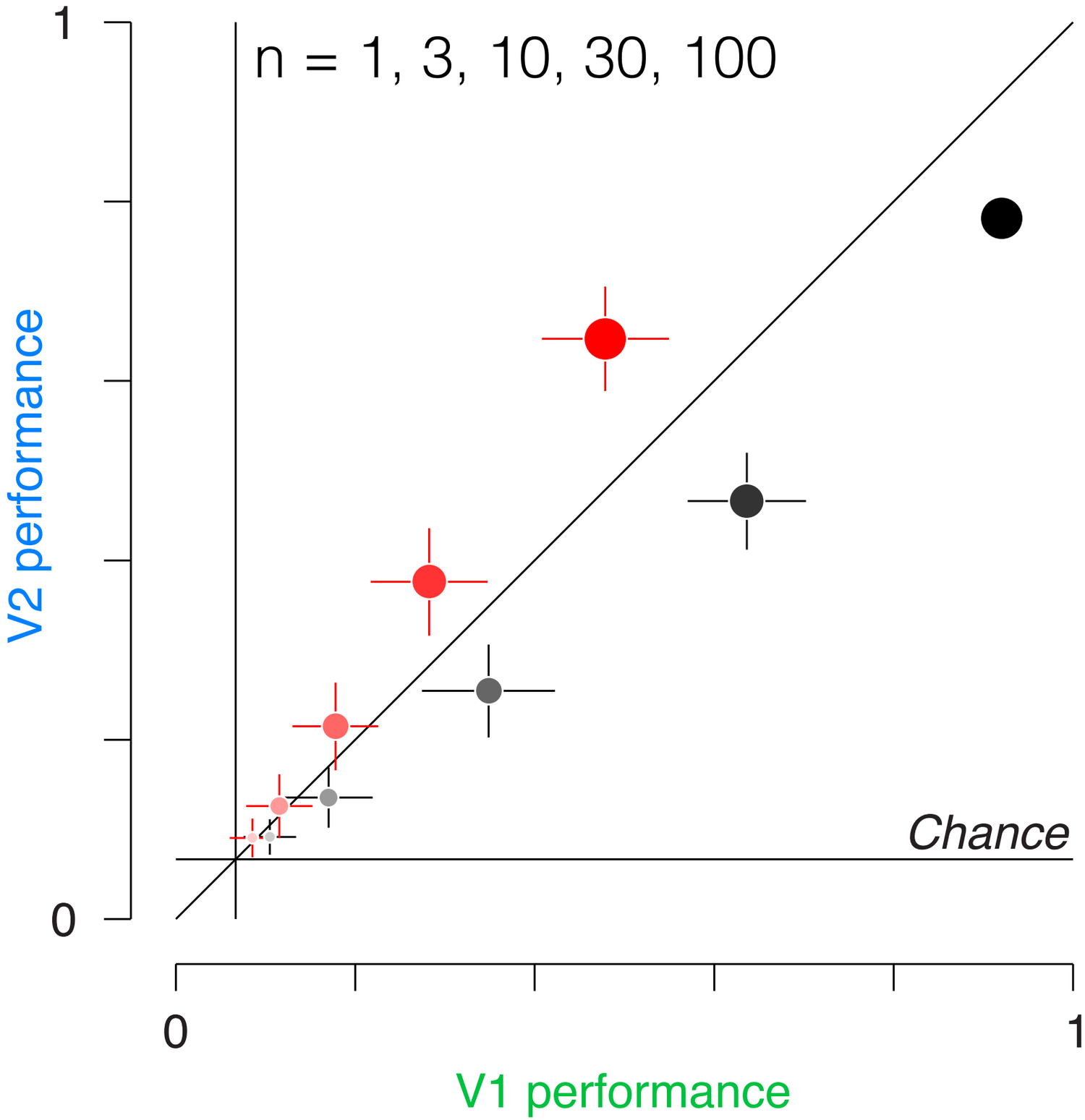
Decoding

Family classification



Exemplar identification





Courtesy of Proceedings of the National Academy of Science. Used with permission.
Source: Ziemba, Corey M., Jeremy Freeman, J. Anthony Movshon, and Eero P. Simoncelli. "Selectivity and tolerance for visual texture in macaque V2." Proceedings of the National Academy of Sciences 113, no. 22 (2016): E3140-E3149.

Portraits of Javier Portilla, Jeremy Freeman, Josh McDermott, Corey Ziemba and Tony Movshon removed due to copyright restrictions. Please see the video.

Resource: Brains, Minds and Machines Summer Course

Tomaso Poggio and Gabriel Kreiman