

The following content is provided by MIT OpenCourseWare under a Creative Commons License. Additional information about our license and MIT OpenCourseWare in general is available at ocw.mit.edu.

JEREMY WOLFE: Good afternoon. In response to-- in response to public demand here, let's see. How many people-- you got-- again, the colors may have faded by now, but check if when you tilt your head. The effect changes. Yeah. How many people still have some sort of effect? It's really nice to know that this course has some lasting impact on people.

And Mara, maybe we should try-- shh! Maybe we should try posting the PowerPoint on the web, because then if you want to build up your own McCollough effect, you can do this and entertain yourselves.

If you have been following the story thus far, you should have been-- you should basically have the idea that at the front of the visual system-- or of any sensory system, you got all sorts of information coming in.

That the job of early processes in the visual system is to-- well, to do the things that the early parts of visual cortex are doing and say, oh look, there's a little line at this point in the visual field. Oh look, there's a dot there and there's a line here and it's moving like that. Little bits of information all over the place. Too much of it for you to handle.

So last time, I talked about this bottleneck of attention that allows only some of it through to processes that would do things like, say, recognition. And then-- I'm going to run out of places to draw, we'll take a little detour here. Somewhere up here, you're going to get perception.

And the job of today's lecture is to convince you that that percept, your current understanding, of the world is always the result of a many, many inferences, many, many guesses about what the nature of the world might be.

Because this information-- not only is there too much information coming in, there's also too little information coming in to specify exactly what's going on out there in the world. Consider just a couple of the problems I think-- are they on the handout? A couple of them are.

Well, one of the obvious ones is the world is 3D. The world-- your job is to figure out what's going on out in the world. The input is inherently 2D. That retina that you've got is a 2D surface. So if you're seeing 3D, you are recovering that 3D information from essentially 2D input. You're collecting information about light intensities. You don't care about light intensities. You care about surface properties in the world.

So every patch that you're seeing-- if you're looking at this spot right here, what you're seeing is the product of the surface-- the properties of the surface and the properties of the illuminant, the properties of what's lighting it up. You don't care about the properties of the illuminant. You want to recover just the properties of the surface. So how do you ignore-- how do you successfully ignore the properties of illumination? We'll say a little bit about that.

The world you're looking at is an essentially stable world. I mean, things move around in it, but the whole world doesn't jump around. But you're looking at it from an inherently unstable vantage point. You're moving around, and more to the point, even if your stock still, the way you're looking at the world is you're moving your eyes around.

Try this for a moment. Look at the lower left-hand corner of the screen-- with that now without moving your head. Well actually, that might be a little large. So look at the lower left-hand corner of this McCullough effect test pattern. Look at the lower right-hand corner of the McCullough effect test pattern. Did the McCullough effect test pattern jump when you did that? No, it didn't.

Why didn't it jump? Because when you moved from-- when you're looking here-- do I have a laser pointer today? No. Oh well. When you're looking here, the image is-- the bulk of that square is to the right of fixation. When you're looking here, the bulk of that image is to the left of fixation. So it's in two different spots on the retina. If I put something here and something here on your retina, it'll look like it's moving. Why didn't that look like it's moving?

The reason is that when you tell your eyes to move, you send a copy of that command, in effect, to visual centers of your brain saying, look, I just told my eyes to move, kindly ignore the resulting smear. In fact, I want you to do two things. I want you-- your visual system-- I want you to ignore-- to shut down during the course of the eye movement. And I want you to compensate for the fact that everything has been displaced.

You can see what would happen if that was not the case by taking your finger and poking your eye. So on your eyelid, you should try this because it's more interesting if you actually try this. If you wiggle your eyeball-- you can do it slowly, or you can do it quickly, look at me and poke your eyeball around. You will notice that all your friends look funny, but you'll notice that things are jumping around.

Why is that? Well, look. Millions of years of evolution did not provide you with a mechanism that said, I am now going to poke my eye, please send a copy of that signal to the visual centers of the brain saying to cancel that out. There's no cancellation signal here. And so you see the image moving around. You don't see the image moving around when you move your eyes normally because you're compensating. You're compensating for it.

All right, so you're collecting all this information, you're doing your best to register it in a stable kind of a way. Oh, look at that. It also says I'm going to demonstrate the visual-- the vestibular ocular reflex. Let me do that just because you might as well get to use your fingers some more.

You also want the world not to jump around too much when you're moving your head. One of the things you do very reflexively is, if you rotate your head one way, your eyes counter-rotate the other. That's a very quick reflex. If you want to see how quick it is, try this.

Hold your finger out in front of you, look at your finger, and move your head back and forth, and just keep your eyes on the finger. No problem, right? Yeah, yeah, you know you can do that. Now, try to, at the same speed, move your fingers and try to keep your eye on the finger. Can't do it.

It doesn't work because that reflexive movement-- or that tracking movement is-- there are more neurons involved. It's a slower process. This vestibular ocular reflex is a very short latency kind of a reflex designed to keep the input relatively stable.

So, you take all that lovely input in, you got all these little bits of information all over the place, and then you've got to make your best guess about what it is that you're looking at.

The reason that this is an interesting picture-- it's in the book, by the way, so when you can't see it here, you can go and study it in the book until you can see the Dalmatian dog that really is there. How many people can see it now? Oh, we got most of them. Just for-- yep. His head is right above my finger, front paws, back paws. He's on a road sloping from lower left to upper right.

The point of a picture like this is that it slows down the process of inference enough that you can feel it happen. Normally when I look out at you, for instance, my visual system kicks up one and only one interpretation of what I'm looking at so rapidly that I never notice all the work that's involved. The purpose of this picture, and really, the purpose of this lecture is to show some of the work that's involved.

So, what you've got here is a bunch of isolated little black and white regions. In order to figure out what's going on here, you've got to decide who goes together. How are we going to decide who goes together? Well, let's think of this in the context of a bunch of little line segments that-- you can decide that some of these guys might go with some of the other ones, but what I'm going to do is rotate all of them by 90 degrees, as I recall. Boop. Well, no, maybe some other orientation, but anyway.

Now, there's a bunch that go together-- they're the same lines on the screen, but now some of them hang together. Got that potato-shaped thing there. Why? What makes these guys go along with each other now in a way that they didn't before?

Well, if you've got an isolated-- if you're a little chunk of brain whose job it is to figure out where contours are out in the world, what you're getting from earlier in the visual system is word that there's a little bit of a contour here, there's a little bit of a contour here, a little bit of a contour here. I wonder if those should go together. Now how do we decide if a little bits of contour might go together?

Well, you might do something like this. If you've got-- so suppose you've got this. Whoop! Come here, where's my little-- there we go. If you've got a line here and you're asking, what's the best bet about where this-- if this is really a piece of a continuing contour, where is this likely to go next? Well it might make a hairpin turn and go off that way, but that doesn't seem really.

More likely, it's going to go off in the direction-- in something like the direction that it's pointing. And that turns out to be-- whoops, these are the random ones. There's the potato-shaped ones. This is a very rule-governed sort of behavior.

If you've got a bunch of little line segments, as long as the deviation here isn't more than about, as I recall, 30 degrees or so from collinear, you're willing to string those together pretty happily. If it starts to be more than that, you're unlikely to string them together. The beginnings of an effort to tie little pieces of information together into larger structures.

All right. What do you see here?

STUDENT: Two lines crossing each other.

JEREMY WOLFE: Two lines crossing each other. A reasonable interpretation, though it's not the only possible interpretation of this. I mean, it could be something like this, two birds kissing each other or something like that.

Why do you see-- and in fact-- this is something like that. And you see that as an X with two lines crossing each other. But if I provide enough other details here, it's a sea lion?

[LAUGHTER]

It's a sea lion. It's a-- I don't know what it is. The point is, he's still a lousy artist, it hasn't gotten any better. The point is more or less the same, that you've got this little-- a process that's worrying about that little piece of line segment, gets to this junction-- it's busy doing the three roads diverged in a yellow wood kind of thing trying to decide which way to go. And it's guessing that, all else being equal, I should probably go with that.

This is-- it says on the handout-- no, I mean on the-- oh yes. OK. Sometimes it goes by the name of good continuation. It's one of a variety of so-called grouping rules that were developed first by the gestalt psychologists starting in the early part of the 20th century. And they're really rules for figuring out how bits of the scene might hang together.

You can see a similar process going on here. You can see that as three isolated line segments, but you probably don't. You see that as a curvy line that's occluded. If I suddenly reveal this, you don't go, oh, amazing! That's what I thought was there.

That's also highly rule-governed. If you've got a line segment and you've got another line segment, you're perfectly happy to see this one and this one is connected. If I-- well, how about this? This will do if I put one up here-- all the more, if I erase the stuff in between, that you're less likely to see as connected.

Now why are you less likely to see it as connected? The rule turns out to be that if I can connect two lines with a smooth curve, I'm in business. I'll be willing to see those as connected. But if I have to put an inflection in it to make it work, then it doesn't look as convincing.

I mean, it's not that I deny the possibility that this could ever connect with that, but if I give people a bunch of stimuli like this and ask, good connection or bad connection? The good connections are ones that can be done with a single, smooth curve that can't quite-- I think you probably have to get an inflection in there somewhere. And if you have to inflect the curve, it fails, or people report it as looking less convincingly continuous.

All right. It's voting time here. Again, so here, we've got a whole bunch of isolated guys, but they do seem to have something to do with each other. If you had to pick, this being organized by columns or rows, how many vote for columns? How many vote for rows? Eh, there ain't much to choose, right?

But if I do this, OK, how many vote for columns? How many vote for rows? So we've now skewed it very heavily in the direction of columns. And all that I've done is change the proximity of elements.

Once the distance from one item to the next item in a vertical direction is closer than to the items in the horizontal direction, this-- it's another one of these gestalt grouping rules of proximity takes over and says, well, all else being equal, if I had to guess who goes with who, the guys that are close to each other, they probably go with each other.

Multiple rules operate at the same time. So I'll keep the proximity rule working here. Now, you have to vote. How many vote for columns? How many vote for rows? So now I've skewed it very heavily in the direction of rows even though the proximity law, the proximity rule is still going for columns.

Those things are closer to each other in a vertical direction than horizontal. In this case, the similarity is trumping that. You can balance these off against each other. How similar does it need to compensate for a 2-to-1 difference in distance, for instance, or something like that?

But the important point here is that what you're trying to do-- these are demonstration versions of presumably what you're doing all the time. I'm looking out there and I'm seeing regions of-- what am I seeing? Regions of redness.

And there's this-- I see a few disconnected regions of redness, but I'm guessing they're all part of her top. Yeah, your-- that's your top there. Because anyway, yeah, it's you. No, not her. That's pink you're wearing. No, the woman behind you. Well, the woman next to you is also red.

But in this case, I'm not doing a proximity-- sorry. You can see the role of proximity here. So the similarity thing is telling me all those red things are tied together and I'm making it into one piece of clothing. The proximity thing is saying, well, I don't think her red top and her red top are the same red top. That would be a very weird assumption to be made.

On the other hand, the guy she's sitting next to is also wearing red, so maybe they're just wearing one garment. No, I'm probably not going to come up with that assumption either.

But it's these-- what my visual system is doing is continuously trying to cut the world up into meaningful chunks that are going to be worth subsequent analysis. I don't want to go off in to analyze every little pixel in the scene. I don't have the brain power to do that. I want to have meaningful chunks that are worth analyzing. So here, I might decide that the meaningful chunks were the rows, that they are something or other.

OK. One of the reasons that you need-- so there are a couple of reasons why you need to do this grouping business over little elements in the world. One of them I was cartooning over here, which is that early in the visual system, the chunks of the brain that are looking at bits of the world are only looking at very teeny bits and you're going to have to tie those together.

The other reason is that out in the world, contours don't behave well. They tend to do awkward things like disappear on you, and you don't want to get the idea that these are-- that a contour-- well, you don't want to get the idea--

If I've got a contour like that, if, for some reason, bits of it are deleted, maybe because there's an occluder or maybe because something just bad happened in the image, you don't want to lose this whole structure because bits of it have been deleted. So you have a lot of clever mechanisms designed to help you find where the edges are of things out there in the scene.

You can find-- and putting the bits of edges together into a long, coherent one. If you play with Photoshop, you can go and find the edges, I think. Isn't there like a Find Edges filter or something? So try that sometime. Do Find Edges on the filter and find edges on an image, say, a person, and you'll see that it comes up with a lot of edges that you recognize as being related to this person, but it's fragmented all over the place.

And in fact, getting your computer to figure out which bits go with each other is a tricky piece of Work it's tricky for you, too. You just don't know that it's tricky because it works all the time.

All right, so this beautifully boring stimulus is there because you see this beautiful vertical contour. It's there. Now all I'm going to do, I'm going to leave it there, but I'm going to put a new background on it. Isn't that lovely? Why is that interesting?

Well, among the reasons it is interesting is because it's the same gray stuff that was there before. So the top of that bar and the bottom of that bar are still identical. All four of those little rectangles are all identical even though they no longer look identical, because what the system is doing is-- actually, I can probably-- can I do this? I forget what I programmed. Oh, there we go. Look, he's going to turn into that one. Isn't that fun? Whoops, get back there. There we go.

But my real point here-- so what that is is a simultaneous contrast effect, by the way. This one-- this square looks bright because it's surrounded by darker stuff. That square looks dark because it's surrounded by brighter stuff. And even though this bar is continuous in its gray level from bottom to top, it picks up its apparent brightness from the surrounding-- the immediately surrounding contours.

The interesting aspect of this from the point of view understanding where edges are is that, well, look, if it's darker-- if the bar is brighter than the background here and darker than the background up there, there must be a place in the middle where it's gone, where there is no contour. But you don't see that. It's actually a fairly-- I don't think-- I was too busy making this little thing move and I didn't put in a-- no, I forgot to do that. Oh well.

It's not a small region. There's a fairly sizable region of the middle there where there is no physical contour. But you fill it in. You know in some fashion that that contour is there. That's called a subjective contour where you're completing a contour that doesn't have any real support in the image. That's the piece that your Photoshop filter will have a hard time doing, by the way.

Oh, you can also-- we can also do-- oops. We can also do a second-order effect here that's-- well, no, we'll do that later. Oh, come on. Go away. You've moved often enough. OK. Let's continue the same point here. All right. Everybody sees this rectangle. And what's-- the rectangle is-- what are those black things?

STUDENT: Three-quarter circles.

JEREMY WOLFE: Three-quarter circles, yes. The literalists have figured out that these are Pac-Man or three-quarter circles or something like that. But you didn't really see that when it came up. You said, oh, it's a rectangle sitting on top of four circles of some variety. And, in fact, you still see it as a rectangle sitting on top of four circles.

In fact, you're probably reasonably convinced that you can see the contour, going, mmm, that's weird. I'm reasonably convinced that there's an interesting artifact on the screen that's creating contours. I don't know what that's about. It looks better-- it looks less bogus over there.

You can probably see the whole rectangle. The white-on-white borders are not there. There simply is no physical contour there. There may be here because the projector is doing something mutant, but there's certainly no contour there even though the rectangle at the center looks somewhat brighter.

Again, what you're doing is making a guess about what is it that actually created that image that's landing on my retina now? It could be a little conference of Pac people, right? Four little three-quarter circles who got together to talk to each other. But that doesn't seem the most likely possibility. What seems more likely here is that it's a white rectangle sitting on top of four black circles. And you end up completing that contour-- or this circle for that matter.

What are you doing here? Oh, you don't need to have fancy computer graphics to do this. One of the advantages of this particular-- of the material in this particular lecture is that it provides great material for doodling in other lectures. So if you like subjective contours, you can make your own.

And so how does that look? Got a subjective contour there? That's some sort of-- yeah, it's not perfectly circular. That's OK and-- but what seems to happen is that you generate a hypothesis that says-- the contours-- the whole problem I got here is the contours don't just end in the world, they tend to continue.

Well, if this guy is continuing, well, what happened here? Well, maybe it ran into another edge and it's being hidden. Well, all else being equal, it probably ran into an edge that's orthogonal to the direction it's going. Let's guess that.

And if we guess a whole bunch of little orthogonal edges, we're back to that earlier demonstration with a bunch of little line segments. I can tie those little bits together. They make a kind of a circle thing. So I end up seeing that imaginary circle. And in fact, I'm going to start filling in the contour all the way around.

That suggests, by the way that if I was to tilt all these lines a little bit off of straight radial so that the virtual line segment was not forming a nice, neat circle, that the impression of a subjective circle would get weaker, and you can decide whether or not that's true here.

So see this circle? Now the question is, is this version will give you three choices here. Stronger, weaker, or about the same? How many vote that this one is stronger than the previous one? How many vote for just about the same? How many vote for weaker? Just about the same. That's a boring demo, I'll have to change that next year because it didn't work well enough.

OK. Another example. This is-- if you're following along on the notes, we have now gotten to these so-called Craik-O'Brien-Cornsweet illusion, named after Craik, O'Brien, and Cornsweet. We're continuing the edge business, but now, what I want to do is tie that into that topic that I mentioned early in the lecture about how it is that-- what you're interested in is things that surface properties in the world and you're not interested in lighting. Lighting is very boring.

So-- well, what do you-- what do you see here? Some bold soul describe this complicated image. Boy. Slow group. Hmm? The gray thing. Two gray things. OK. All right. And if you had to-- one's darker. Oh boy, this is-- I'm going to bring my pliers next time this is like pulling teeth. Yes?

STUDENT: It's-- like the way I see it, it's like-- seem like a top of a pyramid and it's like shading on the other side--

JEREMY WOLFE: Like the top of a-- oh, oh, oh yeah, OK. So the light's coming from the left or something like that. OK, yeah, that's a complicated inference about these things. OK. But what you don't particularly see is this. If I take the edge out, if I take that edge away from the middle here, what you discover is that the whole thing is the same gray. That's not obvious.

If you were to draw the luminance profile, drag a photodetector across this thing, what you would get is something like this. Which side is bright? OK, that side is bright. So it rises, then drops across the edge, and then rises back like this. So that if you take out the actual edge, it's equal on the two sides.

So, that's kind of weird. Why does it look like-- now remember-- yeah, yeah. Here, here, we'll-- they didn't believe me. They thought I was doing something evil here. Whoops. There we go. Look, that's why I had this thing slowly sneak up, so you could-- well, semi-slowly. Fastly sneak up.

What's going on here is that the visual system knows something about edges and about lighting. Edges in the world tend to be fairly abrupt. Lighting changes-- so it's bright here and dimmer over here. Lighting changes tend to be fairly gradual.

So if what I'm interested in is seeing what's on the surface as opposed to seeing what's-- the product of boring light and shade variations, what I might want to do is to look for abrupt changes and to, in effect, suppress gradual changes, and that's what's going on here.

Now, you can also do it in a-- we might as well do an entertaining second-order effect here. Stare at-- remember that negative afterimage thing? Look at red, you see green and stuff like that? What I should be able to do here is produce a negative version of this effect where you'll end up seeing this side is dark and this side is light even though I'm not going to change anything over here or over here.

So stare at the center-- stare at that center line. Right stare rigorously at the center line, and keep staring. And then when I do that. So this is two illusions concatenated on top of it. Do it again, was that? Do it again? This is-- all right. For-- whoops. Whoop, gotta go back this way. For the people incapable of following instruction the first time, try it again. So stare at the center. And hold your fixation there.

Actually, the people who got it the first time, keep staring at the center. But the people who got it the first time can try something trickier, which is move your fixation a little bit and you'll change the proportion that looks light or dark. Isn't that fun? You should understand why that works. If you don't, write yourself a little note on your paper saying, I don't understand why that works, and figure it out.

All right, so, the important thing is that what you're trying to do here is you're trying to get rid of information about the light. You don't care about the light levels. What you care about is what's going on in the world.

Now Ted Adelson in the Brain and Cogs Department here has exploited this fact brilliantly in a variety of gorgeous demos, one of which is this. It is extremely difficult, even if you've seen this before, to convince yourself that the gray levels of A and B are identical, which they are.

In fact, it's so difficult that even if I stick a bar across it, that's clearly the same thing, it's still hard to see that as identical. Your brain wants to do all sorts of things to deny that possibility. What's going on here?

What's going on is that that virtual cylinder is casting a virtual shadow that you are busy discounting. You're saying, I don't care about the shadow. There's a cylinder, there's a checkerboard. I know about checkerboards. Checkerboards go dark square, light square. I can figure this out. That means that B is a light square because it's like surrounded by dark squares and A is a dark square because it's surrounded by light squares.

And it's a particularly lovely example of, among other things, this ability to get rid of information about the illumination. This isn't to say that what you do is somehow just run some sort of "throw away the light source information and don't do anything with it." Whoops, get back there. Yeah, what's it say?

STUDENTS: Cow.

JEREMY WOLFE: Cow. OK. They're only white and black on that screen. Look at the C and ask where that contour is coming from. That contour, particularly the outside of that C, has extremely little support in the image. What you are doing is making an inference-- this time, you're using the shadow information.

You don't want to see the shadow-- how many people see it as not only as cow, but as an embossed word "cow" sticking out a little bit? The reason you're seeing the cow at all is you're assuming that those black things are shadows. If those black things are shadows, it follows the light is coming from the upper-left. If the light is coming from the upper-left, well, we can go and figure out what shape object must have been producing those shadows, and the answer spells "cow."

But if you were to take a look at-- the C is the clearest-- well, I don't know, the O is pretty good, too. And so is the W for that matter. Anyway, look at any of those guys and look at the shapes of the black bits, which is the only thing that stands out from the background, of course. There's nothing else there except the black bits on a white background.

None of those bits say C or O or W. It's a construction based on the assumption that the black bits are shadows. You use that information all the time to do things like see faces. These are so-called Mooney faces, named after a guy named Mooney, where you can see-- you can tell me a lot about the curvature of these faces even though, again, there's nothing on the screen except for black regions and white regions, none of which are themselves particularly face-shaped.

I mean, look at the eye. Are any of those eyes actually shaped-- look at the guy on the right. I mean, his eye-- he's only got one, apparently. The other one is in full shadow. That eye looks like, I don't know, a mutant bunny or something. If I just presented-- I should do this.

If I just presented the eye piece on the guy on the right, that black blob that's defining his eye, if I just presented that in isolation, nobody would say, oh yeah, sure, that's an eye. They'd all be saying dalmatian dog before mutant bunny this time or something.

And it relies on these assumptions about shadows. That's what you're doing here. It survives inversion reasonably, well at least in-- but those don't look like faces very much. What went wrong? Well, you might think, well, what went wrong is I know that shadows aren't red and blue, but that's actually not the problem. Here, they look pretty good. Those faces look OK.

But-- whoops. These faces look lousy. Why do they look lousy? [MUMBLING] Somebody raise a hand. There's some-- yeah, there's a theory.

STUDENT: The shadows are lighter than--

JEREMY WOLFE: Yes. If you make the shadow regions lighter than the lit regions, the brain says, that's not a shadow. I don't care about shadows, I don't want to see shadows as entities in their own right particularly most of the time.

But I'll tell you one thing I know about shadows, shadows are darker than the other stuff. If the shadows are lighter than the other stuff, it's not a shadow. It's something else and something weird, so this doesn't work. But that works. Oh, I suppose the fact that it's set on this one shadow must be darker might have tipped some people off.

OK, let's see here. I think what I will do-- this makes a very natural break point where it says Mooney faces, and we'll go on to this question about making the best guess you can make in the context of 3D-- going from 2D to 3D. But before we go on to that, let's take our brief "stretch your limbs" kind of break here.

Oh, this is-- you're going to finally get yourself paid, is that the idea?

STUDENT: Yeah.

JEREMY WOLFE: Is that a pen or a pencil?

STUDENT: Pencil. Do you have a pen?

JEREMY WOLFE: I probably have a pen. I would try that side if they're not over on this side.

OK, let us gather back together here. By the way, it looks like it's getting to be that time in the term where people are abusing their natural sleep mechanism that I will talk about later on in the term. But looking around at this crowd, I would say that you're not getting the seven to eight you need.

Or, if you are, maybe you really need 10 and you're catching the extra two here. But let me-- there are lots of ways-- so I've been talking about little-- almost like atomic, exact, small-scale examples of these inferences that you make. And now what I want to do is head for the larger picture of how you make an inference about the whole scene. I'm not going to get all the way there.

And I also can't really-- there are many realms I could talk about this in. So I'm going to talk about it in one restricted area, which is this question of going from a 2D image to 3D inferences about the world, something that you do automatically all of the time, and I want to explain a bit about how you do it.

You will see on your handout that this is very useless blank region that says, 1, 2, 3, 4, 5. I'm going to go through a series of depth cues probably more than five of them, and that's what's supposed to go in there. There are lots of different sources of information that you use to go from the 2D world to the 3D world, many of them seen here in this lovely piece of Renaissance art that we will come back to.

But, here is a much more boring piece of art. What do you see? Yeah, a circle, square, and a diamond. And then there's the clever person who's trying to figure out how to describe them as Pac-Man before, but these aren't-- but you see a circle, a square, and a diamond. You don't have any serious difficulty inferring that--

STUDENT: A triangle?

JEREMY WOLFE: A triangle, sorry. You don't see this. For present purposes, the important point here is you can also tell me their depth order. It's, in some sense, so obvious that you never think about it, but it's a very important source of information about depth order that you get simply because you know that solid objects occlude each other.

So you firmly believe that I am standing in front of the screen. I could have-- well, no, I couldn't have. It would be theoretically possible that I have suddenly cut a cunningly wolf-shaped hole in the screen and I'm actually like-- I'm, A, very large; and B, standing over towards East Campus somewhere and you're looking at me through this set of holes. No.

[LAUGHTER]

You don't-- you automatically leap to the assumption that if it looks like it's-- if A looks like it's occluding B, A is in front of B. Oh, my bunnies-- I don't know what happened to them, they got pixelated. All right. Which bunnies are closer to you? The big bunnies are closer to you. Right.

Why do you think the big bunnies are closer to you? Because you're making an inference that bunnies are roughly bunny-sized, and that if one of these bunnies-- what would it be-- in the same way that I am currently making the assumption that you guys are all more or less people-sized. If I did not make that assumption, I would come to some very odd inferences about the current view that I am looking at.

So, the people in the front row-- there's a person in the front row. Her head is about 2 degrees of visual angle. Remember, 360 degrees around the head-- around my head. Each degree is about my thumb, so her head takes up about 2 degrees. And, all right, let's see. There's this guy in the cheap seats back there. His head is only about half-a-degree.

I could make the assumption that he's a pin head, right? And a guy with a real small head sitting out in the deep back out there. Or actually, I wouldn't make the assumption that he was sitting in the back. He's a pinhead sitting at the same distance as large head woman here in the front.

But that's dumb. Your visual system knows that's dumb. Your visual system knows people are roughly people-sized-- not exactly people-sized, but roughly people-sized. And if I see a bunch of small things there and a bunch of big things here, hey, gee, odds are that this is closer than that.

And that's part of what's giving me my current inference that I'm looking at a tilted plane of people in purple seats is this information about size. If I organize the bunnies the way you guys are organized, I get a much clearer sensation of depth from this texture gradient. So now, you should be able to see a tilted rabbit plane. Even though you know objectively it's just sitting flat on the screen, it looks tilted. Was that a hand up there? That was a hand.

STUDENT: I tend to see this sort of tilted thing like a [INAUDIBLE], like some of them being smaller towards the front, or near the front.

STUDENT: Yeah.

JEREMY WOLFE: So-- I'm sorry, I missed that.

STUDENT: So in your previous image--

JEREMY WOLFE: My previous image, we can do that. Whoop.

STUDENT: Yeah. I tend to see just some of the bunny being smaller and more in front [INAUDIBLE]

JEREMY WOLFE: Yeah, that's another possibility. It could be that-- well, I mean, in fact, it is just a flat image-- you actually-- you see, you sound like you're getting a hybrid of a tilted plane with bunnies of a range of different sizes. Actually, we can see that combination here.

So we got a whole bunch of big bunnies and two little bunnies. Which is the smallest bunny in this display? The bottom-right bunny. These guys are identical in size. It's a very minimal display. There's a much more vivid version of this illusion in the book, as I recall. It's a very minimal version of the illusion. Because you assume, if you are--

[LAUGHTER]

All right, so I still can't draw. If I am looking at two bunnies-- if I'm looking at a ground plane, if I'm looking at a ground plane, closer is also lower in the visual field. So you make an automatic assumption-- that was what was giving this woman the notion of a faintly-tilted plane in the first bunny example. You make the assumption that the bottom of the image is closer to you than the top of the image.

Well, if the bottom of the image-- let's go back here. Whoops, not there. If the bottom of the image is closer than the top and these two bunny images are the same size, if this guy is closer, it must also be really small. Remember what I was talking--

So suppose I took the guy from the back row here, his whole upper body fills the top joint of my thumb in visual angle terms. If I moved that image to the front and sat him in a seat here, he would-- he'd be about the size of this woman's upper arm. And I would think, that's a really small guy. The whole guy fits on-- she could wear him on her-- instead of wearing your heart on your sleeve, you can wear the whole guy on your sleeve.

[LAUGHTER]

So that's what's going on here. You're making the assumption that small bunny 1 is closer than small bunny 2. They're the same image size. You, therefore, infer that out in the world this must be a really small bunny and that one's just a reasonably small bunny.

Now the bunny part turns out to be not that critical. Here, you can also see a nice texture-- a nice plane going off into the distance with objects that are clearly not meaningful. Big stuff front and low, small stuff high in the image, and you simply infer that this is close and the stuff up there is far and you get a nice impression of a tilted plane.

This is one of the cues that is interestingly variable depending on where you're from. So the atmosphere scatters light. Particularly, water in the atmosphere scatters light. That's why the sky is blue. The result is that objects that are far away tend to be both hazier and bluer, something that you can see in all sorts of works of art. Go to the museum and you see artists taking advantage of this right, left, and center.

And you can probably get, even in my pathetically reduced version of it, you probably get a sensation of depth here that you don't get here. The geographic aspect of it-- anybody here from Arizona? Doesn't work well in Arizona.

I know this because I went to Tucson-- I went to a meeting in Tucson, and it was boring, so I went out for a walk. And I saw this hill, and I said to the guy who was at the street corner with me, how long would it take me to walk to that hill? And he said, three days, four days?

[LAUGHTER]

It's like 50 miles away and it's like 6000 feet high, but it was extremely crisp and there's no water in Arizona. I don't know why people live there. It's like hot all the time and there's no water, and-- but anyway. And aerial perspective cues don't work, so you're not only are you thirsty, but you're short one depth cue.

Anyway, this works much better in a humid setting than in a non-humid setting, but the point is that, again, you know about-- this is your visual system making use of the physics of the situation in order to infer something about the depth of the situation.

You also know about the geometry of the world. So this is an extremely limited picture. If I say, oh, this is a highway going off to infinity somewhere in Arizona or something like that, yeah, that's not a great, vivid picture, but you can believe that because you know implicitly that parallel lines in the world, if they're in depth, will look like they are converging towards a vanishing point somewhere.

Now it is sometimes claimed that-- this is known as linear perspective. That linear perspective was discovered by artists during the Renaissance. That's only sort of true. What happened in the Renaissance was that they made this knowledge explicit and became able to use it, for instance, to make their artworks.

But your cat and your lizard and stuff know about linear perspective, they just know it implicitly the same way they knew about aerial-- the aerial perspective and size cues and occlusion cues and things like that. You didn't have-- it wasn't that we woke up one day in Renaissance Italy and suddenly we could use this depth cue. It's what we figured out was how to paint with this depth cue. And you can do all sorts of amusing things with this depth cue.

So for instance, lines look more or less the same size. I just drew them with a chalk. They are about the same-- yeah, yeah, OK, good. Let me see if I can change that here. The inelegant way to-- the sub-PowerPoint demo. Whoops. It's not a great ruler.

All right. Even though we are using crude materials, let's do a forced choice vote here. If you now-- you know they're at the same size, so it's boring to ask if they're the same size. But if you had to vote bigger or smaller, how many people would vote that the bottom line now looks bigger? How many would vote at the bottom line now looks smaller? Well, I guess that worked, cheap chalk and all.

This is an illusion known as the Ponzo illusion. There are a number of ways to account for it. But one of the intuitively appealing ones, at least, is to say, what this is doing, even though you're not particularly seeing it as a depth cue, is it's telling the chunks of your brain that are trying to figure out 3D.

Hey, I see these two parallel-- well, they're not parallel lines. I see these two converging lines. If they're parallel lines in the world, they must be going off into depth. If they're going off into depth, then this thing is further away than this one. Well, if this one's further away and they're the same size on my retina, well, this must be bigger. If that's not intuitively obvious to you, think about this as train tracks.

So here are train tracks going off into the distance. And ask yourself, which maiden here tied to the tracks is in more distress? It's obvious that this must be the bigger person if we interpret this as train tracks going off into the distance even though we know that they're essentially the same size. So you're getting the results of the inference are then influencing what you see-- they're influencing other inferences that you make about the image.

Now you can exploit these rules of linear perspective in much more elaborate fashion than that. And the great master of that game is Escher. Here is one of Escher's pictures. What you want to do-- again, I think the image looks rather sharper up on these side guys, but ask yourself where the vanishing point is in this image.

So look at the first floor of that structure. And it's pretty clear that those railings are converging to a vanishing point off to the right somewhere. Well, now look at the top floor. Well, that's converging to a vanishing point off to the left somewhere. And you say-- and then when you try to put the whole thing together, you've got a structure that doesn't quite hang together quite right.

This tells you a couple of things. Thing 1 it tells you is Escher was a very clever draftsman. Thing 2 that it tells you is that you do a lot of these calculations about perspective very locally. What you do is you say-- you're attending, let's say, to the lower floor there, and you say, yeah, this all adds up, it makes sense. and then you direct your attention to the upper floor, and it all adds up, it makes sense. It's only when you try to combine all of that across the whole image that you realize that the whole image somehow doesn't quite make sense.

And that's how Escher can do things like have staircases that always go up and waterfalls that fall apparently in an infinite loop and things like that. Grab yourself your favorite Escher website and/or your favorite Escher book and you can watch him manipulating these depth cues endlessly. It's great entertainment.

So here-- now what I want to do is to bring together the themes of the lecture at this point in a single demo. At the moment, that doesn't look like much of nothing except that-- well, I don't know, what does it look like? Cubes, all right. Oh, cubes, that's interesting, because-- why is it-- I don't see no cubes. Where is that inference coming from?

Well, you're getting those-- what you really have is a bunch of Y-junctions. And you know about Y-junctions, those are probably corners, and they look like they might be cubey corners. But what I'm going to do is rotate each of those. Now this is a cool stimulus for a variety of reasons.

First of all, now you're really seeing a cube, right? No problem. Second of all, there are two cubes, right? There's the cube with its face pointing-- here, let's try this. Yeah, there's this-- that face pointing down and to the right, or there's-- whoops, come back. Here, mousey, mousey. Thank you. There's that face pointing up and to the left.

So you've got two cubes. This is an ambiguous, bistable figure. It's known as a Necker cube, just that-- well, we can quickly draw one of those. Endless fun for doodling again. You can make yourself ambiguous figures instantly. That figure by itself is known as a Necker cube after a guy named Necker.

So, you're inferring the cube. You're inferring one of two cubes. The cube isn't particularly there. The black stuff is-- all you're seeing is the vertices of this cube, but you're still managing to infer the rest of the cube. You can probably see the lines of the cube in the black region, right?

STUDENT: Yeah.

JEREMY WOLFE: In fact, you can see the intersection. If you strayed up from here, you can see the intersection of two lines that aren't there you're managing to create. And I can make those lines go away.

Now take a look at that cube-- at the figure and imagine that what you're looking at is a cube-- a wireframe cube that's behind a sheet of black Swiss cheese. You're looking at it through holes. Can you get it to go back there? If you get it to go back there and you can hold it back there, you probably noticed that the subjective contours pretty much disappear on you. Why is that?

Well, if it's behind, there's no reason that you should be seeing those contours. They would be invisible. And so the invisible contours become invisible. Now if you bring the cube back in front in your perception, you'll see, oh yeah, now if that cube is floating in front, I ought to be able to see the whole cube, and now I can see the subjective contours.

So you can make the subjective contours contingent on whether or not-- on which particular interpretation you care to give to the image. So I think this illustrates very nicely the notion that what you are seeing is your current hypothesis about what might be generating the image on the screen.

The other thing that-- another thing that it points out is that you are only willing to entertain a limited set of hypotheses. It is extremely hard to look at this and see it as however many it is-- eight little disks with chicken feet in them. With little Y's in them of some. It's very, very hard to see that even though that is a perfectly consistent with the image hypothesis.

You are out there trying to make a guess about the world and you're not willing to entertain all of them. At least most of them are immediately relegated to the realm of the very unlikely. A few of them-- and normally out in the world, what happens is that a single hypothesis immediately pops to the fore and you accept it.

In weird situations like this, we can get you to entertain a few of them, but it is very-- you immediately narrow down the realm of possibilities to a few, not to an infinity even though there is an infinite number of possible ways to generate this thing.

Shadows, I've already shown, are a depth cue. The reason for putting this nice piece of Renaissance art up there is to point out that while shadows are a depth cue, you're not actually terribly picky about the physics of the situation, at least not the global physics of the situation.

So where is the sun here? This is an outdoor scene. Where is the sun coming from? Well, if you look at the people in the lower-left-- or the people on the ground plane there, it's pretty clear that the sun must be down and to the left somewhere. Well, look at the shadow underneath that portico. Well, I don't know, the sun must be up and to the right there somewhere. There's no consistent source of illumination in this image.

But you're perfectly willing to accept-- but you're perfectly willing to use the shadow information to give you depth information. It's giving it to you locally. The fact that it doesn't add up globally doesn't bother you, and it doesn't even bother you to the extent that it bothers you in the Escher picture where the building was actually impossible. Here, you don't recognize the impossibility at all unless it's pointed out to you directly.

And finally, I should add to the list three more that are rather hard to demonstrate I put up just as PowerPoint slides. When people think about depth perception, if they think about depth perception, it's stereopsis, binocular vision, that they typically think of.

Your two eyes are in two different places in your head for most of us. If you blink from eye to eye or just cover your one eye after the other-- it's actually better if you hold one finger out in front of you and blink from eye to eye, you'll see that the image in the two eyes is not the same.

The difference in those two images is highly geometrically regular, and you make use of that regularity as a depth cue. It's so-called binocular disparity, very useful depth cue. It's what's giving you the *Magic Eye* demos that you get-- those posters that if you cross your eyes just right, they jump out in depth? Those still around? OK.

It's a very useful depth cue. It's a little on the overrated side because it's a lot of fun to study. People who think that binocular vision is the be all and end all of depth cues should cover one eye and ask if the world suddenly looks very flat. It looks a little flatter, but I can still perfectly well tell who's in front of whom.

On the other hand, if you want to see what stereo is doing for you on a beautiful day like today, go outside, lie under a tree, close one eye and look up into the branches, and try to figure out which twigs are in front of other twigs. And you'll have a very hard time doing it. Open the eye and the whole thing will jump out at you in depth. That's the information that stereo is giving you.

You don't have to do it with stereo. Motion parallax, to jump to the bottom of that list, is a similar geometric clue. If I'm here and then I'm here, the image changes in a geometrically regular way. You guys are sliding around on my retina in such a way that these guys are moving more than you guys out in the back on my retina.

And I know that, again, implicitly, like I know linear parallax, and I can use that to infer depth. Try this under the tree and you can see the same thing. Close one eye, look up into the branches, the branches look relatively flat. Now, rather than opening this eye, just move your head back and forth, and the tree will jump out at you in depth. It's actually quite striking, you should try this sometime. I don't know if anybody ever does take me up on this suggestion. So if you actually try it, let me know so that I know that somebody actually tried it.

Oh, and vergence is another one of these geometrical cues. Hold your finger out in front of you, look at your finger. Now move the finger towards you, holding it as a single finger as long as you can. So I can watch you go cross-eyed. What you are doing is converging your eyes. And if you could move your finger further out, you would be diverging your eyes.

Well, what you can think of that as doing is taking a pair of sticks and pointing them at the object. This is the axes of your-- your visual axes, in a sense, you're pointing at the object there. And the angle formed by those sticks is narrower if you're looking far away than it is if you're looking close up. And you have a fairly impoverished ability to use that as a depth cue, too.

If you were a chameleon, you'd be much better at this. Chameleons have eyes that move independently and are very sensitive to the angle that their eyes are pointing. And in fact, that's how they go-- you've seen the Animal Planet kind of videos where the chameleon's tongue goes out the length of its body and it grabs a fly or something like that? How does it know where to fly is?

Well, it knows by measuring the angle of its eyes. How do we know that? Well, we know that because somebody went and put glasses on a chameleon that diverged the eye so that in order to point its eyes at the fly, it had the angle wrong, basically. So you put a fly on a popsicle stick, you put the glasses on the chameleon, and then you film the chameleon's tongue, and the chameleon keeps missing.

[LAUGHTER]

If I put those prisms on your eyes, you will adapt and you will eventually be able to catch the fly again, should you be so inclined. The chameleon turns out to be a less adaptable creature than you and will not adapt. Similarly, you can do the same kind of game with chickens. Put a pair of prisms on your eyes that divert everything off, say, 15 degrees to the left, and then if I tell you pick this up, you'll reach 15 degrees in the wrong direction, but eventually you'll learn.

Put the same prisms on a chicken, put some grain down, here are the grains here, the chicken sees it over there, chicken's pecking over there, chicken's not getting any grain. Chicken will do that forever.

[LAUGHTER]

And will apparently not learn to get it right. OK. Now what I want to-- oh, good. I left myself with enough time to talk about inferences in a more global sense of combining information from across the senses. So you've got this job to try to figure out what's going on. I've been talking about doing that specifically within the visual system.

But your gain-- you're collecting information from multiple sources and taking whatever the best information is. So if I show you a movie, you get captured by the visual information and you're perfectly happy to see the word-- or to hear the words coming out of the mouth of the guy on the screen, even though it's coming out of some speaker on the side.

And it doesn't matter if they've got fancy Dolby stereo or something like that. Use a cheap, simple speaker sitting off to the side and you'll still hear it as coming out of the guy's mouth if you're watching it. So you're combining information from multiple senses.

The particular example I thought I would discuss with you is the ever-pleasant example of motion sickness. So when I first came to graduate school, my lab was doing research for NASA on the effects of looking at large fields that were rotating-- this is Omnitheater stuff. If you look at a whole field that's rotating counterclockwise around your line of sight, you feel like you're rotating clockwise.

Meantime, guys up at Brandeis who we were collaborating with were doing the same sort of thing, but their dependent measure was how long it took you to throw up. Fortunately, that was not my introduction to graduate school. But it will make you sick.

How many-- well, we might as well collect some data here. How many people have ever been motion sick here? OK. What made you sick?

STUDENT: Motion?

JEREMY WOLFE: Motion. Yes, thank you.

[LAUGHTER]

Could we get a little more specific while keeping within the realm of good taste here? Everybody got-- OK, yeah?

STUDENT: Reading on a bumpy bus.

JEREMY WOLFE: Oh, reading on a bumpy bus. So this is-- that's one good example. Anybody got another good-- yeah, OK. Maybe yes, no?

STUDENT: [INAUDIBLE]

JEREMY WOLFE: Oh, OK. She was going to give the answer.

STUDENT: Motion sickness feels like watching something, but not actually moving yourself.

JEREMY WOLFE: So the answer that she's giving-- well, all right, we'll take one more here.

STUDENT: Spinning around in a circle.

JEREMY WOLFE: Oh, spinning around in the circle. Did you get sick while you were spinning around in the circle or after you stopped?

STUDENT: Stopped.

JEREMY WOLFE: After you stopped-- and when you stopped, you felt like you were spinning around in the other direction. So when you're spinning yourself around in a circle-- you have in your inner ears these tubes filled with fluid. And-- a quick drawing here. Quick bit of vestibular physiology, you need a nice fat piece of chalk.

So-- oh, it's yellow. You've got these tubes inside your ears that are filled with a fluid. And inside a little space in there are these hairs. If you bend the hairs, it sends a signal off to your brain. That's the transducer, the equivalent of the photoreceptors in the eyes.

If you rotate your head-- well, you can imagine-- imagine taking a bucket and starting to move. The fluid stays behind for a little while and sloshes around? That's why it's hard to carry buckets of liquid around if they're too full. So if you rotate your head, the fluid tends to stay put, and it moves over the hairs, bending them, and telling you that you've moved your head. That's what's telling you about this motion.

Well, if you spin around for a while, and the fluid in here eventually catches up with you and starts moving, then you stop, and the fluid keeps going, and you say, oh, no, no, no.

The other way you can do this, by the way-- so this very carefully calibrated system, turns out that alcohol is lighter than this fluid. If you drink, the reason you get dizzy is not because you've pickled your brain, which is also true, but because you've diluted this fluid with alcohol and this stuff is uncalibrated, and now you move your head a little, and the brain says, oh, man, we just moved a lot, man! Oh, don't move that head. No, no don't do it. So that's--

[LAUGHTER]

What she was pointing out is that what turns out to be the great stimulus for making yourself motion sick is a mismatch between the information that, in particular, this vestibular system, this balance system and your eyes are giving you. So the example given here of reading on the bus is a marvelous example because what you're doing is, in order to read, you're holding this thing so it's a relatively stable visual stimulus. Nothing's happening here.

But your vestibular system is saying lots of stuff is happening, and your digestive system is busy saying bloop! The same thing happens on a plane. The problem on a plane is that when the plane bounces up and down-- I mean, short of catastrophic bouncing up and down. When the plane hits turbulence and is bouncing up and down, what do you see? Nothing. There's nothing happening visually. What do you feel? You feel ba-doop, ba-doop, and and then bloop.

And the interesting question is, why do you get sick? It's fine to say that the mismatch of visual and vestibular information turns out to be nauseating, but why should it make you sick? The answer is, it's another inference, but the question is, what's the-- so the next question, then, is, what's the inference? What are you guessing? Anybody got a-- you're guessing that the airline food was lousy. Yeah?

STUDENT: That you're poisoned.

JEREMY WOLFE: Yeah. You're guessing that you were poisoned. Why are you guessing that you were poisoned?

STUDENT: Because your vision is strange.

JEREMY WOLFE: No, it's not-- it's the mismatch that's critical. If you just make your vision strange-- I could show you-- I could show you weird stuff you hadn't seen before, and you wouldn't all like throw up. And I can also bounce you around, and until I do fairly dramatic stuff, you won't throw up.

But, if I mismatch what you think-- what your visual system is telling you about your body and what your vestibular system is telling you, this is a protection of sorts against neurotoxins. How do poisons work? Well, some of them work by attacking your nervous system. A lot of them work by attacking your nervous system.

How are you, the owner of the nervous system, going to know? You're going to think, oh, I can't integrate anymore. Now I can no longer remember that I love my mother. The sorts of things that immediately present themselves to you as you're being poisoned are, your senses are coming unglued from each other.

And so if your eyes are saying boingy, boingy, boing and your vestibular system is saying this, your body infers that you have been poisoned, and a useful idea if you've been poisoned is to get rid of whatever you just ate. So that's what's-- and why does this happen on airplanes and stuff like that?

You weren't built to be flying around at 30,000 feet bouncing around in the clouds. It just wasn't what nature set you up to do, and so you get the unfortunate situation where, even though you haven't been poisoned, you get sick. OK. Enough of that cheery topic.