# Visual Perception

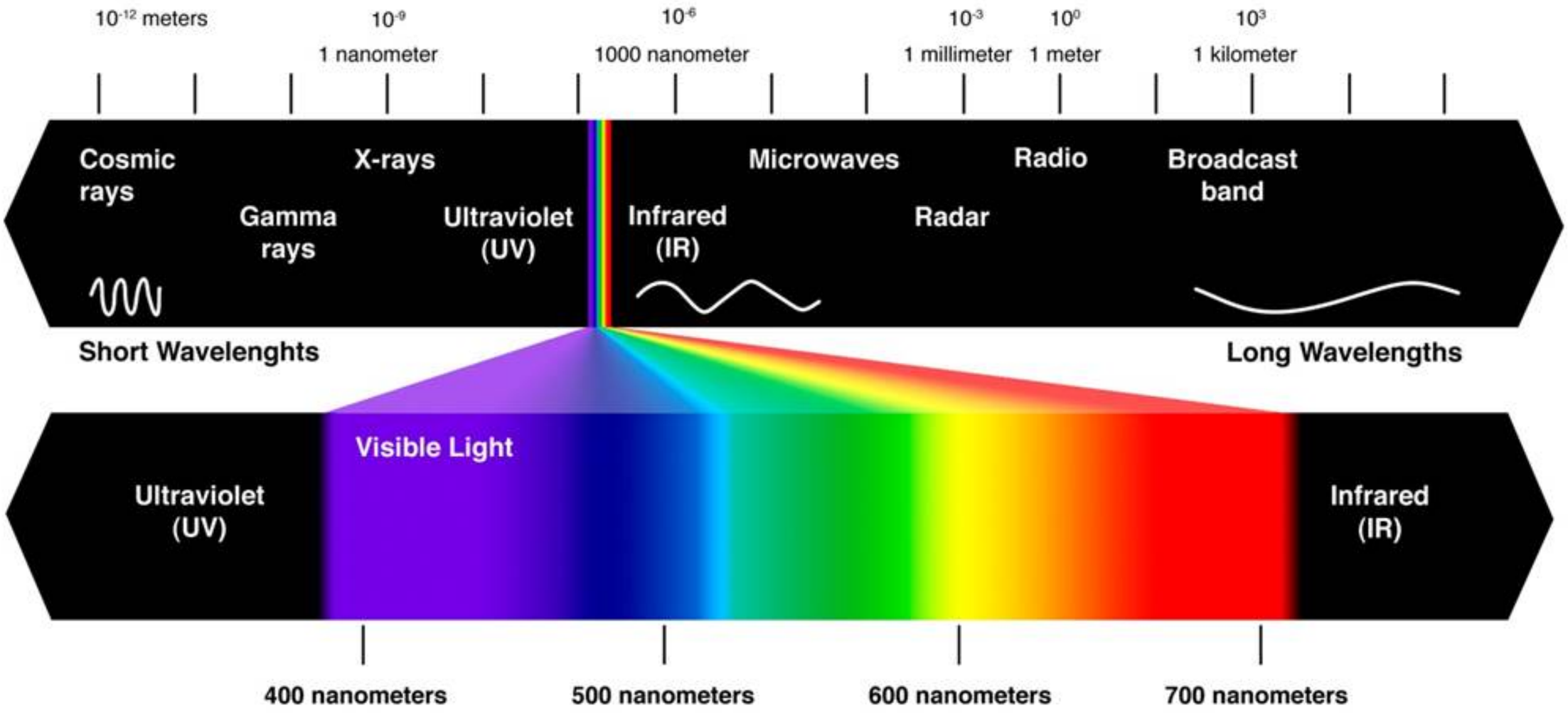
## COMP0160 Perception & Interfaces

**Matthew Caldwell, UCL, 17 January 2023**

# 2.1 Preamble

# Vision

- How we obtain information about the external world using **light**
  - In particular, **visible** light: EM radiation roughly 400–700 nm


$10^{-12}$ meters
$10^{-9}$
1 nanometer
$10^{-6}$
1000 nanometer
$10^{-3}$
1 millimeter
$10^{0}$
1 meter
$10^{3}$
1 kilometer
Cosmic rays
Gamma rays
X-rays
Ultraviolet (UV)
Infrared (IR)
Microwaves
Radar
Radio
Broadcast band
Short Wavelenghts
Long Wavelengths
Ultraviolet (UV)
Visible Light
Infrared (IR)
400 nanometers
500 nanometers
600 nanometers
700 nanometers

# Light

- Light is a **great** medium for gathering information
  - Propagates fast, over long distances, even through vacuum
  - Travels in straight lines
  - Interacts with matter, participates in chemistry — biology can detect
  - Plentiful, at least during the day, thanks to the sun

# Vision is useful

- So useful it has evolved many times
- Implementation details vary widely, eg with different eye configurations

# Vision is useful

- So useful it has evolved many times
- Implementation details vary widely, eg with different eye configurations

- But some of the underlying biochemistry is relatively well conserved

# Vision is (too) useful

- So useful it is a bit overpowering
- Dominant sense in most humans with a lot of resources dedicated to it
- Radically shapes our understanding of the world
  - Both literally and metaphorically
- Vocabulary of “vision” is pervasive
  - encompassing clarity, comprehension, ambition, **divine revelation**

# Vision is culturally important

- A lot of human culture revolves around the visual
  - Painting, drawing, photography, movies, video games
  - Practical necessities come wrapped in visual aesthetics
  - Written word is primarily mediated through our eyes

# Vision is technologically important

- We build lots of technology to enhance vision
  - Microscopes, telescopes, cameras
- And to produce visual stimuli
  - Paper and pencils, printing and screens, HUDs and holograms and VR
- Tends to be one of the main ways we interact with computers and software

# Vision is unreliable

- Complex and fragile and lots of things can go wrong
- Impairments are very common
  - Focal range, resolution
  - Lighting conditions, colour
  - Obstruction, tissue damage, neural pathways
- Some problems can be mitigated with technology like glasses, but not all

# So…

- Given centrality to human endeavours and frequent failings…
- Important to understand how it works, what it does, strengths & weaknesses
- So we can make the most of it while also maximising accessibility

# 2.2 Optics & Images

# Light

- Light exhibits characteristics of both particles and waves
- Some properties are better thought of in terms of one or the other, eg:
  - **brightness** is a function of number of particles — photons
  - **colour** depends on wavelength/frequency
  - for given colour, greater brightness means **more** photons, not brighter ones
- NB: woefully misusing terminology here — brightness and colour are really **perceptual** properties, not actually attributes of light at all!

# Energy

- Energy depends on wavelength — or reciprocally, frequency
  - Short wavelengths have high frequency and high energy
  - Longer wavelengths have lower frequency and lower energy
- High energies can be damaging to biological tissues
- Long wavelengths can be harder to detect
- Visible light is in a sort of Goldilocks zone of harmless detectability for us
  - Range is physiological happenstance, not physical, differs for other animals

# Refraction

- Light travels in straight lines at **constant speed** — in a vacuum
- Matter complicates things
- Can think of speed changing according to medium, leading to path changes
- Eg: heat haze — density fluctuations warp vision
- In particular, light rays bend at distinct boundaries, eg from air to glass

# Localisation

- To get info about the world, it helps to know where light is coming from
  - Not absolutely essential, may still be some value in sensing lightness and darkness without localisation — some animals do this
- If light from multiple locations hits a detector, the info about each all piles up together in the detector's response in an inseparable **blur**
- Ideally we would like to spatially organise the light so that all light coming from one location in the world arrives at the same place in our detector, and light from other locations goes to other places

# Pinholes

- One solution is to force light to pass through a single point to reach detector
  - a **pinhole aperture**
- Light arriving at any detector position must come from direction of pinhole
- Nice and simple, and does occur in nature, especially in simpler organisms
- But discards a lot of potentially informative light
  - almost all paths are blocked

# Lenses

- Alternative is to use refraction at the curved surface boundary of a lens
- Light travelling by different paths from same point is differentially bent
- All these rays are brought back together — into **focus** — at the detector
- Lens systems are harder to build than pinholes, both for human engineers and for biology
- But they capture more light and more information

# Image formation

- Detector at a focused point has good information about a single location
- With many detectors at many points we can form a **spatial map**
  - Each receives information about a different point in the world
  - Collectively they gather optical information over a whole **visual field**
- In other words: an **image**

# In and out of focus

- Focus is imperfect and not the same everywhere
- Lenses will only bring to focus some parts of the world
  - Some range of distances from the lens
- Detectors at well focused points will get strongly localised information
- Detectors elsewhere will receive light from a wider spread of locations, and the image will be correspondingly more imprecise, more **blurry**

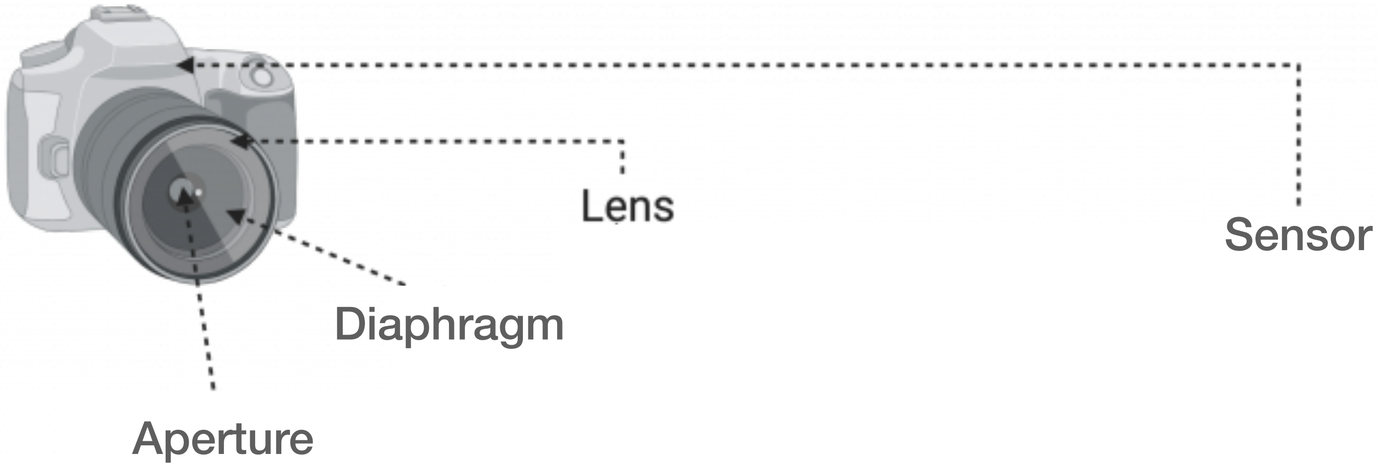
- We’re used to this kind of locally-focused imaging from using **cameras**
- Human vision is different from photography in many ways
- But cameras can still be a useful analogy when thinking about **eyes**

# 2.3 Eyes

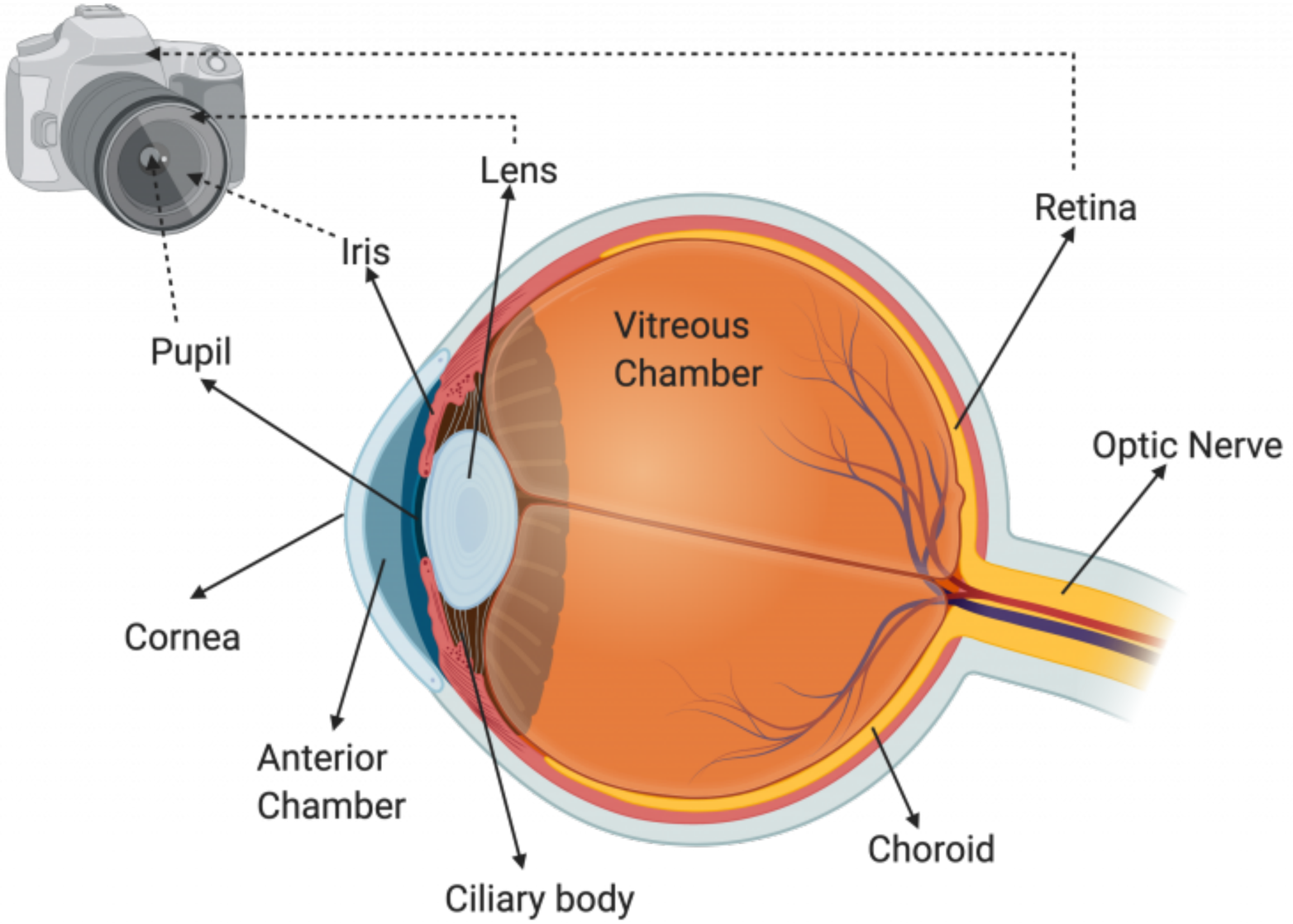
# Camera structure

- An **aperture** to control amount of light entering
  - Usually adjustable by means of a **diaphragm**
- A **lens** or system of lenses to capture and focus light
- Some form of chemical or electronic **detector**
- All wrapped in a dark box or chamber
  - from which the term **camera** derives

# The 'camera' eye



# The 'camera' eye



# Eye-camera equivalences

- The eye’s aperture is the **pupil**
  - Adjusted by the dilatory and sphincter muscles of the **iris**
- There are two focusing elements:
  - The rigid, fixed focal length **cornea** does most of the work
  - The deformable **lens** provides adjustability

# Accommodation

- Ring of ciliary muscles around the lens contract to increase curvature
- Decreases focal length
- Allows nearer objects to be in focus
- This adjustment is known as **accommodation**

# Focal range

- Maximum distance in focus with least accommodation — when ciliary muscles are fully relaxed — is the **far point**
- Minimum distance that is in focus with most accommodation — ciliary muscles fully contracted — is the **near point**
- These bound the range that can possibly be brought to focus
- Range depends on shape of the eye, focusing power of cornea and lens, deformability of the lens, strength of ciliary muscles

# Out of focus

- Focal range impairments are very common, including:
  - **Myopia** — nearsightedness — far point much closer than infinity, distant objects focus in front of retina
  - **Hyperopia** — farsightedness — near point far away, close objects focus behind retina
  - **Presbyopia** — “old eye” — age-related farsightedness due to stiffening of lens and weakening of ciliary muscles
- Mitigable with glasses or contact lenses, shifting baseline focus of whole

# Eye movement & control

- Eyes are directional
- As well as muscles for focus & aperture, there are six (per eye) for movement
  - Acting in oppositional pairs to rotate around (somewhat) different axes
  - Movements are also coordinated across both eyes
- Several very distinct types of movement

# Saccades

- Fast skips or jumps between static fixation points. Eg:
  - Looking from face to face in conversation
  - Scanning along words in a line of text
  - Glancing between slide and speaker in a lecture like this
- A key feature: the jump is basically **invisible** (saccadic masking)
- There are also **microsaccades**, which are a bit different, but also invisible

# Smooth pursuit

- Continuous movements tracking a moving object

# Vergence

- Opposing angle changes of both eyes to point at the same thing
  - Rotating closer together for nearby objects
  - Rotating further apart for more distant ones

# Gaze stabilisation

- Compensate for changes in head position or view content
- Maintains field of view
- Key example: **vestibulo-ocular reflex** rotates and shifts eye position to oppose rotations and shifts of the head
  - Driven by positional information detected in the inner ear
  - So works (up to a point) for induced, involuntary movements

# Out of control

- Most of these movements are not under conscious control
- You can choose to **look at** some location, not point your eyes in arbitrary directions
  - …or rotate them around the viewing axis
  - …or wilfully contract your irises — let alone each one separately!
- Can consciously trace a path in a sequence of saccades, but not **smoothly pursue**
  - for that you need to watch **something** moving and let your brain do the work
- Vergence and focus are *somewhat* controllable, but mostly by imagining targets

# Key Grips

- Brain imposes a lot of mostly-unnoticed structure on processes of vision
- Lots of work going on behind the scenes to create your experience of seeing
- Visual system is like a slick Hollywood production team:
  - saccades chop up the world like a skilled editor
  - vestibulo-ocular reflex is your own personal steadicam operator
- And this is just **mechanics**. We haven't even got to transduction yet!

# 2.4 Retina

# At long last, transduction

- Retina: thin tissue layer at the back of the eye
- Responsible for transducing incoming light to electrochemical nerve signals
  - but also for several additional integration and processing steps
- Retina is **not** just a simple sensor
- Part of the CNS, not peripheral — an extension of the brain
  - Grows out of (what becomes) the brain during development

# Layer Cake

- Organised in functional layers containing different types of neuron
- **Inverted** in humans (and other vertebrates)
  - Processing proceeds from **back to front**
  - Detectors are at the back — light passes through all layers first
  - Transduced signals propagate back upward
  - Results at top layer have to go back through retina again to get to brain
    - …creating hole with no detectors — the **blind spot**

# Rods and cones

- Two distinct types of photoreceptor neurons in back layer
- Known as rods and cones after shape of **outer segment**
  - the end that’s furthest away, containing the photosensitive pigments
- Differ in sensitivity and spatial distribution and connectivity
  - we’ll come back to this shortly

# Polarisation

- Unusually, rods and cones are **depolarised** at rest
  - Inside of cell is less electrically negative relative to outside than normal
- Depolarisation drives neurons to release neurotransmitter
- So in the absence of light stimulus, photoreceptor neurons release a lot
- Incoming photons are absorbed by photosensitive protein complexes, causing the **retinal** component to dissociate from the **opsin** component
- Causes an amplifying signalling cascade that leads to **hyperpolarisation**

# Propagation

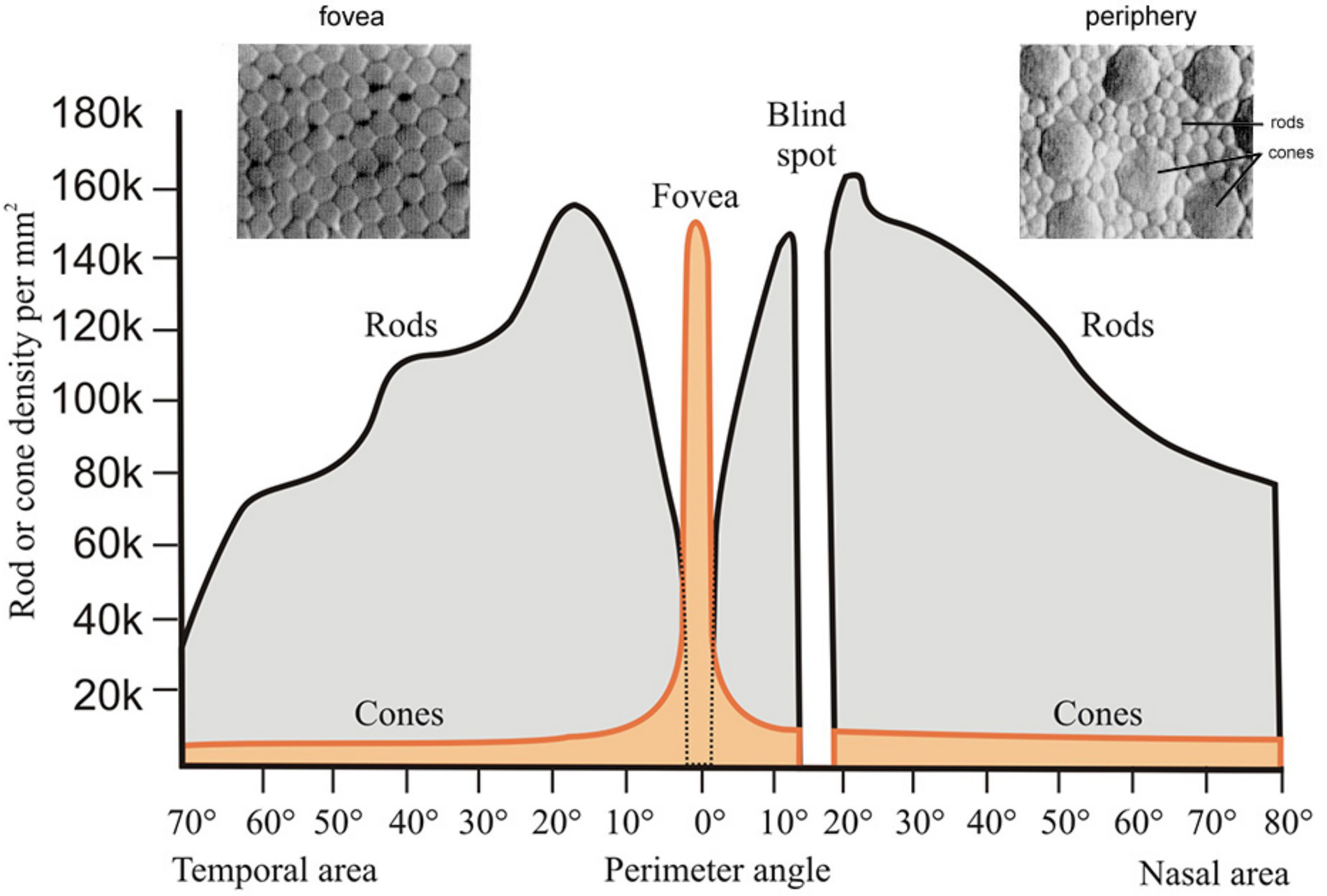
- Release changes in the rods and cones are received by the next layer
  - (of **bipolar** and **horizontal** cells)
- …and passed on to the uppermost layer
  - (of **retinal ganglion** and **amacrine** neurons)
- …which then sends the processed signals onward to the brain

# Local neurons for local people

- Horizontal and amacrine cells are **interneurons**
  - They mediate locally between neurons in the same layer or region, rather than projecting elsewhere
- Interneurons are primarily **inhibitory**
- Crucial for pretty much every kind of finely-tuned process in the brain

# Distribution

- Photoreceptor distribution over retina is extremely non-uniform
- Small central region — **fovea** — contains only small, densely-packed cones
- Outside fovea there's a mix of rods and cones, with rods dominating
  - ~10-20× more rods than cones overall
- Cones rapidly drop off
- All cells become larger and less dense towards periphery
- None at all in the blind spot


fovea
periphery
rods
cones
Blind
spot
Fovea
Rods
Rods
Cones
Cones
180k
160k
140k
120k
100k
80k
60k
40k
20k
Rod or cone density per mm²
70° 60° 50° 40° 30° 20° 10° 0° 10° 20° 30° 40° 50° 60° 70° 80°
Temporal area
Perimeter angle
Nasal area

# Convergence

- Rod and cone signals route to the retinal ganglion cells
- But there are about ~100× more rods+cones than RGCs
- So what gets sent to brain is not 1:1 reportage of raw detection
- Data are summarised and transformed, picking out interesting features, smoothing out noise, amplifying
- Many photoreceptor outputs may **converge** on a single RGC
- Convergence is low in the fovea, high in the periphery

# Convergence vs acuity

- Aggregating many inputs improves detection and SNR but sacrifices detail
- Cones in fovea have very low convergence and are small and tightly packed
- So visual acuity in this region is typically very high
- In periphery, hundreds or thousands of rods may converge on one RGC
- Density is also low, so aggregated acuity is low

# Sensitivity

- Rods are much more light sensitive than cones
- Cones respond well in bright (**photopic**) conditions, rods are overwhelmed
- In dim (**scotopic**) conditions fovea is pretty unresponsive
- Scotopic vision dominated by the high convergence low acuity periphery
- Some objects may only be visible **out of the corner of your eye**
- Cones also responsible for colour detection (see later), so night vision is largely monochromatic

# Desensitisation

- Photosensitive pigments that detect light are **bleached** in the process
- So light exposure leads to **desensitisation**
- Pigments are continually replenished, but the process takes time
- So sensitivity of retina depends on how much light has been seen lately

# Adaptation

- Sensitivity comes from interplay of rate of bleaching and rate of replenishment
- When moving from one lighting environment to another rates will change, and it will take time to reach equilibrium — this is known as **adaptation**
- Eg, moving from sunny outdoors to darkened room: sensitivity will have been depleted by brightness and you won't see much
- If you've been in the dark for a long time, pigment levels will be much higher due to reduced bleaching, and you'll have dark-adapted vision

# Receptive fields

- Retinal ganglion cells receive signals from multiple converging photoreceptors
- So RGC response depends on the whole patch of retina containing them
  - and hence the corresponding patch of the visual field
- This is known as the **receptive field**
  - The portion of the visual field to which the cell is receptive
- Inputs from the receptive field are aggregated in non-trivial ways

# Centre-surround antagonism

- Most RGCs have a centre-surround structure to their receptive fields
- Inputs from the centre of the field are treated differently from the surround
  - in particular, the two are opposed, exhibiting **lateral inhibition**
  - signal evoked by photoreceptors in the central region are inhibited by signals from its neighbours in the surrounding region
- RGCs may be **on-centre** (responding maximally to bright centre and dark surround in receptive field) or **off-centre** (dark centre and bright surround)

# Feature extraction

- Centre-surround antagonism picks out differences — **edges**
  - (or spots or corners or whatever — it is not orientation sensitive)
- It serves to extract features of interest from the image
- Driven by perceptual utility not physiological necessity
- Edges are potentially important boundary markers of things in the world
- Regions of uniform illumination are perceptually dull!

# Chorus

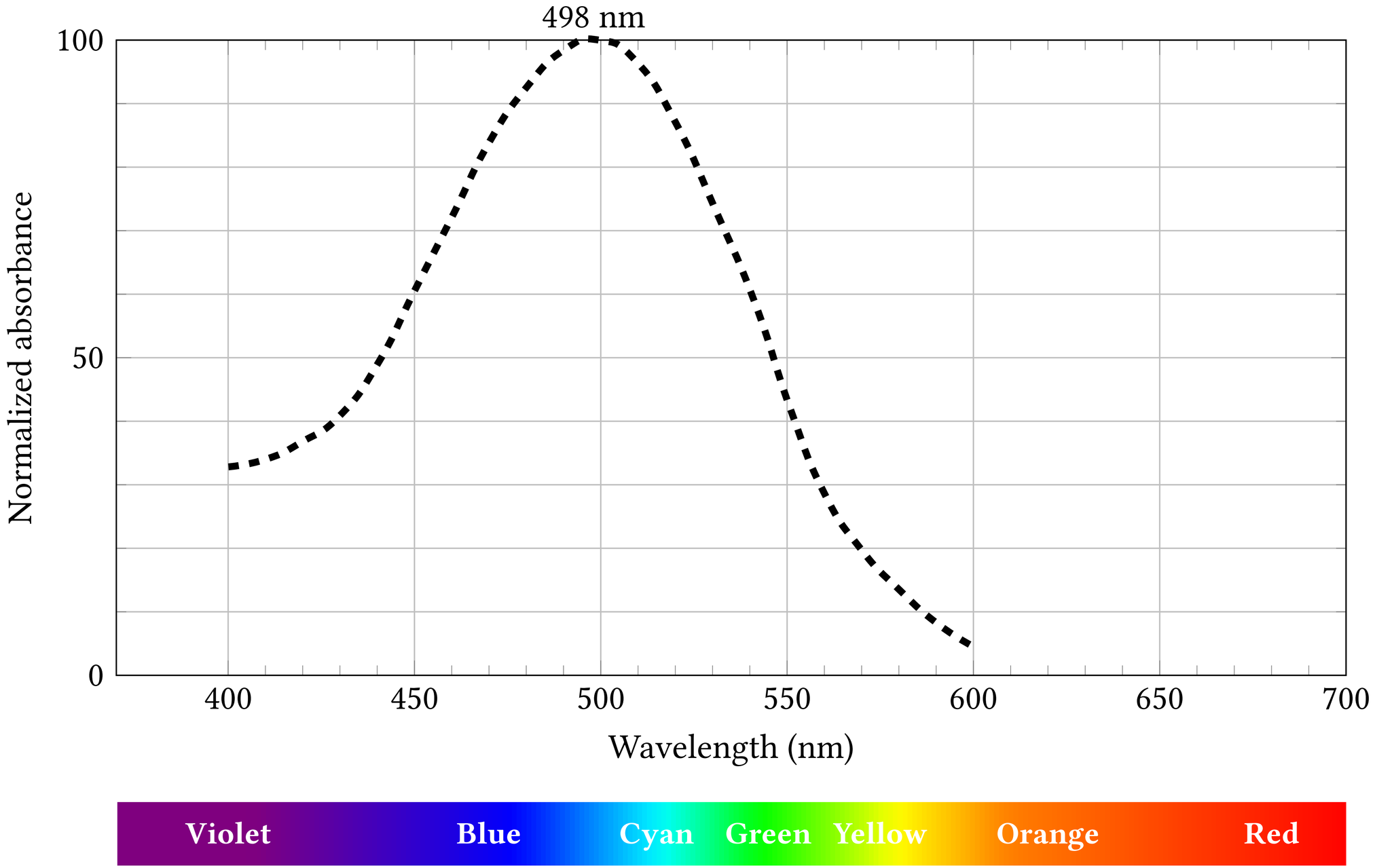
- Once again, we note that perception is an active, constructive process, not one of passive detection and consumption
- At every stage it makes choices, imposes structure, sifts and selects
- Even before we've left the eye we're already making value judgements

# 2.5 Colour

# Spectral sensitivity

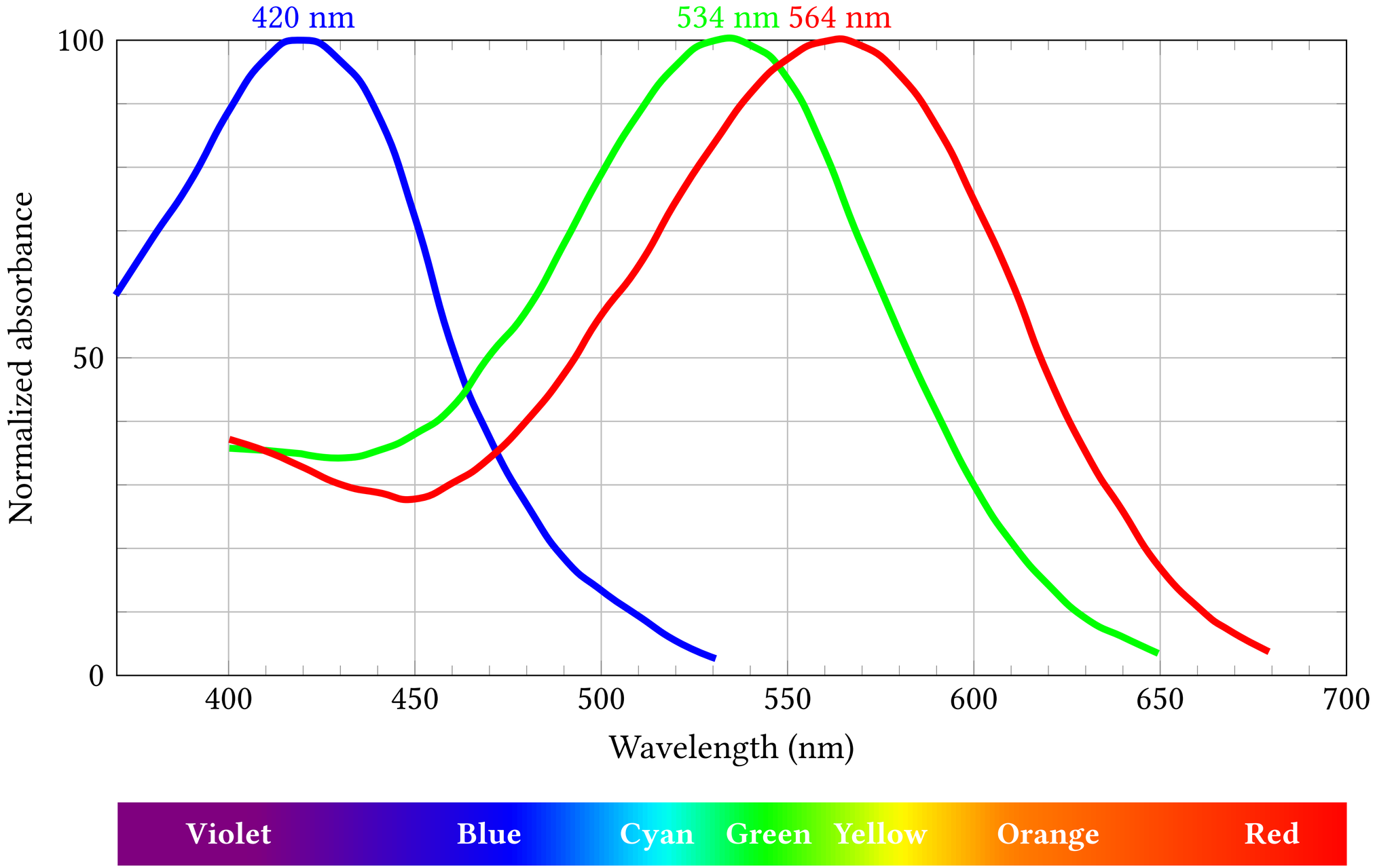
- The photopigments in rods and cones either absorb a photon and dissociate, or they don’t, there’s no middle ground
- They are not equally sensitive to all wavelengths, but the result is the same either way — the wavelength isn’t captured in the response
- The pattern of neurotransmitter release from a rod or cone tells you how strongly it is being stimulated, but not *why* the stimulation is strong:
  - Few photons of wavelength to which pigment is highly sensitive?
  - Or lots of a wavelength to which it is weakly sensitive?

# Rods are all the same


498 nm
100
50
0
Normalized absorbance
400
450
500
550
600
650
700
Wavelength (nm)
Violet
Blue
Cyan
Green
Yellow
Orange
Red

# Cones come in three flavours

- The absorption spectrum is different for each

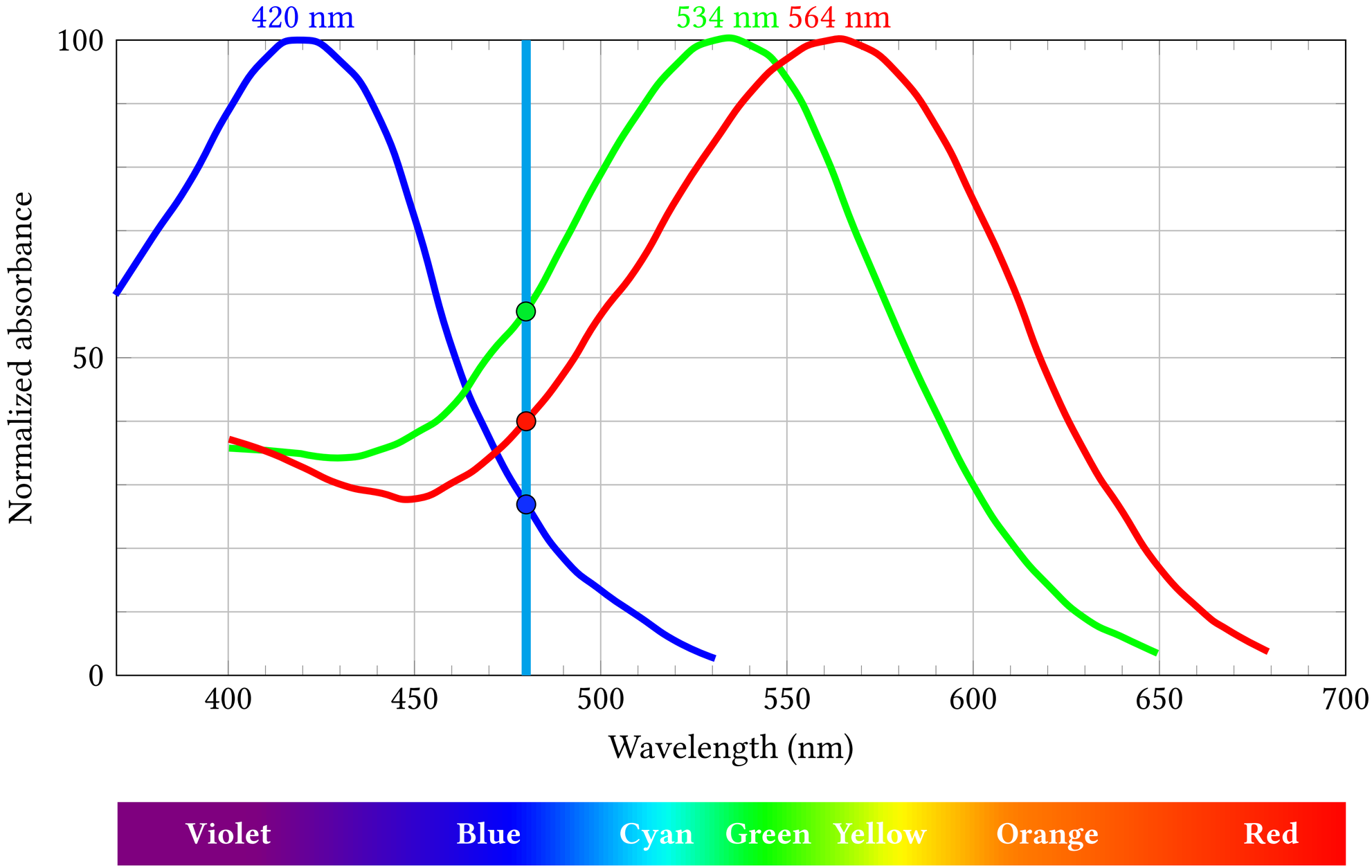

420 nm
534 nm 564 nm
100
50
0
Normalized absorbance
400
450
500
550
600
650
700
Wavelength (nm)
Violet
Blue
Cyan
Green
Yellow
Orange
Red

# Cones come in three flavours

- The absorption spectrum is different for each
- Sometimes termed red, green and blue…
  - (for evocative but **obviously wrong** reasons)
- …but more accurately short (S), medium (M) and long (L) wavelength
- The difference is the basis for our perception of **colour**

# Better living through spectroscopy

- If cones of all three types are illuminated by light of the same wavelength and intensity, their responses will differ
- By comparing the responses we can **disentangle** the effects of wavelength & intensity
  - Similar to the way a smart watch measures blood oxygenation
- And thus figure out what the wavelength actually is
  - At least, sorta kinda


420 nm
534 nm 564 nm
100
50
0
Normalized absorbance
400
450
500
550
600
650
700
Wavelength (nm)
Violet
Blue
Cyan
Green
Yellow
Orange
Red

# Inference

- Short wavelength S (or “blue”) cones are actually stimulated least
- Long wavelength L (“red”) are stimulated a bit more
- Medium M (“green”) most of all
- But this isn’t **winner takes all**
- Importantly, no other wavelength produces this ratio of responses
- So we can unambiguously identify the light as **blueish** 480 nm

# Except…

- Inference assumes single wavelength light
- This is never really true, but it’s not a terrible assumption…
  - …when broad spectrum white sunlight bounces off stuff that absorbs most wavelengths and leaves only a narrow band
  - **mostly** the case **most** of the time for **most** of the history of life on Earth
- So it’s a decent model for evolution to have hardwired into our physiology

# Nowadays…

- We spend a lot of time looking at light from artificial sources
- Model doesn’t hold at all
- But our visual system doesn’t know that, it’s still stuck on the veldt
- Good news for technology and culture!
- We can bamboozle our brain with cheap parlour tricks

# Cheap parlour tricks

- Take 3 sources that **mostly** stimulate the long, medium and short cones
- Shine them in the right proportions into your eyes
- Behold!

# Primary colours

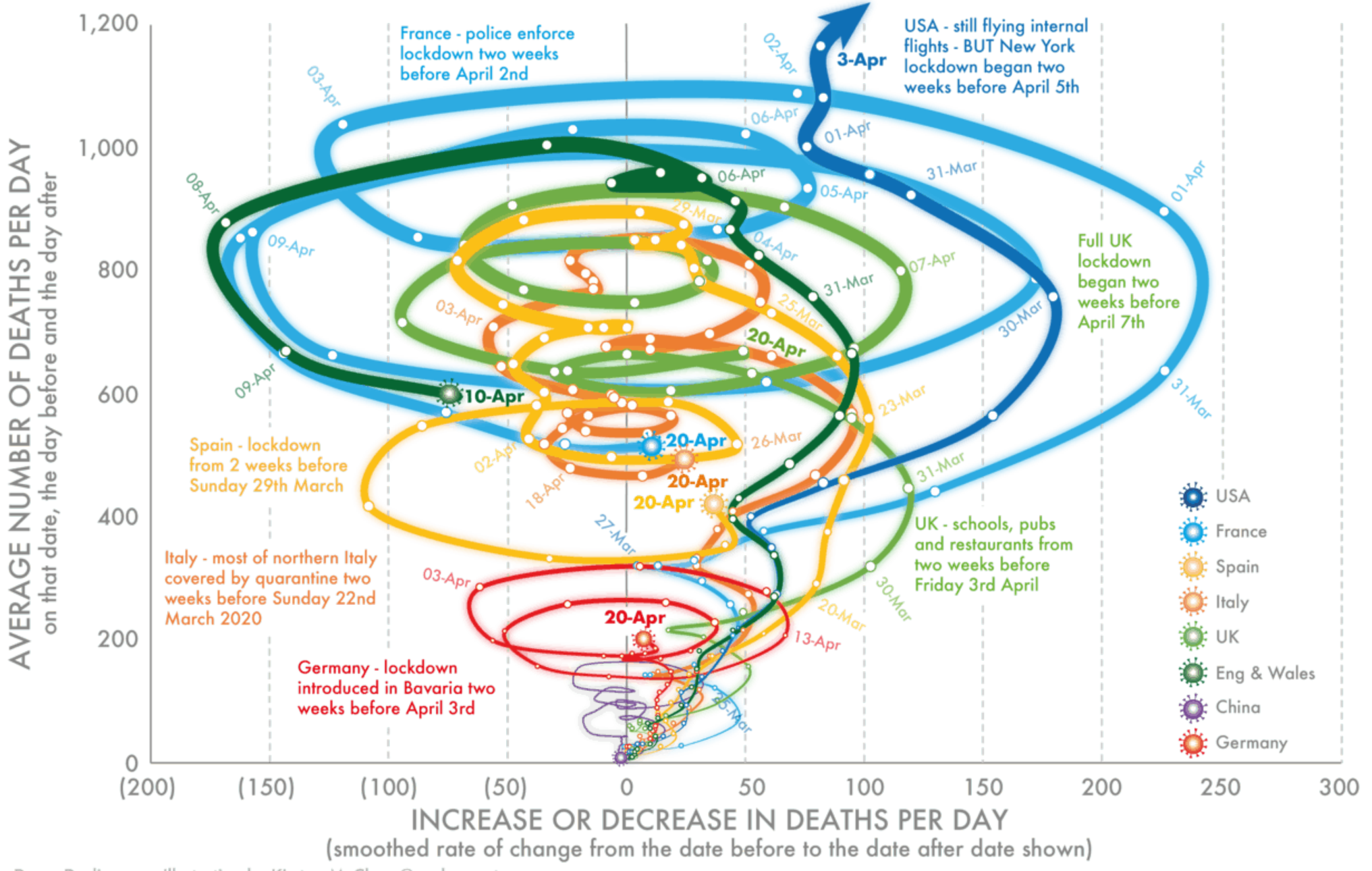
- Nothing special about these wavelengths in terms of physics
- Wavelengths don’t mix to magically produce other wavelengths
- Violet light is violet, not some mixture of red and blue

# What is violet anyway?

- As noted earlier, colour is a perceptual property, not a property of light
- 425 nm light definitely isn't a mixture of red and blue
- Violet is whatever I think it is
- Primary colours *are* primary
  - We can mix to create perception of (roughly) any other colour
- Primariness is a function of physiology and neural processing

# Impairments

- Colour perception is not universal
- Can’t rely on other people seeing colours the same way as you
- Mutations affect cone cells, leading to deficiencies
  - Particularly involve X chromosome proteins, so more common in males
- Range from mildly reduced ability to distinguish some colours (**anomalous trichromacy**) to (rarely) complete **monochromacy**
- Unfortunate: **colour** is a very handy and widely used carrier of meaning




DannyDorling.org. Illustration by Kirsten McClure @orpheuscat

# Handle with care

- For UI design and especially data visualisation, colour is really useful
- Do use it, but thoughtfully
- Try not to make it the sole locus of information
- There are tools to help
- We’ll look at this a bit in this week’s practical

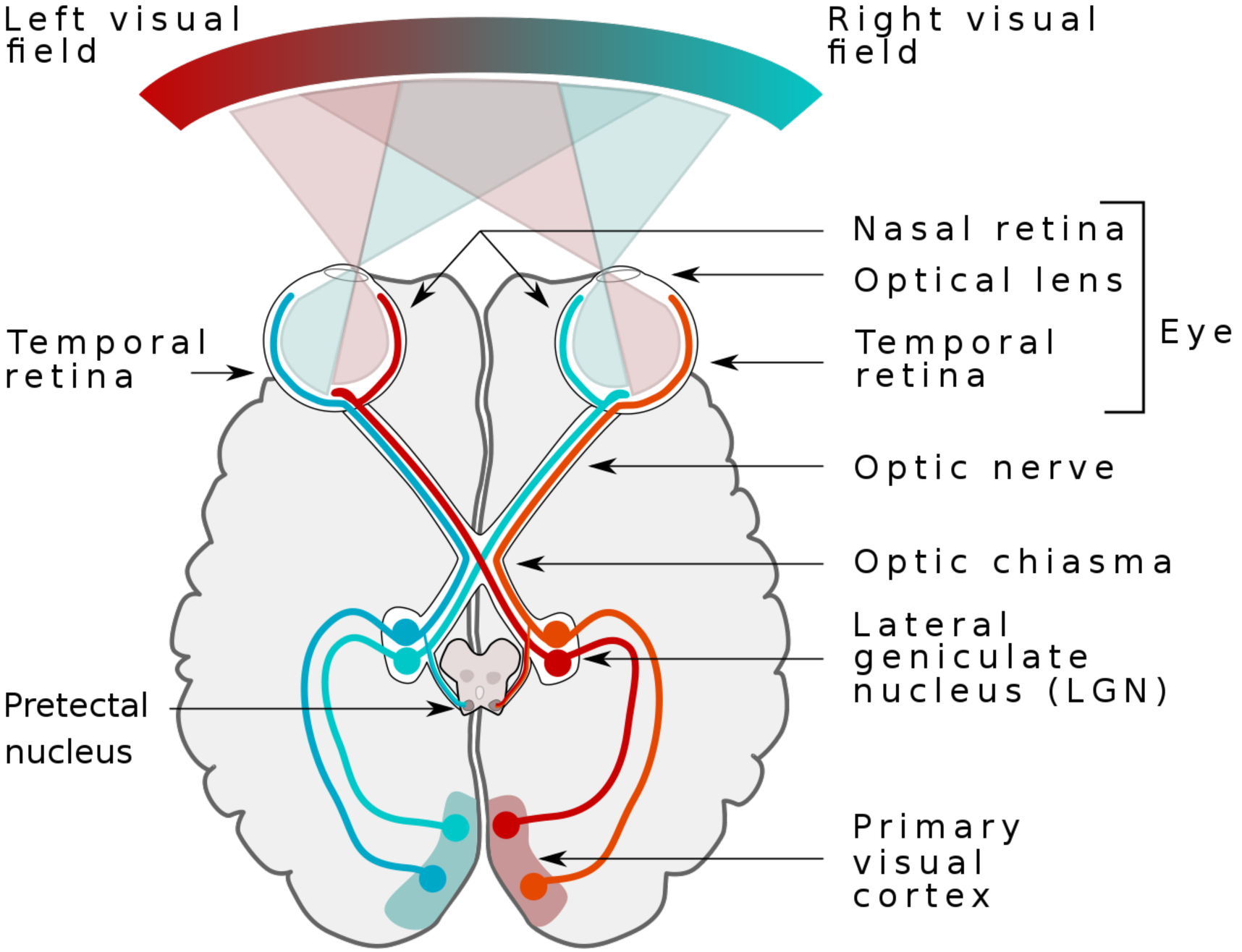
# 2.6 Side By Side

# Stepping onto the visual pathways

- Signals from the retina *finally* travel out along the optic nerve…
  - (through the blind spot)
- …for lots more processing in the brain
- Much of which is — frankly — not very well understood

# Crossroads

- First stop is the **optic chiasm**, a sort of traffic intersection
- Signals from both eyes come together
- …only to be split up and shipped out again in different company
- Signals from **nasal** half of each retina cross over to other side of the brain
- Joining the signals from the **temporal** half of the *other* retina


Left visual field
Right visual field
Nasal retina
Optical lens
Temporal retina
Eye
Temporal retina
Optic nerve
Optic chiasma
Lateral geniculate nucleus (LGN)
Pretectal nucleus
Primary visual cortex

# The world divided

- Nasal right retina and temporal left retina both capture the right hemifield
- Nasal left retina and temporal right capture the left hemifield
- From this point on, signals are grouped according to which side of the world they come from, not which eye
- Everything on the right hand side gets processed by the left hemisphere
- Everything on the left hand side gets processed by the right hemisphere
- Each hemisphere pretty much **loses sight** of the other side

# Filling in

- Hemifields from each eye overlap quite a bit, but jointly cover more territory
- Only the nasal edition has a blind spot
- One of the reasons we are almost never aware of this hole in our vision is that it can be filled in from the other eye
  - Though more fundamental is that brain doesn’t want to be aware

# Zaza

- Eyes are spatially separated, so they see things from a different angle
- Comparing and contrasting the views allows stereoscopic depth perception
- So bringing together information from both eyes for processing make sense
- Hemifield split is unintuitive, though
- Can lead to odd pathologies

# Hemianopsia

- If you lose or start without sight in one eye, you still see a whole world
  - More limited field of view, poor depth perception, but still see left and right
- If visual processing pathways disrupted downstream of optic chiasm, eg by stroke or traumatic brain injury, may impair ability to see a whole side of the world
- Eyes work fine, still sense both halves of the world
  - but can’t perceive one of them

# Grand Central

- Next stop after optic chiasm is the **thalamus**
  - main routing hub for sensory information into the brain
- Visual signals go specifically to the **lateral geniculate nuclei** (one each side)
- LGN does a lot of processing that is only poorly understood
- But notably the neural inputs to the input are not just **feedforward** from the retina but also a lot of **feedback** input from later visual processing

# Shock of the new

- One thing LGN is believed to do is take a kind of **temporal derivative**
  - comparing current inputs to **lagged** versions — what was seen before
  - picking out **differences** — things changing
- Analogous to spatial differencing performed in the retinal ganglion cells
- Once again, processing is tuned for perceptual utility
- **Novelty** is interesting

# Early onset attention

- Another thing LGN is thought to do is apply some **attentional filtering**
- Selecting & deselecting aspects of the visual signal according to what we're interested in perceiving right now
  - What we are **paying attention** to
- This is not fixed feature extraction, like in the RGCs, it varies over time
- Feedback signals from higher process are saying:
  - … boring, boring, blah, boring, ooh *hang on*, tell me more …

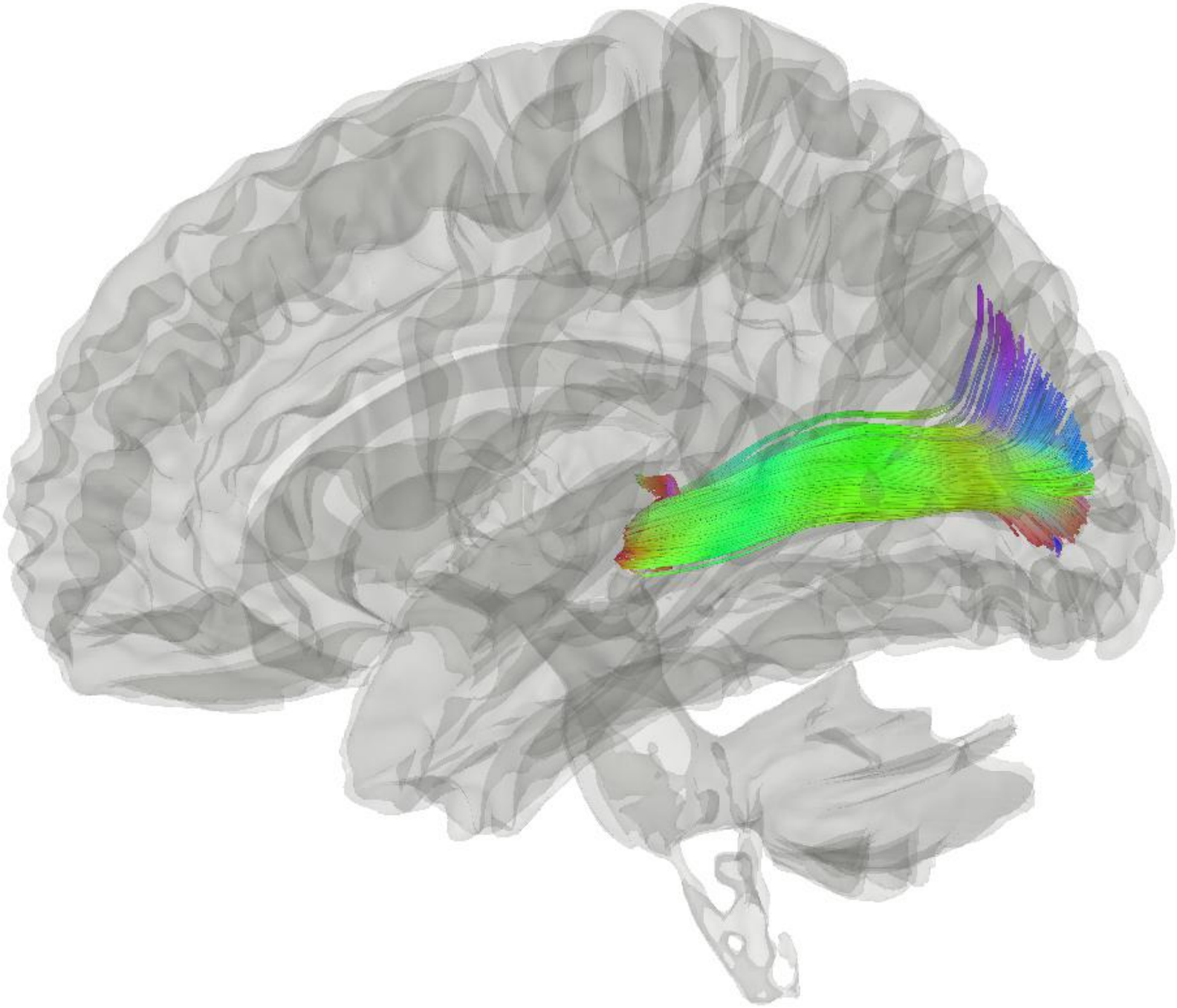
# Refrain

- All this processing is still happening at a very low level
- Tweaking nerve firings from to tiny patches of space and moments of time
- None of it is yet assembled into objects and events, not yet translated into **knowledge** or **behaviour**
- But these fragments are the **raw materials** of visual perception
- Yet again, brain is right there in the mix, picking and choosing, amplifying and suppressing, making connections and — as necessary — **making stuff up**

# 2.7 Doors & Corners

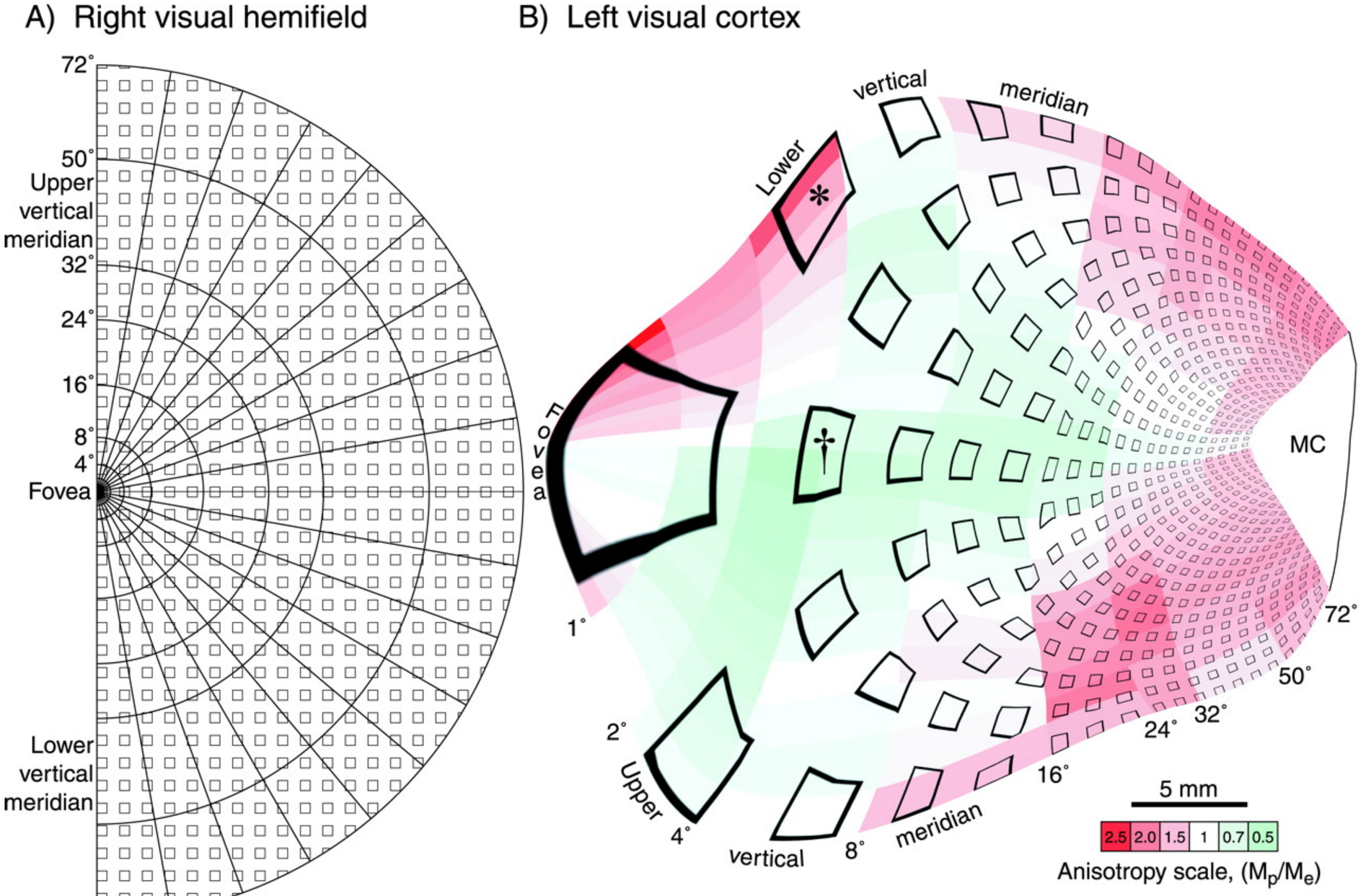
# Radiation

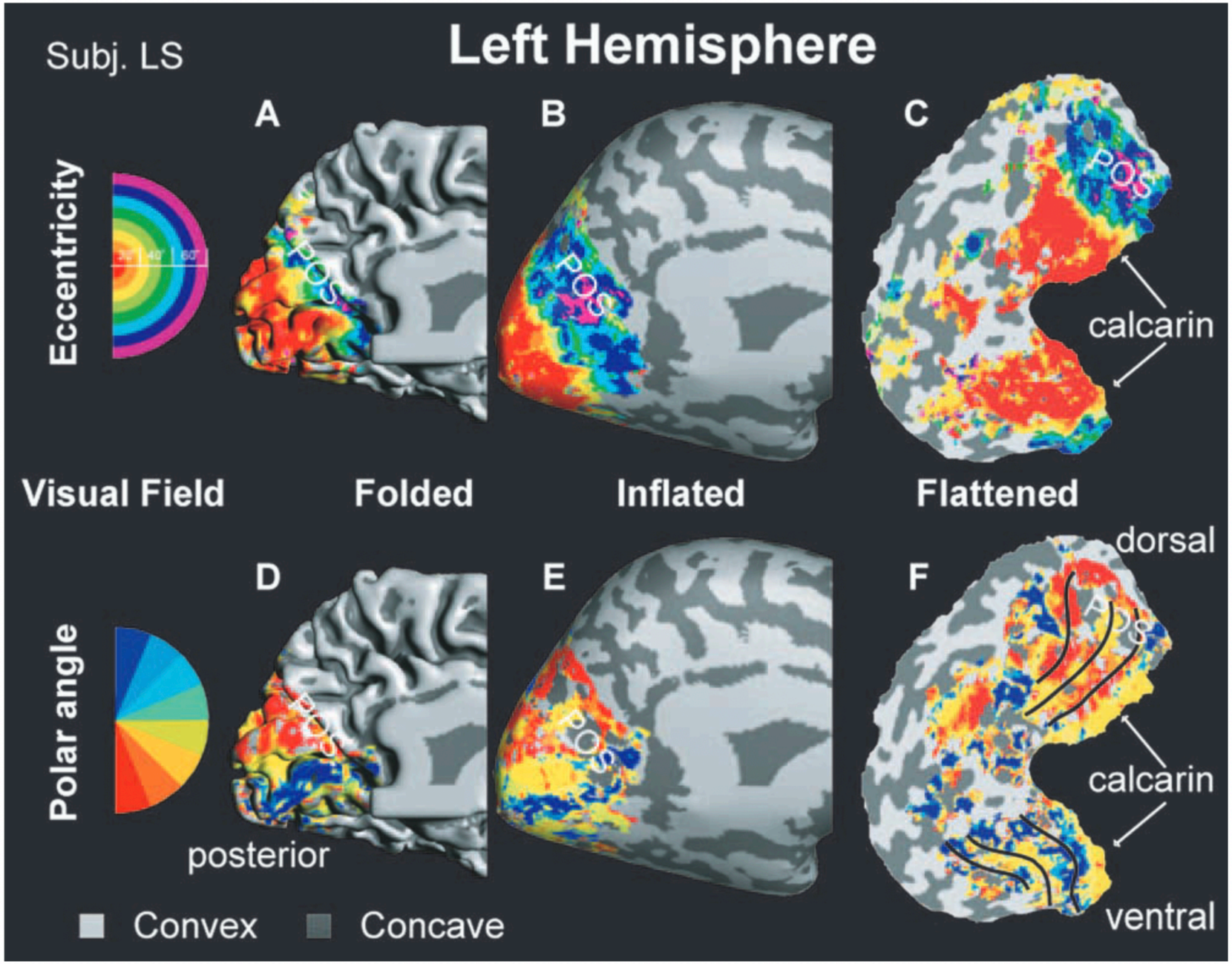
- Processed signals from the LGN propagate onwards in what is known as the **optic radiation**
  - Nerve fibres fan out to a relatively broad region of brain surface at the rear of the brain
  - **Striate cortex** or **primary visual cortex**, or often just **V1**



# Retinotopy

- The distribution of visual signals over the extent of V1 is highly non-uniform
  - Recall the wildly varying receptor densities over the retina
- But, spatial coherence is relatively preserved
  - Signals which originate close together in the image space tend also to be close together in the visual cortex
- This spatial arrangement is known as the **retinotopic map**


A) Right visual hemifield
B) Left visual cortex
72°
50°
Upper
vertical
meridian
32°
24°
16°
8°
4°
Fovea
Lower
vertical
meridian
vertical
meridian
Lower
Fovea
1°
2°
Upper
4°
vertical
8°
meridian
16°
24°
32°
50°
72°
MC
5 mm
2.5
2.0
1.5
1
0.7
0.5
Anisotropy scale, (Mp/Me)


Subj. LS
Left Hemisphere
A
B
C
Eccentricity
POS
POS
POS
calcarin
Visual Field
Folded
Inflated
Flattened
D
E
F
dorsal
Polar angle
POS
POS
POS
calcarin
posterior
ventral
Convex
Concave

# Pragmatism rules

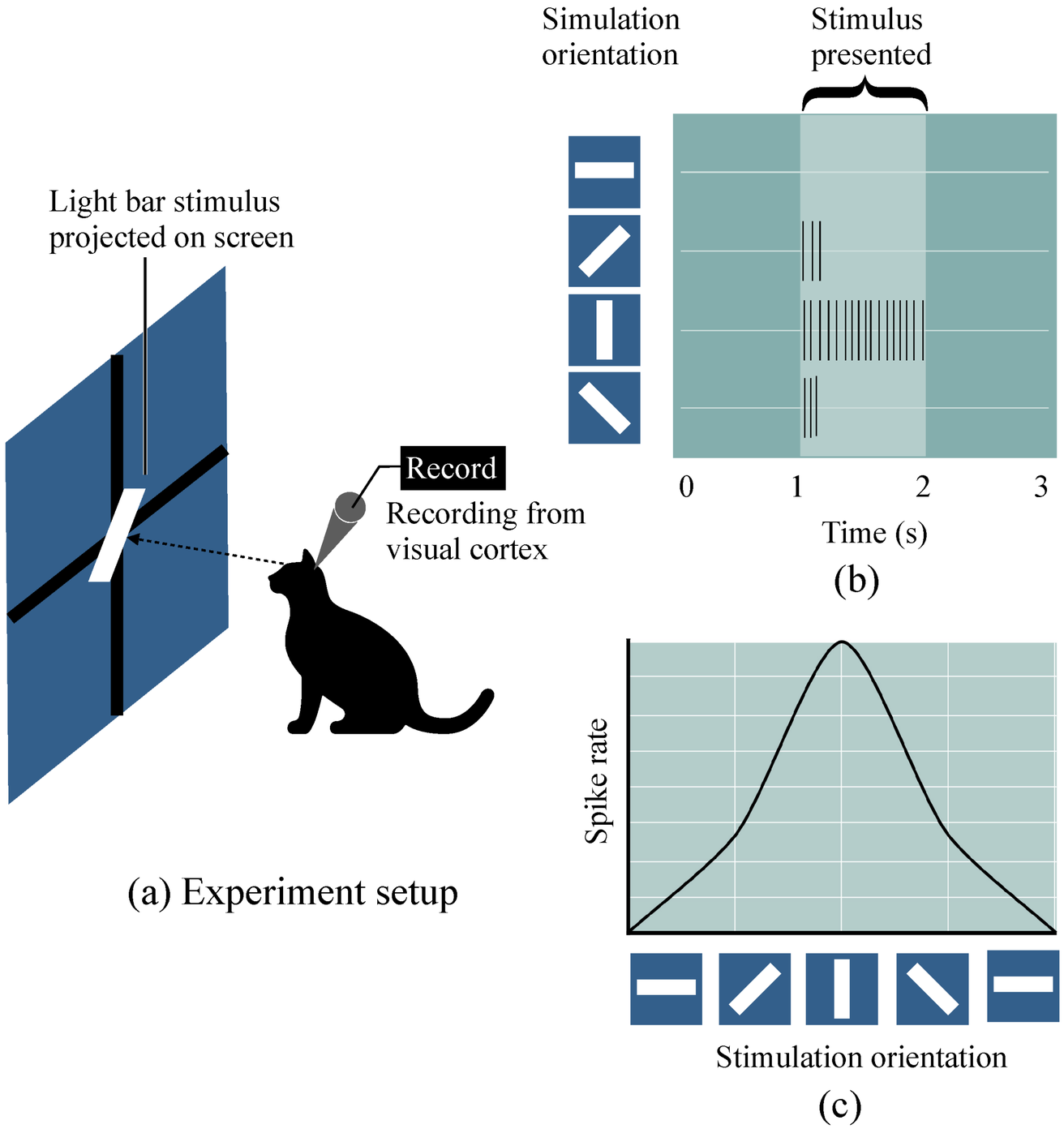
- We keep coming back to this, but:
  - Retinotopic spatial coherence is not a physiological necessity
  - Neurons could be wired up any old how
- This structure is an empirically successful evolutionary strategy
- It has utility because spatial relationships really exist out there in the world
  - Relationships that are meaningfully represented in the detectable image
  - If the information is there, use it!

# Feature extraction redux

- Recall discussion of receptive fields and pattern sensitivity in RGCs
- Local connectivity within retina means receptive field is a (small) spatial region
- Centre-surround antagonism **tunes** the RGC response to pick out brightness changes in that region
- But RGC responses are extremely simplistic: any boundary will do
  - edge, spot, corner, who cares?
- In particular, they are not sensitive to edge **orientation**

# Orientation matters

- This is obviously not true of our vision
  - You and I can see orientation just fine even if RGCs can’t
- And it turns out — famously — not to be the case in visual cortex
- As demonstrated by Hubel & Wiesel in 1959, V1 contains neurons that spike when exposed (in their receptive field) to edges of a particular orientation


Light bar stimulus
projected on screen
Record
Recording from
visual cortex
(a) Experiment setup
Simulation
orientation
Stimulus
presented
0
1
2
3
Time (s)
(b)
Spike rate
Stimulation orientation
(c)

# Building blocks

- First clear example of complex features being built up from locally-connected integration of simpler features
- This pattern of combining spatially (and/or temporally) related features into composites is believed to recur throughout visual processing
- Also the inspiration for a lot of computer vision and deep learning methods
- Maybe fundamental to thinking in general?

# Aggregation grows receptive fields

- Inputs to a neuron in V1 — or later — are outputs from earlier neurons
- Those neurons have their own receptive fields
- Anything in any of those fields may affect the downstream neuron
- So the receptive field of a V1 edge detector neuron (or any complex feature extractor) is the **union** of the receptive fields of its antecedents
- If those are non-identical — which they always are — the union will be **bigger**

# Aggregation increases complexity

- Combinations of simple features — like orientation-independent differences in RGCs — make more complex features — orientation-sensitive edges in V1
- Once we have complex features, we can combine them too
  - Combinations of edges make shapes
  - Combinations of shapes make spaces
  - Combinations of spaces make… everything?

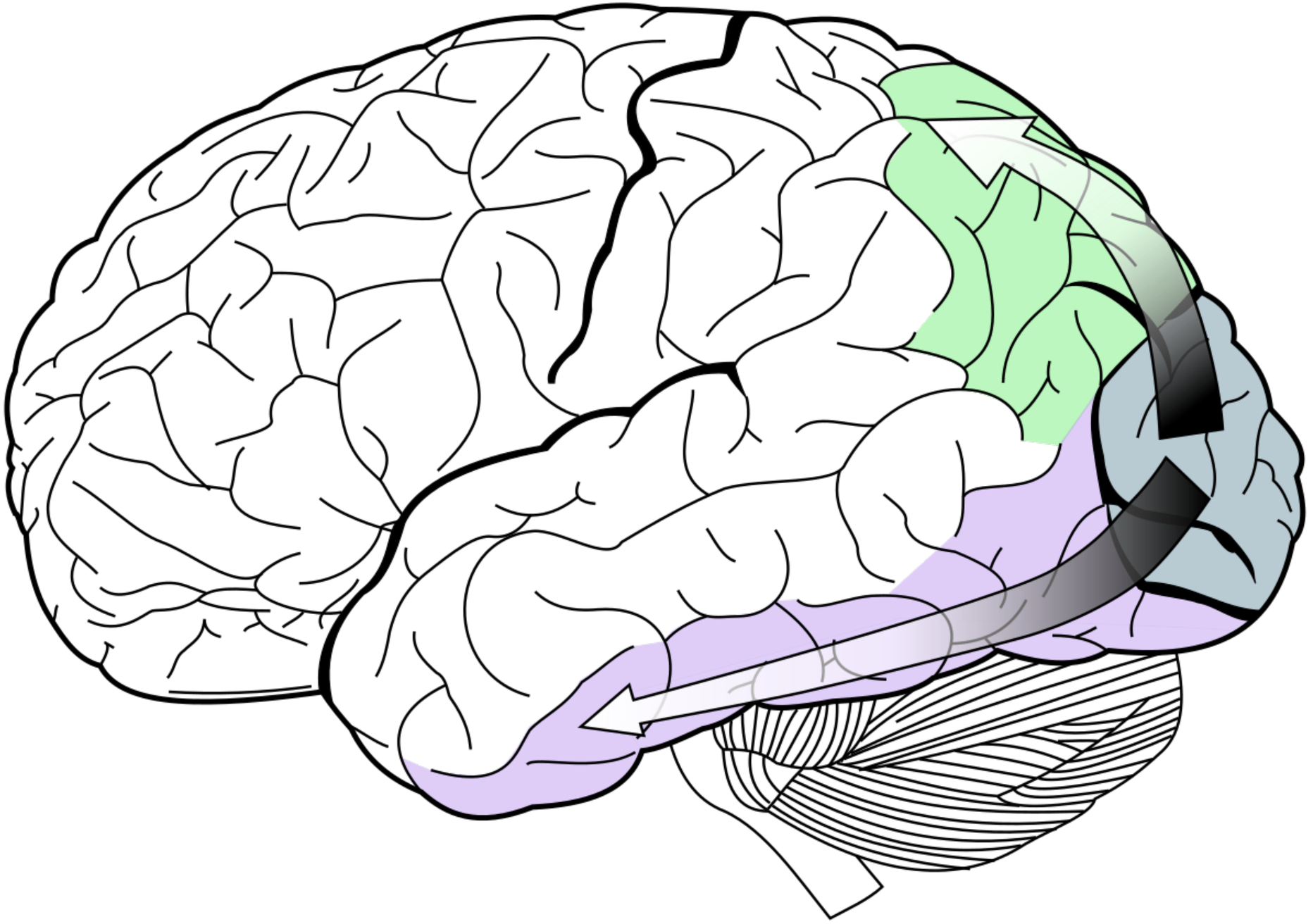
# Platitude for the day

- This is sort of the **lesson** of visual perception
- We can model the whole world as the (weighted) sum of its parts
- Provided we have a coherent relational structure for those parts

- Especially if we’re willing to make stuff up to fit our model

# Deeper into the vortex

- Signals go through several layers of processing in V1, percolating through vertical slabs (knowns as **columns**) of related cells
- And they propagate onward all over the shop, to all sorts of places, modelling all sorts of things
- But in particular there are two reasonably distinct pathways that seem to have different purposes



# Two streams hypothesis

- **Ventral stream**
  - via underside of brain to temporal lobe
  - **what** — object perception and recognition
- **Dorsal stream**
  - via top of brain to parietal lobe
  - **how** — spatial organisation and behaviour

# Fin?

- So:
  - Visual information is gathered, aggregated, folded, spindled and mutilated in the visual cortex
  - Passed forward to other areas for increasingly complex and specialised processing with increasingly specific purposes
  - Nearly all of this we don’t really understand at all, so beyond acknowledging that it somehow happens, is there anything more to say?

# 2.8 Constancy

# Reality is (somewhat) stable

- We like to believe the world is, mostly, coherent
- Things exist out there — and persist
  - They don’t suddenly vanish or turn into something else when we look away
- There is structure, continuity, to the ways in which reality behaves
- If this was not true, life would be pretty difficult

# Appearance is variable

- What things look like changes over time
  - Illumination
  - Viewpoint
- Even simple objects may manifest radically differently on the retina

# Hobgoblin of little minds

- **Constancies** are visual perception's attempts at apprehending what is consistent about the world in changing conditions
- Parsing the extremely different visual stimuli into the same conceptualisations
- The brain is astonishingly good at this
- So much so that we don't realise we're doing it
- Or even that it's something that needs to be done

# Excursion into white balance

- Colour film gives really varied results according to lighting
- Film pigments are balanced for specific lighting conditions
  - Usually couched in terms of black body temperature
- Other illumination gives weird casts without careful filtering
- Early video cameras required explicit calibration for white balance
- Our eyes mostly don’t notice these lighting changes at all

# Constancies are omnipresent

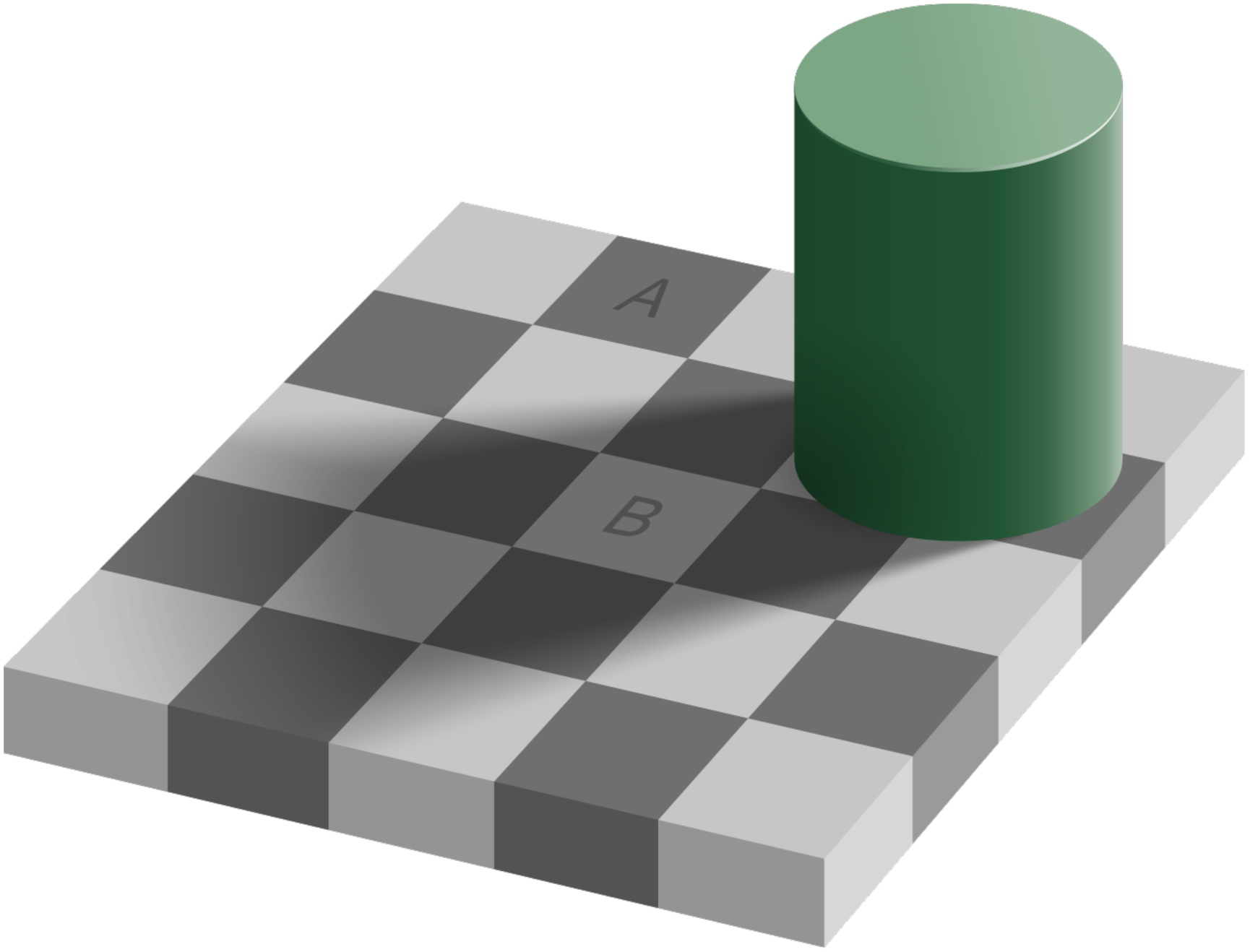
- You mostly can’t help seeing that this screen is **rectangular**
  - From wherever you are in the room
- There is no position where the image on your retina is rectangular
- And even if it were that wouldn’t help because the sensory map doesn’t remotely preserve angles or distances
- Every thing you know about the shape and size and colour and shading of things is a constructive interpretation

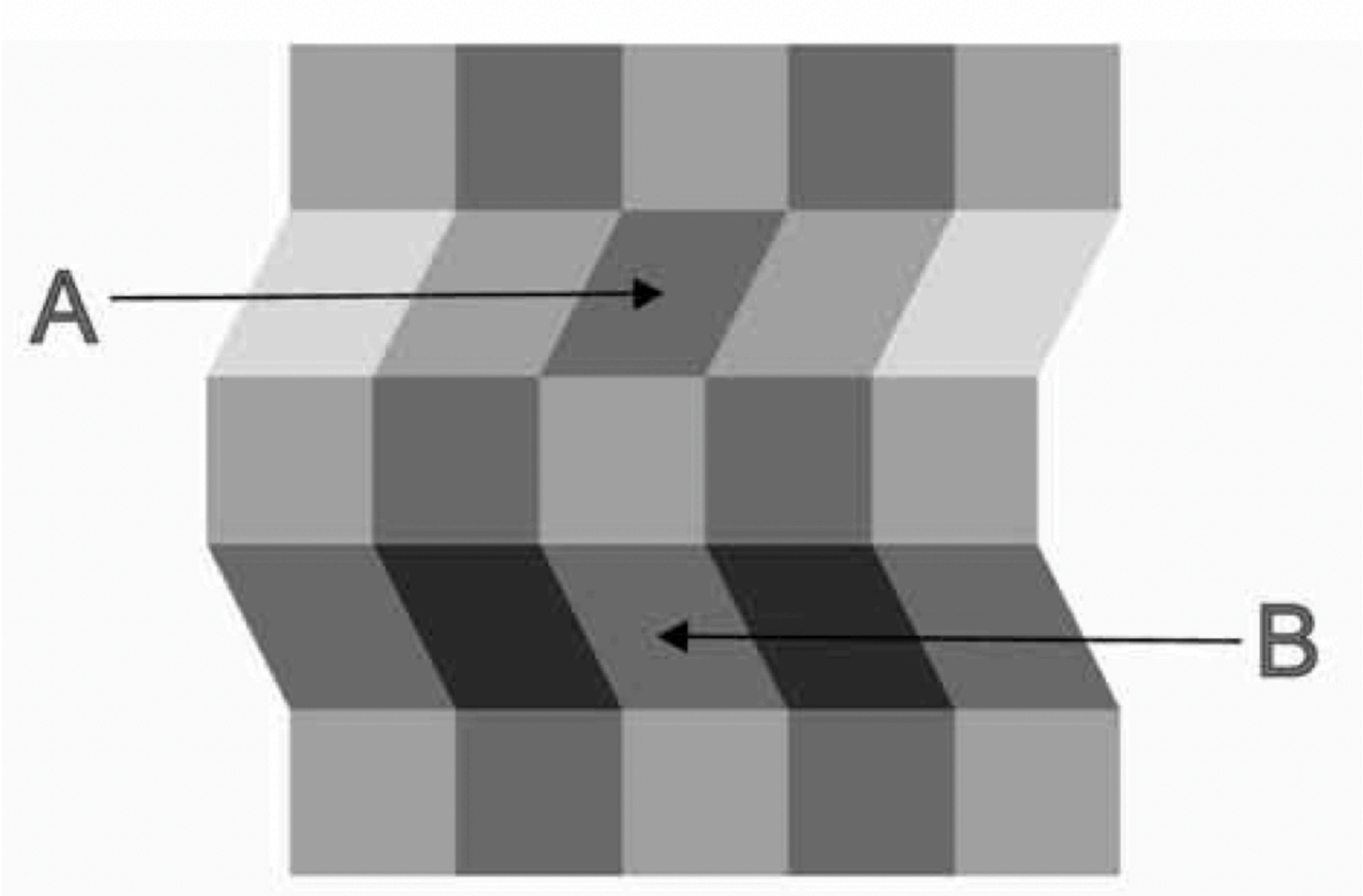
# Check

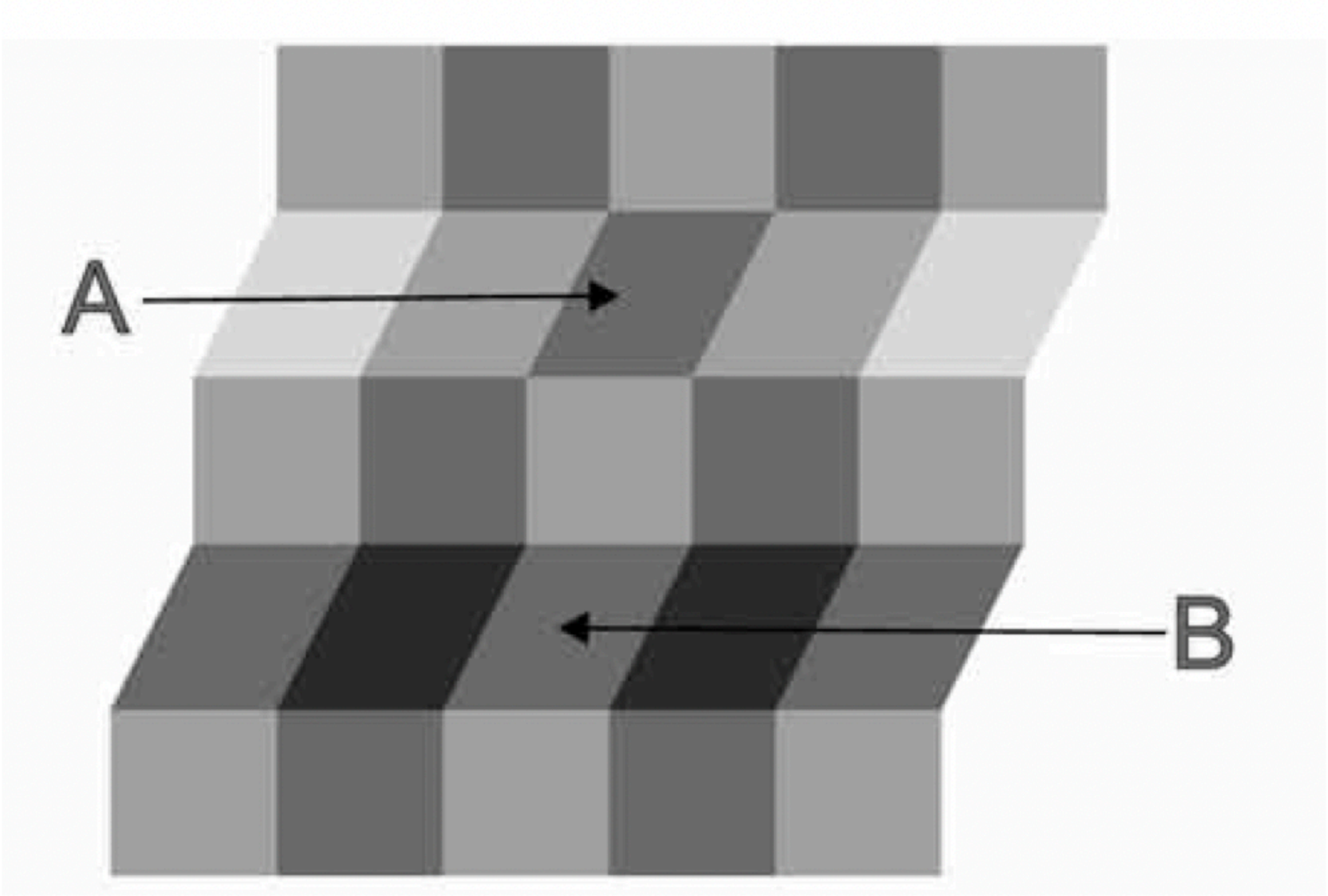
- In the checkerboard example, you will likely see:
  - The white squares are white
  - The red squares are red — and all the same red
  - The squares are squares
  - And they’re all the same size

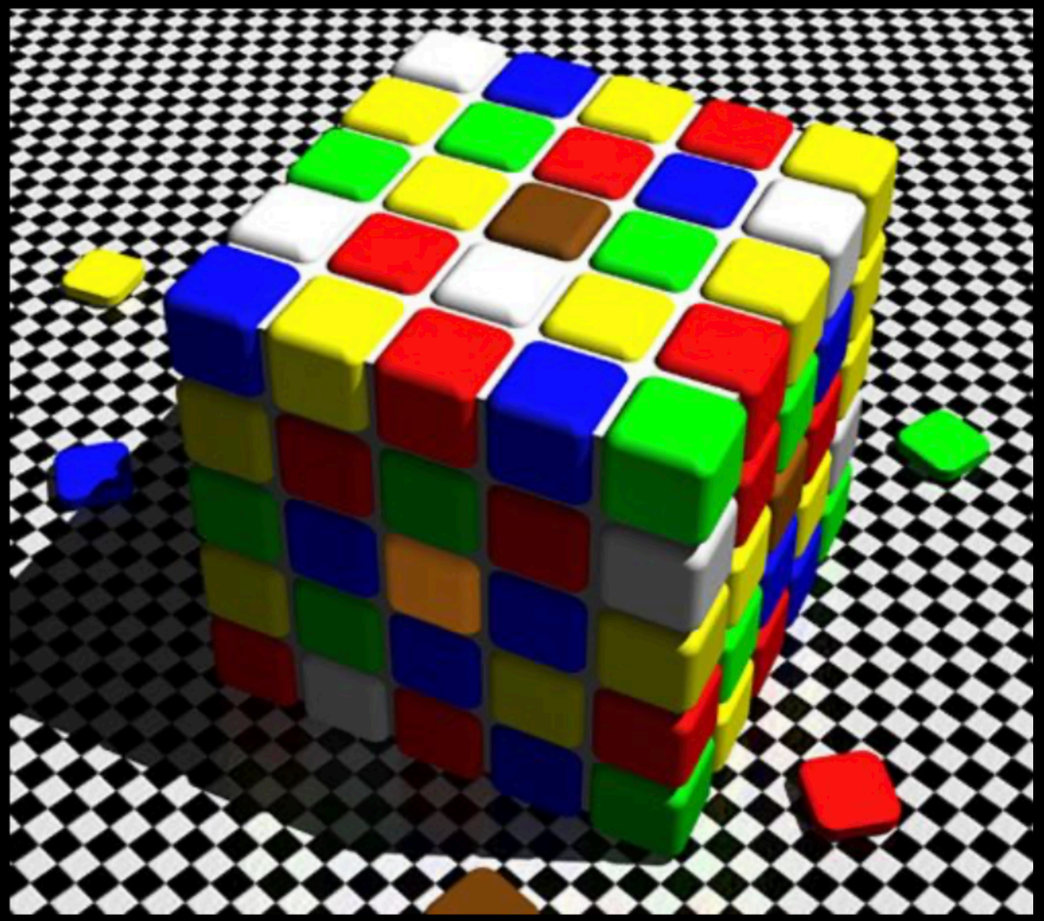
# Mate

- None of these things are literally true in terms of the image contents
- Or the responses being evoked in your retina or early visual system
  - There are no straight lines or right angles in your physiology, only the idea of straight lines, the notion of right angles
- All of this is imposed in the process of interpretation
- But for most people it’s really hard not to do


A
B


A
B


A
B



https://www.echalk.co.uk/amusements/OpticalIllusions/colourPerception/colourPerception.html

# The Dress


祖
祖
ワィ
ワィ

# What does constancy tell us?

- In a sense, nothing new
- Visual system constructs your perception of the world based on its models
- Often takes the form of a huge pile of hacks that mostly kinda work
- Some of these are extremely sophisticated, some are cheap as chips
- You will almost never know which is which
- Mostly you can just rely on them to maintain sanity in an insane world

# 2.9 Attention Must Be Paid!

# Teleology

- We’ve mentioned attention before
  - Ifat will talk a lot more about it in a few weeks
- Integration is the way visual perception builds the world
- Constancy is how it imposes some kind of sense on it
- Attention is how it gives its world modelling **purpose**

# The Distraction Club

- You are bombarded with sensory information every minute of the day
- Some of this will make you happy
- Some of it may kill you or save your life
  - Rare, but really important!
- Most of it is of no interest at all
- You’ve seen it all before, it’s just the world ticking over

# But which is which?

- How do you know which droplets in the torrent to pay attention to?
  - Spoiler: you often don’t, and neither does anyone else
- But: you might be able to pin the problem down a bit with goals

# Who are you? What do you want?

- Things that can help with filtering the oncoming storm:
  - What am I doing?
  - What are my goals?
  - What information do I need for those goals?
  - What do I not?
- There is no one static answer, these things change moment to moment
- But in the context of a moment the focus may be clear

# Cocktail party problem

- We don’t know how attention works, but it’s extremely clear that it does
- A famous example is the **cocktail party problem**
  - You are surrounded by the good and the great, yakking away
  - But you can still have a conversation, focus on your companion, shut out the clamouring hordes
  - The noise is still reaching your ears and triggering nerve impulses, you’re just not paying attention to it

# Simons & Chabris selective attention test

- Very famous example, many of you may have seen it before
- Poses a challenging task of picking out behaviour amongst distractions
  - Count basketball passes when other passes are happening around
- Your job is to shut out the noise


copyright (c) 1999 Daniel J. Simons. All rights reserved.

# Inattentional blindness

- Lack of — or misdirected — attention may lead to a failure to perceive things
- Experimental examples may be contrived, but this happens all the time
- People miss their station or crash their cars talking on mobiles
- People are robbed in plain sight and don’t notice for hours or days

# Attention in interfaces

- Harnessing attention may be crucial to getting people to use your UI
- Or to gathering experimental data
  - We saw in last week's lab how easy it is to get bored with repetitive tasks and for your attention to wander — what does that mean for results?
- So it's helpful to at least think about how people will attend to what you are setting before them

And you will read this last

# You will read this first

**And then you will read this**

Then this one

# 2.10 Coda

# What you see…

- We have barely scratched the surface of visual perception
- Completely omitted important topics like **visual grouping** and **optic flow**
  - Barely mentioned **motion**, so many fun optical illusions missed

# What you see…

- We have barely scratched the surface of visual perception processing
- Completely omitted important topics like **visual grouping** and **optic flow**
  - Barely mentioned **motion**, so many fun optical illusions missed
- Huge tracts of visual processing in the brain are dedicated to **faces**
  - Incredibly important to our existence as social animals — nada

# …is what you get

- What I hope you come away with is a sense of how **non-trivial** vision is
- How much amazing effort most people's brains are putting in all the time to construct and deploy the magic of visual perception
- To whatever extent you have this ability, don't take it for granted
- Think about what it is doing for you and how you can make the most of that
- And make an effort to support those whose vision is impaired in any way