

The following content is provided under a Creative Commons license. Your support will help MIT OpenCourseWare continue to offer high quality educational resources for free. To make a donation or view additional materials from hundreds of MIT courses, visit MIT OpenCourseWare at ocw.mit.edu.

PROFESSOR: OK. Today, teaching equations and how to teach equations in a way that promotes long lasting learning and understanding. OK. So the first example-- I'm going to give you two choices for starting the example.

So this is example one for teaching equations. This is Hardy-Weinberg equilibrium. So here's one way you could start. So if a locust has n alleles and the organism is polyploid-- so that means it has, let's say, C copies of chromosomes.

OK. So there's n alleles. So the allele frequencies are P_1 through P_n . Then we're going to deduce the genotype frequencies.

So this is the multinomial middle coefficient C . Choose K_1, K_2, K_3 , all the way to K_n . And K_1 through K_n are the number of copies of each allele with K_1 plus blah, blah, blah, to K_n , all equal to C .

So you could actually state that and then prove it. That's option A. So now I'll give you option B for starting the same topic. So option B is from Hardy. So in Hardy's *Mathematician's Apology*, he writes, I have never done anything useful. No discovery of mine has made or is likely to make, directly or indirectly, for good or ill, the least difference to the amenity of the world.

That's a very interesting, very bold statement. And amazingly, for such a brilliant mathematician, it's wrong. And today, what we're going to do is we're going to look at an example of the discovery by Hardy-- the Hardy-Weinberg equilibrium-- that does make a difference to the amenity of the world and helps you actually understand genetics.

OK. So the option B is to quote Hardy and then point out the contradiction between

what he said and what we're going to do. Now, also, in option B, you can go a bit farther. You can say, well, why would Hardy have said something like that? And indeed, why would he have said something like that in his *Apology*?

Well, it's actually probably hard to appreciate on this side of the Atlantic, because since the Civil War, we've never had a war that just basically devastated entire countries on our own soil. Whereas in Europe, the memories of World War I are very strong.

So if anyone's from Europe, you know that. In every English chapel, there's names on the wall of all the people who died in the, quote, Great War. So World War I had a very strong effect on European society.

And one of the effects was on European science and people's attitude towards science. So poison gas was invented partly by German chemists. Hopper was one of them. And in the end, he committed suicide, partly maybe because of what he had done. No, sorry. His wife committed suicide. I forget if-- I should check that.

But basically, people in the family were so unhappy about what had happened just there that there was a suicide. And furthermore, science after that was considered-- World War I was considered the chemists' war. So there was a wish to distance oneself from those kinds of horrible effects.

And the quote from Hardy is actually a reflection of that. He himself was very strongly anti-war. He left Cambridge because Cambridge fired Bertrand Russell for protesting World War I. So Hardy left Cambridge for a professorship in Oxford and only came back basically 12 or 13 years later to Cambridge.

So that statement of his partly a wish that he's hoping nothing he's done has any effect on the world because in his mind, maybe most effects are bad. Well, but actually, he did have an effect. The effect is Hardy-Weinberg. Even if you discount everything you did in number theory, there's Hardy-Weinberg, and it has an effect and we're going to look at it.

So let me ask you rhetorical question. If you're a student, which would you find more

engaging, more inviting? This one or this one? Who votes for that one? OK. Who votes for that one?

Yeah. Now, why? So now think of-- so the question I'm going to ask you to think about just for a minute with one or two of your neighbors is why. So yeah, this is definitely more engaging. In terms of the principles we talked about last time, what makes this way of introducing it more engaging, more likely to bring students in, more likely to make them want to learn about Hardy-Weinberg.

OK. So find a neighbor or two we think about the principles behind option B.

OK. So I'll rudely interrupt you. And good thing I did the voice exercise so I can project all the way into the back. So what reasons for option B or against option A, which is the same thing?

Yes. Can you tell me your name?

AUDIENCE: Brian.

PROFESSOR: Brian. I'm going to try to learn people's names. So go ahead.

AUDIENCE: The reason against option A-- when you start out that way, you are activating the ability that students have to transfer material from the blackboard to their notes and then pass it to their brains.

PROFESSOR: Right-- through the wave. Yeah. I'm activating the ability of the students to basically take dictation. Or as someone said, teaching is an excellent way to transfer material from the notes of the teacher to the notes of the student without it passing through the minds of either.

So I'll call that not A. So that's the not symbol. Not A is-- so I'm basically asking to do dedication which doesn't necessarily lead to any kind of learning. They could, for example, copy all of that down but not really understand of it.

OK. Other-- yes, can you tell me your name?

AUDIENCE: Susannah.

PROFESSOR: Susannah.

AUDIENCE: This is right brain? So it's easier to remember and ingest it. And maybe even identify the theorems.

PROFESSOR: OK. Right. And so B is the story. I'll give you an example of how stories can be so powerful, and just even the word.

So a big commercial publisher-- this was several years ago-- they wanted me to write a freshman physics textbook. I had actually just put up a proposal on the web saying, we should write a freshman physics textbook that is based on the history of science.

And they saw it and they thought, oh, that's great. So they flew me from England to California to talk to them. And they liked it mostly. But they said, history-- so there was the word. They liked the idea as I talked about it, but the word history really frightened them.

And then, I'm still amazed-- I had this amazing insight of just two letters. I thought, oh, wait. I can actually explain to them what I mean. I said, well, actually, if you just take away that part, what I'm really talking about is that.

And it was interesting. As soon as they saw the word story and they saw that history doesn't have to be names, dates, facts-- I think that's what they were seeing it as-- they saw it as story, which is actually its origin in French or Latin-- [FRENCH].

They all of a sudden were totally convinced. They said, oh, yeah. That's exactly how we should do the textbook. For various reasons, basically because I wanted to be a freely licensed book, we didn't actually sign a contract. But it's an example of how powerful story is and how people who actually spend their lives thinking about teaching and reaching students-- in other words, this educational publisher knew the importance of story.

But if you present it as just dry history, it's not so interesting. So story-- yeah. So

let's see. Who haven't I heard from? I haven't heard from Adrian. Yeah, and then you're next.

AUDIENCE: So the fact that it's a contradiction creates some sort of tension from linear thought. And it's more of a exploratory.

PROFESSOR: OK. So there's a contradiction. And the contradiction creates some kind of tension. So the tension creates interest. So every good story needs some kind of tension. So the tension here, the contradiction, is that Hardy wanted to do nothing that could harm people, basically. And so he was like, OK, well, anything could be used, even if it's for good, for ill. So let me just back off from all of that and say I'm not going to do anything that has an effect, an application. I'm a pure mathematician.

His most famous book is *A Course of Pure Mathematics*. Well, maybe his most famous book is that book that I quoted from, *A Mathematician's Apology*. His most famous math book is probably *A Course of Pure Mathematics*.

And that's what he wanted to be known for. And that's in contradiction with the fact and no, even Hardy couldn't help an application-- an actually very, very common one. It's taught in every single introductory biology course, probably.

OK. So you need tension. Now, that's a general principle of learning there-- the idea of story and tension and paradox. So the one way that, as a undergraduate in physics, I learned a ton of physics, me and my friends were doing problem sets together. And a lot of problems were just grinding through math. So we didn't learn a hell of a lot from that.

But we were in the library late at night, the physics library, ordering pizza, and trying to do a problem set, and we just got to making up physics paradoxes-- perpetual motion machines. We'd invent perpetual motion machines and try to get the other person to figure out what was wrong with it. Sometimes, we didn't even know what was wrong with it and we'd both try to figure it out.

So from that kind of tension-- tension is almost, in a way, self teaching, because as

long as the tension's there, you know you're not at the end of the story. You know there's more to do.

So the same thing-- as long as there was perpetual motion going on and we hadn't found the reason, we knew we weren't done with the problem. We didn't have to ask a teacher to say, well, is this right, because it was basically self teaching. As long as the tension was there, we knew that we weren't finished.

So yeah-- another general principle, stories and tension. Let's see. There was a comment over-- yes. Tell me your name?

AUDIENCE: Wing Ho.

PROFESSOR: Wing Ho?

AUDIENCE: Yeah. So this along the line that the teacher should make themselves more human, because Hardy-Weinberg sounds like such a big name you would tattoo. That student may be driven to believe in the equation just by the matter of the fact that it's authority. But then making a story makes--

PROFESSOR: OK. So the teacher makes himself more human and makes one of the people who invented the Hardy-Weinberg equilibrium more human. So it's actually something much more easy for the students to connect to. Yes, can you tell me your name?

AUDIENCE: Gregor.

PROFESSOR: Gregor.

AUDIENCE: Yeah. So my problem is that two different options are very, very different. And the amount of legwork it is, it's maybe more tense to me, because you can discuss it without legwork, but I think also that option A is pretty interesting, because you can approach this biological problem from more of a qualitative approach.

So depending on what kind of learner you have and what kind of question you choose, A would be also very interesting. But at that point, with the background I have right now, it's very difficult to pursue that.

PROFESSOR: OK. So Gregor's point, which is a good one, is that B doesn't have the same quantitative depth as A. And actually, what I'm going to show you is a way to get to this starting from here and using the principles that we're talking about so that by the time you get here, or to something like this, it actually makes sense to the students-- so that you can have both.

So that's a promise. And hopefully, I'll deliver on it for you. But does that address what you're talking about?

AUDIENCE: Yeah.

PROFESSOR: --that B doesn't have the quantitative depth-- and that's true.

AUDIENCE: Actually, I mean, you can talk about one hour on B. But if you really want to deliver an idea or a concept, that is as quantitative as option A, then you have to get to A at some point.

PROFESSOR: So the point made is that if you want to get here, you do have to say this at some point. You can't just say the story for now. Well, it depends on the purpose of the class. If the purpose is the history of biology, maybe you'd continue with the story or the history of science and war.

But if you want students to be able to solve problems in genetics, maybe you need to go here. Or what we'll find is I'll continue with option B moving towards A. And I'm going to show how you can get there and still have the opening of B preserved.

OK so let's see. There was-- yes. Can you tell me your name?

AUDIENCE: Roger. I think one of the reasons we like B is because we see A all the time. So I don't know to what extent that's [INAUDIBLE].

PROFESSOR: OK. So B is new. Right. So against A, A is very familiar and common. So in fact, yeah, I chose this-- I didn't really make a straw man here. This is how a lot of things are introduced.

For example, in mathematics class, theorem proof-- so the theorem will be

introduced without any of the struggle, the wondering that led to the theorem. Why would anyone even care about such a thing? Or if they cared about it, why would they come up with something like that? How can you see that?

So yeah, A is seems very familiar and therefore, maybe not as interesting. Although, it may also be intrinsically less interesting for the other reasons. But I take your point, too, that it could be just familiarity. Yes. Can you tell me your name?

AUDIENCE: Yeah, I'm Meg.

PROFESSOR: Meg.

AUDIENCE: It seems that you could really incorporate the quantitative side into B if you wanted to by having the way the proof occurred [INAUDIBLE]. And then you wind up reversing the traditional order, where you have a proof and then a theorem, because the way it happened in reality and the way it happens when you're going to do research, and the way it happens to students when they're trying to work it out themselves. And so I feel like that's much more of a natural flow [INAUDIBLE].

PROFESSOR: OK. So your point, which is an excellent one as well, is that B doesn't preclude the quantitative. What you could do is you could talk about the history, start with the history of Hardy saying that, which is not chronological, because he said that in 1940. But that's OK. It doesn't have to be chronological just because it's history.

Go to 1940 and then backtrack to well, how did this problem actually come to Hardy's attention, which actually turns out to be quite an interesting story. And then talk about how they solved it and what was the publication history. And all the quantitative ideas would come out in that way.

And it would be almost backwards from the usual way. The usual way here is completely general. And that's probably not the way it was first figured out. And the advantage of telling about it that way is that you're preparing students themselves for a research career, because nobody comes up with theorems full blown like Athena out of the head of Zeus. You come to them from struggle and wondering.

Hm, I wonder.

Everything has some historical background to it. So it's called the, you could say, the genetic approach. So in biology, there's a slight, not misconception, but a saying that ontology recapitulates phylogeny-- in other words, that the organism, as it develops, say, in the womb, goes through all the evolutionary stages that it went through over the last 200 million years or whatever to become, say, a person. You're at first a fish and then maybe you're a monkey and then you're a person.

So roughly speaking-- it's not exactly true. But there's a lot of truth to that for learning ideas-- that you have to recapitulate the history of ideas to really understand where we are now, the ideas today. So a great example of that is Newton's Second Law of Motion, the idea that force and acceleration are connected.

So for thousands of years, millions of years, people thought force and velocity were connected. And it's actually force and acceleration. So actually, you can guide people to the understanding of force and acceleration being connected by showing them the history of how people thought it was force and velocity, because for those same reasons, those are the reasons that students will think it, too. So you're actually helping them overcome their misconceptions.

So the history actually, generally speaking, helps overcome misconceptions. So I'll put that here. And the history is actually quite interesting. I think it was Punnett from Punnett squares. He played cricket with Hardy and Hardy loved cricket. And he just asked him about this problem.

And Hardy said, oh. Yeah, no, it's just this, and sent off a paper to *Science* or *Nature*. So basically, it's because they were cricket playing colleagues in Cambridge, that's how a mathematician ended up interested in a biology problem. Punnett was a biologist.

OK. So any other reasons? Yes.

AUDIENCE: I'm going to play devil's advocate for a second here.

PROFESSOR: Sure. Can you tell me your name?

AUDIENCE: Paul.

PROFESSOR: Paul.

AUDIENCE: So when I had biochemistry the first time, the first month of the class, the teacher would tell the history about the people developing quantum mechanics. And that was the lectures. We didn't touch the material one bit. And because his philosophy was the material was in the book, you don't need to teach it. You just go home and read it.

But I remember being in that class-- and maybe that's because I normally don't see things taught like that-- but everyone was outraged, because we didn't feel we learned anything.

PROFESSOR: Right. OK. So the comment is that you were in an actual class that was purely taught about the history. People were outraged. They felt they weren't learning actual content.

And there's two answers to that. One is that actually, I don't recommend if, say, you're teaching a physics class, that you teach it purely historically. But to the extent that you do teach the history, you can actually teach the content through the history. So it's not that the history and the equations or the ideas are separate. It's that you can use the history as a means of making the equations come alive with a deeper understanding.

So it's possible that the teacher you're talking about actually didn't do that and just talked about the people who did this and did that. But he didn't really know what they did or why. So you're not forced to do that.

OK. So I'll give you an example. I'll continue with this in just a moment after I take any other questions, showing you how you can start with the history, and then you can show the content. You can actually show people exactly why that would be true.

And I'll show you that in one second.

OK. There was a question there. Yes. Sharon. Oh, sorry. Not Sharon. Can you tell me your name?

AUDIENCE: Me?

PROFESSOR: No, in front.

AUDIENCE: Cecilia.

PROFESSOR: Cecilia. Yes.

AUDIENCE: I was wondering that Option B-- it's like targeting another audience. I come to a class and I'm not very interested in the class. I need to be motivated here, Genetics or something. If I already signed up for genetics-- because if I were being in a class interested in that, I actually want to learn things. I want to know the rules. So I don't know if I want to be distracted by some story that is-- when you first said that, I was like, oh my god, if this is going to be some other big story, he's going to have the equation at the end. I'm not going to be understanding what's going on or whatever you say.

And I think that could be a little bit of a turn off more, because-- try to make it look easy. You have avoided writing the equation on the board.

PROFESSOR: Yeah, I'm going to do that next.

AUDIENCE: OK. And also, I guess there are two kinds of stories you can tell. I don't know what story you're going to tell exactly, but this seems to be removed from the actual equation. This is more Hardy's philosophy. It's not really--

PROFESSOR: It's not directly about that.

AUDIENCE: Something you did a second ago was more about the equation and more about the concepts involved with the equation.

PROFESSOR: That's true. So one question is how related should the stories be? And there's a lot

of freedom in that. I would say the answer is that you want it-- if it does this, if it creates interest, then it's already done something.

Now, your other point was that it depends on the audience. Suppose the audience is all graduate students in genetics or people who want a crash course in genetics. Maybe actually, they want to just know the formula. And they would actually be offended if you started telling them stuff that wasn't the formula, because they feel like they're time's being wasted-- a little bit like what you were saying.

So now, did you have a comment about that?

AUDIENCE: No.

PROFESSOR: OK. Then I'll come to you one second. So one comment about that is that actually, most of the students, at least in a big class in genetics, are going to be students who are-- for example, at MIT-- are there not because they love genetics and are going to continue in genetics. They're doing it because they're required to take introductory biology.

And your job as a teacher is to actually show them that this is a really fascinating subject. It's the same in physics. It's often taught as if the only students in a physics class were the physics majors.

Now, when I was an undergraduate in Stanford, there were 1,600 undergraduates and 12 physics majors every year. So 1% of the student body was physics majors. So they were actually teaching in a way, generally speaking, for 1% of the student body.

What about the other 99%? It was really important for them to reach them as well-- important, also, for the health and growth of the field, because the field needed people who, even if they weren't professional physicists, saw the value in the field and the value in physics. So generally speaking, it is actually a wise way to reach all of them.

Now, suppose you have people who just want this. The next thing I'm going to do

will show you a way of teaching to them that they'll actually learn this better, because I guarantee you, almost everyone, if you just tell them this and they take dictation, now they go away, and you say, OK, everybody, what did I write down? Most people can't reconstruct it. They'll say, well, was this a C or an N?

In fact, if you look in Wikipedia, the Wikipedia entry is incorrect. Actually, I think it has an N over here. It seems plausible, but if you actually understand the equation, you think, that can't possibly be right. And I think it has an N over here.

So they actually do not really have a command of the equation. So even for them, you want to not-- the story isn't the proof of it for them, but the continuation will be valuable for them. So again, I'll promise that.

Now, there was a question over there. Yes.

AUDIENCE: The story, you can basically use to-- and if you see you lose your students when you go on this quantitative approach, then it's maybe helpful to loosen up a little bit and basically bought you interest to see what's a big context.

And the other point was--

PROFESSOR: So you're first point was that you can add stories as needed.

AUDIENCE: Yeah. So depending on what it's like. And the other thing is when I make my comments related to the problem, I see it more from a various student point of view. So you point out that, depending on what the audience is, you teach differently.

PROFESSOR: Yeah, that's true. Yeah.

AUDIENCE: So when you give your examples--

PROFESSOR: Oh, I should tell you who the audience is.

AUDIENCE: --interesting to know at a certain point what kind of audience it was.

PROFESSOR: Yeah, OK. Good point.

AUDIENCE: So I'm undergrad. So I don't know how it is to teach first year university. I've never had students. So how much basics or how much story you have to put in compared to how much math. So it would be very helpful to see what our target is.

PROFESSOR: OK. Your comment is that it would be helpful when I'll give you an example like this to say, OK, for the particular audience, what would you do? Yeah. And I'll try to do that. That's a good point.

Now, there was another question-- behind Cecilia. Did you still have a comment?

AUDIENCE: Oh, I just wanted to mention that in choice A, there's a lot of jargon.

PROFESSOR: OK. So there's a lot of jargon in choice A, which makes it much harder to understand. You have almost-- for example, the multinomial coefficient-- you have to keep that in your head, as well as all these subscripts, keep it all in your hand and try to manipulate that object-- very difficult.

One of the points about that related is the chunking. So the chunking paper-- the idea is that these, for a student, each of these things is going to fill up one of the chunk slots, almost, because it's all new to them.

You've flooded the chunking system. You've overflowed it. And they can't actually manipulate this as one object anymore, because it's far too many chunks.

Now, you, as a professional in the field, are like, well, of course. It's a multinomial coefficient. What else could it be? For you, that's one chunk. For the student, that's is it C or N? K_1 ? K ? Why K ? What the hell is K here? Is it commas or spaces? Parentheses? What about brackets? I've never seen something like that. Shouldn't there be as many things up here as here?

So the way they look at it is completely different. Every symbol is almost a chunk. So this is just massively overflowing with chunking. So this is sort of related to the jargon. Yes.

AUDIENCE: Scott.

PROFESSOR: Scott.

AUDIENCE: I have a meta comment. In this example, if you don't know how you're actually going to complete these, it's created a lot of tension. I'm dying to know how you're going to do it.

PROFESSOR: OK. Fair enough. Would you be very offended if I gave everyone a break for a few minutes and then finished it?

AUDIENCE: OK--

PROFESSOR: OK. So finish your comment.

AUDIENCE: My question is in a situation like this where you have people who are making so many comments, and the other half's just dying to know, how do you know when to stop?

PROFESSOR: Good question. So how do I know when to stop? Because I don't want to exhaust people's patience, but I do want to take questions. And in this case, there's the American teaching phrase saved by the bell, which is that I don't have to make the decision too hard, because it's probably time for a break anyway.

But how do I make the decision? Partly, I listen to people's comments. For example, suppose I got yet another comment saying, well, we still don't know the equation. And I'm saying for the third time, oh, wait. I'll show you a way to get to the equation. I think, you know, it's probably time for me to do the equation and then I'll take more questions later.

So listen for the tension in people's voice, which is also a good reason to have everyone do the voice exercise first and free the tension so that if the tension creeps back in, you know it's something you've done.

OK. So I'll take more questions after the break. But it's 10:02 by that clock. 10:05, we'll start again. You can jump up and down, do jumping jacks or whatever it takes to get the blood flowing. I'll take a couple more questions and then I'll show you how to continue it to get towards that in a perceptive way.

OK. So let me, as promised, continue along the lines-- not continuing the story, but continuing that approach, the alternative mirror image approach of A, which is to say, well, what would I do after telling the story?

So first of all, I would try to fix some of the problems in just basically blasting people with that. So I try to make it as clear and unjargony as possible-- so as you notice, jargon up there.

So imagine a gene with two flavors-- sickle cell or not. And two chromosomes-- so just like people have two copies of each chromosome.

OK. So now, before I continue, what have I done just by doing that? Well, first of all, I've made something concrete. It's sickle cell or not. So right away, you can imagine it. So that's continuing the idea of a story. It's much easier to imagine a concrete situation. Either you have sickle cell anemia or you don't. Or you have a gene for it or a gene that doesn't cause it.

And the two chromosomes-- so rather than having C chromosomes, I'm just restricting it to two. Now, you might think, well, that's a terrible restriction. That's a specialization. Therefore, it's bad. No. Actually, it's good for that reason, because it makes it possible to understand, in the next step, the idea behind that equation.

And once you understand the idea, then that equation is not so mysterious. And furthermore, what's nice about two chromosomes? Yeah, it's a specialization, but--

AUDIENCE: Same two.

PROFESSOR: Yeah, we all have two. There's some plants that have more and some plants have less, I think. But people have two. So it's already interesting to us just for that reason.

OK. So then the question is how frequent are the combinations? So we want to answer how frequent are the three combinations sickle-- lowercase S is not sickle. So this is S. This is lowercase S.

What happened to the fourth combination? Why are there only three? Yeah. It's the same. S, S-- this and this are the same. OK So I would actually ask that, too. Any time there's something interesting-- it should have been four. Plausibly, it could have been four, but it's actually three. Why three? OK, it's because we're actually lumping these two together.

OK. So how frequent are those three combinations? Well, that's provided by this formula. P^2 , q^2 , $2pq$ -- so p and q are the frequencies of S and little S.

So this is the-- OK. So now, because you've asked them about the three versus four, this two makes sense, already somehow. But how could make sense of the p^2 , $2pq$, and q^2 ?

And so that's where I would continue with the following, which is this. So here is 0 to 1. And this is p . And that's q . So this is the frequency of sickle cell. This is the frequency of not sickle cell.

So right away, you have $p + q = 1$. I'll draw this up here too. So I make a square. And I have four regions.

So it's not magic. Why am I making a square? What about the problem tells me square? So I would ask the class that.

I'm looking for area. And why an area? Why two dimensions?

AUDIENCE: Because there are two chromosomes?

PROFESSOR: Because there are two chromosomes. So this is your C. The number of dimensions is eventually going to become this thing, so you can see how we're going to get there eventually.

So I'm looking for an area, because I have two chromosomes. So now, let's just look at these various areas. Here's a p^2 . What's this area? There's q by p . So that's a pq . That's also pq . And that's q^2 .

OK. Let's look at all our pieces. Oh, we have a p squared there. We have a q squared there. Oh, and pq , pq -- $2pq$. So in fact, this picture explains the entire set of frequencies here.

And if you see this picture, there's nothing really new to understand here. In fact, all this picture is showing is that-- it's a picture of p plus q squared equals 1. So if p plus q is 1, its square is also equal to 1. And you break this up into the four terms, group two of them together, and you get three terms, three different kinds of terms. You get three different frequencies.

So now, this square is actually very easy to use. Suppose someone tells you that sickle cell anemia, people who have both genes, are 1% in the population. Well, what fraction of the population has no sickle cell gene that all?

OK. Well, we can just look at the square. The information was that this is 1%. So that means this 0.1. And that's 0.9. That's 0.9. So this here is people with no sickle cell at all. It is 0.81.

And so the square actually makes everything really easy to understand. So now, the question is how do you generalize it? Well, what are some of the-- the first generalization I'll do is I'll say, suppose that there are-- well, I'll ask you. Should I generalize it to more dimensions or more different copies of the gene? What's easier to draw?

More copies, because more dimensions-- I don't know how to draw in three dimensions. But I can draw three copies. That's pretty easy.

So now, take a piece of paper and just on your own, draw this same figure for three copies of the gene with frequencies p , q , and r .

So people have a picture? Does anyone want to describe their picture to me? Yeah.

AUDIENCE: I did a three by three square with p squared and q squared and r squared diagonal [INAUDIBLE].

PROFESSOR: OK. So I do the same. I make a p, a q, and an r. And then I do the same here-- qr.
OK. So there's a p squared, q squared, r squared.

Now, how many guys-- so we have nine square, sub squares. Three of them are this. There's six more. Now, which of them look similar? So this is a pq. Are there any other pq's?

Yeah, one other, right? pq over here. . And here is a pr. Is there a pr? Yea, another pr over here. And here's a qr and a qr. So we have p squared. And that's that guy. We have q squared, r squared, and then 2pr, 2qr, and 2pq.

So that's the generalization to three copies of the gene. So it turns out that all these coefficients-- 1, 1, 1, 2, 2, 2-- those are binomial, multinomial coefficients.

So now, let's generalize one more. So what have we done? There, before, we've written out this. And those are the nine terms grouped into six. So this was our C. So now, we know how to generalize.

p plus q plus r to the C equals 1. If this is two chromosomes-- this is C chromosomes. So this is C chromosomes with three copies. Well, what happens if we have n copies? $p_1 + p_2 + \dots + p_n$ to the C equals 1.

And all you have to do is expand that out. You can use a square in higher dimensions. Or you can actually use formulas for math. And that's where these come from-- p_1 to the 1 power, p_2 to the next power, all the way to p_n to the next power.

And these guys are the coefficients that count the multiplicity. Yes.

AUDIENCE: Why do you have to add up to 1?

PROFESSOR: Oh. They have to add up to 1 because p plus q plus r, if you only have three flavors of the gene-- either you have sickle cell A, B, or C, let's say-- then these are probabilities, probability p.

AUDIENCE: Oh, so parentheses says how often it's done?

PROFESSOR: Yes. Frequency and probability-- so $p + q + r = 1$. So if you square it, you still get 1. So you start from the idea that $p + q = 1$, which is what we did over here. So $p + q = 1$. Then you square it. It's still equal to 1. So now, you just another way of writing 1.

So it turns out that Hardy-Weinberg is all just fancy ways of writing 1. Starting from this picture, we've successively complicated it. This is $C = 2, n = 2$. There's $C = 2, n = 3$.

Here is the same thing again. Here is $n = 3$, general C . Here's general n , general C . OK so what we've done is we've basically got here-- by stages of successive approach, one step at a time. So at every stage, it's clear what is going on.

And what's the core idea? The core idea is the one you just asked a question about, which is that the frequencies add up to 1-- $p + q = 1$ -- so if you square the frequencies, you still get 1. And there's nothing more to Hardy-Weinberg than that. Yes.

AUDIENCE: I'm Julie.

PROFESSOR: Julie.

AUDIENCE: My question has to do with the original way you presented the problem-- using the word flavors instead of alleles. And I've always been fond of teaching that you should always use the probable vocabulary of your students.

PROFESSOR: To switch.

AUDIENCE: So you got the same thing as the--

PROFESSOR: Yeah, good question. So I used flavors instead of alleles-- so right here, imagine a gene with two flavors. So actually, probably the best way to do it is to combine the two. So you say flavors, because-- so this is a question of transmitting information to the student without noise on the channel.

So if you say the word allele-- so this is again related to chunking. If you say the word allele, the problem is that now, you're expecting them to try to understand this new idea as well as this new item for taking up one of their chunks that they have available.

So when I initially presented it, I would use flavors. And then I'd say, OK, now we understand the whole idea. The thing has kind of seeped-- it's not really part of short term memory anymore. It's connected to something else that they know-- for example, this. $p + q = 1$, so $p + q^2 = 1$.

Now, it's not taking up so many chunks anymore. Now, they're ready to hear the word allele. So I'll say, OK. Colloquially, we could say flavors, but actually, the word in the literature is alleles. So at no point, you're overloading the system.

So again, it's philosophy based on, say, history and chunking. Yes.

AUDIENCE: When you presented this whole second way, you made, actually, use of the equation that you wrote down in the--

PROFESSOR: Oh, over there?

AUDIENCE: Yeah.

PROFESSOR: Oh, no. I was just saying that for your benefit.

AUDIENCE: But I thought, actually, for me, it was very useful to see what the kind of equation was. And then as you were going through the argument, I could reason out myself--

PROFESSOR: Good question, good point.

AUDIENCE: --where everything was coming from.

PROFESSOR: Yeah OK. That's a good point. So it actually was helpful to actually see the final goal to know where you're going. Actually, that's a good suggestion. It's actually not bad to say, OK, this is what we're going to try to understand. I don't expect to understand it now. It's full of all kinds of squiggles.

So for example, maybe a really good way to do the whole thing would be to start with the Hardy story and say, we're going to talk about Hardy-Weinberg. Say, what does Hardy-Weinberg say? Well, in its full generality, it says, blah. I don't expect you to make any sense of that right now. Let's talk about what the core ideas of it are.

With this, you approach it, creep up on it bit by bit, successive generalization, and then you get there. So actually, it's a good idea. Leave that on the board the whole time so you have a context and a goal, like a mountain peak you're trying to scale.

Thanks for the suggestion.

Yes. Tell me your name?

AUDIENCE: Scott.

PROFESSOR: Scott.

AUDIENCE: Last week, you did present an example where you wrote down this complicated physics--

PROFESSOR: The Navier-Stokes equation.

AUDIENCE: Yeah, you wrote down the Navier-Stokes equation, and then you said, yeah, I sometimes do that just to create this tension in the students [INAUDIBLE]. So given that when you complete this method-- you actually didn't use the story at all. It seems to me that if you had presented that and then you stood back and made a joke about [INAUDIBLE], it would have been perfectly satisfactory.

PROFESSOR: I think that's true, too. So a lot of teaching is preference of the instructor. I, for example, happen to really like that story because I taught in Cambridge for a long time and I can imagine Hardy running into Punnett at the cricket ground. So it's interesting to me, so I can tell it with enthusiasm. And I think it says something about the relation of science on society, which I think is important for students to learn.

So I might use that story. But an instructor with different purpose might do it, as you say, which is make a joke about how impenetrable this is and say, it seems really

cryptic, but we'll actually understand this by the end. Don't worry.

And that's a good way of-- actually, I think you're right. It's a good way of creating and releasing tension. Yes.

AUDIENCE: Eric.

PROFESSOR: Eric.

AUDIENCE: So what level of background do we assume the audience has? Because you use sickle cell as an example, and it's obviously not critical to the example that they know what sickle cell is. Do you assume that they know that?

PROFESSOR: Right. So what level of background? So if they don't know what sickle cell is, or maybe I'd use cystic fibrosis, which they're maybe even less likely to know.

AUDIENCE: --presentation.

PROFESSOR: Yeah. I could be adding noise. So it's a flip side. It's a two edged sword. So are you adding noise by saying sickle cell to someone who doesn't know what it is? Yeah. It would add noise. So maybe it's actually worth saying just one sentence. What is sickle cell anemia? It's a mutation of the red blood cells that makes them take a different shape and makes you unable to transfer oxygen as effectively and makes you more resistant to malaria.

So actually, that part, I wouldn't say at the beginning. I'd say, well, if it's so bad-- can't transport oxygen so well-- how come it's still around? Just let people think about that for the day and come back the next time and tell them the answer.

But yeah. So I'd probably say one sentence if the people haven't heard of sickle cell. When I was saying it, I was assuming mentally that they know what it is. But I think you're right. Many people wouldn't know what sickle cell anemia is, and it's worth saying just one sentence.

So let me just see if there's anyone who hasn't made a comment or question. Yes. Can you tell me your name?

AUDIENCE: [INAUDIBLE].

PROFESSOR: [INAUDIBLE].

AUDIENCE: Could you point to the last little bit, because I'm curious to see [INAUDIBLE].

PROFESSOR: Oh, how do you get this piece?

AUDIENCE: Yeah.

PROFESSOR: OK. So how do you get this whole thing? OK. So let me do that by analogy. So for example, suppose-- let me do C equals 3. I'm running out of board. Let me erase this one.

And in fact, let me just use n equals 2, because it doesn't illustrate any new ideas to crank n up. But the cranking the c up, the number of copies, actually makes it important to see why you need to add them all up.

So let's do p plus q cubed. So this is an organism with three chromosomes. But it's either sickle cell or not at each spot. OK. Well, let's actually just write this whole thing. You could do it by looking at a cube and seeing all the chunks, because you could actually just do it with algebra, too.

So there's a p cubed. There's a ppq . So there's a ppq , pqp , and qpp . So those are all contributing to p squared q , and there's three of them.

OK. Now, the three isn't what we're talking about. We're talking about this exponent here. Why is K_1 all the way up to K_n equal to C ?

Well, let me just put that 1 in there. What does this equal? The sum of those guys is 3. That's p to the cubed q to the 0. That's also 3. What's the next term? Well, there's a 3 p cubed squared, and there's a q cubed.

Well, that's q cubed p to the 0. That's 0 plus 3. This is 1 plus 2. It's always three.

And it's always the C . Why is that? Well, you only have three products-- p plus q . So you have three factors. And you get to choose, when you're writing out all the

terms-- there are eight of them and we've combined them into four groups. There's three here, three here, one here, and one there.

You get one from each of the factors. So you get one exponent from each factor. So the total of all the exponents has to be 3. So once you understand it for 3, then this is just for C in general. Does that help?

So that's a good example, actually. Someone probably would ask that question or should ask that question. And that's how I would answer it. OK. Now, Cecilia, you had another question. Yes. Did you have another question?

AUDIENCE: What are we proving? For example--

PROFESSOR: Yeah. You'd think, on one hand, what are we proving? Well, we just proved $p + q + r^2 = 1$. So it looks like there's no content.

AUDIENCE: [INAUDIBLE].

PROFESSOR: How has that done anything?

AUDIENCE: The key fact is how you prove that the 1, the 3, the 3, and the 1 in each [INAUDIBLE]. For example, are we supposed to prove that today or break even?

PROFESSOR: Well, I would say there's no less information-- well, what we've learned is that, for example, we've learned this. Why are there various products here? And then what this thing is-- we've learned intuitively what this thing does.

This thing counts for the number of copies. So in the original-- $n = 2$, $C = 2$ -- it was either there's two copies or one copy of each of these guys. So it adjusted for the number of copies. When you have $p + q + r$, there's one copy, one copy, one copy. There's two copies of that one, two of that one, two of that one.

So it's that factor. It's the number of copies factor. And then in probability course, you learn how to count those factors in terms of factorials. And that has a definition in terms of factorials. But I wouldn't focus, for example, in a biology course on why it's factorials, necessarily.

I'd want them to understand that this thing counts for the number of factors. But actually calculating the number of factors for general C and n wouldn't be my first goal in the course.

AUDIENCE: So if I didn't have that in mind, why would I-- so the general purpose is teaching the terms in the multinomial expression.

PROFESSOR: Right. So it's terms in a multinomial expansion-- so basically, it's from writing out 1 equals 1. So it seems like, oh my god, we've done nothing. What have we learned? We've just learned 1 equals 1, which we already knew.

But actually, by splitting up 1 on the other side, you've actually given meaning to each of the terms on the other side. So this is the frequency of probability of having three copies of gene type A and none of the other type. This is the probability of having two copies of gene A and one of gene B. And this is the probability of having one of gene A and two of gene B.

So there are three, because there's different ways of doing it.

AUDIENCE: So do you have these chromosomes with different flavors [INAUDIBLE]?

PROFESSOR: That's right. Because these organisms we're talking about come with multiple copies of chromosomes. Yes.

AUDIENCE: [INAUDIBLE].

PROFESSOR: What is that? Yeah. So what's the phenotype in the end? And you're right. So with the sickle cell-- so actually, that is an advantage. So the point is you need to link this phenotype and the genotype. And that's true, actually. I've erased the sickle cell, but actually, the sickle cell is a good example for doing that.

If you have no copies of the sickle cell gene, then you're perfectly healthy. If you have two copies of the sickle cell gene, you have sickle cell anemia. What happens if you have one copy of the sickle cell gene?

Well, you do have some symptoms of it. So the question is why does that gene

survive? Well, it's because you're actually more malaria resistant. At least, that's one of the theories. So somehow, I guess the malaria parasites can't eat those red blood cells as well because they have a different shape or can't invade it. So it actually gives you some advantage and some disadvantage, but they're sort of balanced.

And yeah, if you have full blown sickle cell with two sickle cell genes, then you're in trouble. But that's much rarer than the 2pq, as long as q is small enough. So then you can actually continue that example. You can look at the frequency of sickle cell gene in different populations and say, OK, well, is it higher in areas where there's malaria? And test that theory. Yes.

AUDIENCE: I noticed that the first one that you did was to stay mathematical with the number line. When I've, in the past, taught [INAUDIBLE], I always stick with conceptual-- p is this allele, and then you've got a phenotype in the end. I'm just wondering if you chose this because you know that you're teaching mostly to people with math rather than scientists. And if you were teaching to, say, English majors--

PROFESSOR: Good question. Yeah, so I think you're right. I did this kind of mathematically, compared to maybe how you've taught it to biology students. And yeah, I guess I'm implicitly assuming-- but I didn't say, and I should say-- that this is MIT students. I'm just, in the back of mind, thinking MIT students.

But I didn't make that assumption explicit. So MIT students are perfectly happy with squaring binomials and trinomials-- trinomials, usually. And to the C power-- sort of stretching it, but it's OK. But generally, that way of doing things for them is good.

If it's people, English majors, yeah, you're right. I would try to actually do even more conceptually. But one thing that's good for all of them is the picture, because once you see the picture, then you actually understand the idea in one grasp.

It's really just one chunk now. Oh. It's just a square in four pieces. Oh, OK. It's just a different way of writing 1. This is one area 1 and there's four pieces. So that I would keep, no matter who I would talk to. And the question is how much would lead up to

it? Yes

AUDIENCE: So I think I just had an idea here that I think highlights why you can't just go in with this, because there's a disconnect between what these terms mean at the end.

PROFESSOR: Why you can't go in with A.

AUDIENCE: Right. So I think you have to draw a connection between, say, the areas of the boxes in the diagram and the phenotype frequency.

PROFESSOR: You're right. So what you're saying is that you can't just launch in with this, because even though you could give-- say you gave really exact definitions of what all these things are. It's not clear that exact definitions are computationally productive for a student. Yeah, they may have the exact definitions, but they can't actually use it.

For example, let's do an exact definition of chess. I'll tell you all the rules of chess. And that's enough of an exact definition to be able to decide what the best first move in chess is. But it's computationally useless. I still don't know what the best first move is, even if I know the rules of chess.

So just telling the student all the rules that, say, define what a genotype is and what's polyploidy, doesn't mean they can actually use it in any problem. So if you want to transfer, all these things have to have meaning for them. And that's what the goal of this approach is.

Now, I think the approach has been improved from your suggestions. I just wanted to show you a direction to go. But I think you've actually taken it farther and extended it and improved upon it. And the general principles among all of them are that you want to-- I would say one of the key ones is chunking.

You want to not overload the chunk system. You want to somehow bring people in so they'll even listen to you. If they don't listen to you, if they don't care, the learning is going to be so much less. So all those are for that.

So the dictation and jargon oppose chunking. And these all go together. OK. So now, what I want to do is give you a short answer to one of the questions that was

raised earlier. How you become a good teacher?

The reason I want to do that is that is exactly the same thing-- if you understand how you become a good teacher, you understand how you become good at anything-- how you become good at chess, how you become good at biology, how you become good at solving physics problems, how you become good at playing concert piano.

So if that's what you want to teach your students-- to be good at those things-- well, you want to understand that in a context, say, that you're working on-- say, being a good teacher. And to do that, there is the following set of experiments.

So projectors-- OK. So the way I'm going to illustrate this is I'm going to show you a chess position. And the goal is to try to remember the chess position. I'm going to give you two more seconds.

OK. So everyone got the position? Now, I'm going to ask you instead of to remember exactly, to reconstruct it. How many pawns were there? That would be E.

So who votes for A? OK. Who votes for B, nine pawns? Who votes for C, 11 pawns? Who votes for D, 13 pawns? Who votes for E, none of the above?

OK. So let me show you. Then I'll explain why I asked you this particular question. So there are actually 11. Now, it's a very hard task. So this actually, this very task-- not counting the number of pawns-- I made that slight variation. But the so-called reconstruction task was given to chess players of various abilities.

Grandmaster slash master is one group. What are called experts-- experts are people who are not quite chess masters, but close, in chess lingo. And then class A players, which is pretty strong tournament players. So they were given the task of looking at a position for four to five seconds. The position was knocked down, and they were asked to reconstruct it as accurately as they could.

OK. So it's even harder than the task of counting the number of pawns. So the results are very striking. So by level of chess player-- so class A is the strong

tournament players, experts or grand master or master.

So the percent correct-- 51% of the pieces correct, 72% or 93%. So the grand masters and the masters were amazing. And in fact, for the number of pawns, I don't think they ever make mistakes on that, because pawns are one of things that you just know as a really strong chess player.

So 93% percent correct-- and that's amazing. So there's an related story, which is about the memory of chess players, which is the Bobby Fischer-- yes.

AUDIENCE: So I'm thinking the positions, the positions wouldn't work. It wasn't possible. They wouldn't do it as well.

PROFESSOR: I'll come to that in one second. So Bobby Fischer was in a tournament and some strong master was playing in it. And Bobby Fischer went to the bathroom. And as he went to the bathroom, he happened to see that master playing a game and just continued walking. And then about six months later, he ran into him at another tournament.

Fischer said to him, oh, in that position in that tournament, did you play blah? And the guy said, well, actually, I had no idea what the position was, even. Fischer said, oh, and he set up the board and said, well, see, this was the thing that you really needed to do. And by then, he sort of remembered. But Bobby Fischer remembered at one glance six months later.

So what's a natural conclusion from this data? You'd say, well, naturally, grand masters and masters-- they were born with better visual memory. But in fact, the crucial experiment was then done in 1973. So this experiment was done in 1948 by de Groot.

What Chase and Simon did in 1973 was that they showed positions that were random. So they redid the experiment-- they confirmed these results. And then, as you suggested, they showed just positions where the pieces were scattered arbitrarily over the board. And then everyone was basically at 12%, give or take 1% or 2%, just random variations.

So what does that show? It's not that the-- maybe Bobby Fischer was an exception. But for almost everybody else, even the very strongest players, it's not that they're born with a better visual memory. It's that they've learned somehow a way of looking at chess positions that there's less to remember for them.

OK. So now let's compare. Here, when the student sees this, there's a ton for the student to remember, just like when we look at a chess position, every piece is separate. But what does a chess player see? When a chess master looks at a chess position-- so I'll put the position back and I'll show you what the chess master sees.

The chess masters see something very different. They see groups of related pieces together. So for example, here, the chess master sees-- this king here is not a surprising thing for this chess master. That's when you castle your king. That's where it goes. Then the rook goes next to it, and then you usually move your rook into the middle. So that's not surprising.

This rook is also not surprising, the second one, because it usually comes from this corner into the middle. So all of this almost contains no information for the chess master. Here, these three pawns are very common, with the castled king on that side.

But then it looks kind of strange. There's some new information there, because maybe this king castled, but then the rook went all the way to the corner. So actually, maybe black never castled. And his kind just sort of wandered into this area.

What does that suggest? This suggests that the black king is really vulnerable. Maybe it's time for an attack. And in fact, I'm pretty sure this is a position from one of Garry Kasparov's games. And in fact, that is the right conclusion. The right conclusion is that it's now time to sacrifice your knight and take this pawn and draw the king out. And he actually won using that-- by sacrificing his knight.

So the chess master looks at it completely differently than the novice. I'm a novice

when I play chess. To me, every piece is a new bit of information. I'm way overloaded past my chunk threshold. I can't hardly remember the board at all.

So your students are in exactly the same position when they're learning material that, for you, you're the chess master. So you're teaching Hardy-Weinberg. Well, clearly, you've been appointed to teach Hardy-Weinberg because you have a Ph.D. In biology. You know a lot of biology. You're the biology master.

So this doesn't surprise you that much. But for the student, every single almost letter in there is news to them. So what you want to do is you want to find ways of thinking about it that you can group the ideas into chunks.

So here is almost one chunk-- for example, the idea that really, it's just p plus q squared. And there's a picture for it. And once you understand that, there's another chunk. There's another idea which is well, you could actually have three kinds of flavors instead of two.

OK. p plus q plus 4-- oh, I can transfer it there. Oh, and then once you understand that, you can transfer it to n flavors. Now, you can increase the number of copies of the chromosome-- so into two dimensions, three dimensions, four dimensional picture.

So then, you can actually make sense of all this. You can have a way of understanding the position.

And not too long ago, there's, I think, a not well enough known paper-- I'll put the reference on the website-- which shows the relative importance of symbolic calculation versus perception. So this is a perceptual mode.

So much of our teaching is, let's say, left brain-- very, very symbolic. Well, there was this really interesting study done of chess grandmasters-- in fact, of the strongest chess grandmaster today, Garry Kasparov.

What's the relative importance of perception versus analysis in his really strong chess playing? So the way they tested that-- there was a really good experiment-- is

he would play simultaneous exhibitions all the time. So the way you do a simultaneous exhibition is there's, for example, let's say, 10 opponents. And you would just go around one component after another.

They have the full time, say, three minutes, to think until you come back. But as soon as you get to a board, you just think for about five or six seconds, maybe 10, and make a move, and go to the next board, so that by the time you come around back to that same opponent, they've had their couple minutes to think.

So now, he played simultaneous exhibitions against very strong grandmasters. And you can then measure his performance there. So why is that a good experiment?

Well, he's now not able to do all of his calculation that he does normally. He's normally able to think for three minutes, maybe five minutes, and do a whole bunch of analysis, symbolic computation.

But when he has five seconds, 10 seconds to think, mostly it's perception. Well, his chess rating effectively dropped by maybe 50 or 60 points. So 50 or 60 points, to give you an idea-- his chess rating is the highest in the world. It dropped to a level which only five people in the world are higher. So it's a very small drop.

So he still plays incredibly strong chess, better than almost every other grandmaster on the planet. So almost purely with perception-- so what that shows is that the way Kasparov has become so good and in general, experts have become so good is it they look at the world different. Their perception is different.

So how do you do that as a teacher? That's one question. How do you promote that in your students? Well, you want to give the ways of looking at the world that change their perception. That's why I'm so focused on the story, the tension, the human, the right brain, the pictures, the chunking, because it's those that are actually producing long term expertise, whereas this is producing what would be the equivalent of programming a chess computer.

But that doesn't work. That may work for chess computers to play good chess. But it doesn't actually work for people to be able to use the knowledge later.

So now, what produces that? So there's one short answer which is that for teaching, you want to change your perception of how students think. If you have a correct, new, good perception of how students are thinking, then you're actually able to make teaching judgments on the fly.

You can plan chess move, your lecture, like a chess game. Your intuition is right. So you want to tune your intuition.

Well, that is why I do this. I've found the single most important thing that has improved my teaching, and I highly recommend, is the feedback sheet, because for example, I learned what was confusing in question one. And in question two, I learned what works and what doesn't work.

So as I see what works and doesn't work, I start to build up a more and more accurate model of you. And I start to be able to plan and reason about how to teach you and, in general, how to teach students.

So I'll talk about that more. That's the idea of deliberate practice and expertise. I'm going to talk about that more in the subsequent sections in more detail and show you some of the studies around that. But the general rule is you want reflective, quick feedback on what you're doing in order to become an expert. And that's true whether you're in teaching, concert piano, physics problems, whatever it may be.

So with that said, if you can fill out the sheets so that I can become a better teacher, that will be very helpful. And one announcement, which is that next week is a Tuesday. MIT is open on Tuesday, except it's Monday's schedule of classes. So we don't have a class next week.

The week after that, I'm a witness in an administrative law trial. So I'm not here. So there's no class for the next two weeks. So we'll meet again in three weeks. And I'll post some readings and a short problem set for you to work on-- some readings growing out of what we've done today.

OK. So if you could bring up your sheet and your index card to separate piles, that

would be very helpful. And there's going to be another class coming in-- a big class, I think. So I'll just go outside and answer any questions that people I have right outside so that the new class can come in.

SPEAKER: Answers from Lecture 3 to questions generated in Lecture 2.

PROFESSOR: I'm going to first answer questions from before, since there are lots of questions, and all interesting. And I'm also going to do another equation example. There was lots of requests for another equation example to see how it plays out in a different field and a different way of approaching equations, not just from the entry point of a story. So I'll show you that.

And then we're going to look at misconceptions in various fields and the fundamental importance of understanding that so that you can understand how to change your teaching and how to reach the students. Basically, if you can't understand where they are, you can't come to them.

So questions from before-- one comment was that I don't often enough summarize the end result with the transferable lessons for later. So thanks for that suggestion. I'll make sure to do that.

Another question was graduate versus undergraduate classes. We talked a fair amount about audience a bit last time in response to questions. So how do you change your teaching in response to questions, in response to the change in audience?

So the particular question here is graduate verses undergraduate classes. And the sense I got from some of the questions was that somehow, it's harder to do what we were talking about last time, which is teaching equations in an intuitive way, in a graduate class than a undergraduate class.

Actually, in some ways, it's the opposite. It's true that generally, in graduate classes, people just put up a ton of equations. For example, in quantum field theory, you just get a gazillion integrals with epsilons floating all over the place and then path integrals. You integrate this. And there are some 2 pis and you take a bunch of

limits. And it seems like a whole bunch of methods gymnastics.

But A, it doesn't have to be that way. And also there's another characteristic of graduate students which you don't have so much with undergraduates, which is that graduate students know how to read. Now, this may seem like a bizarre statement, because surely, everyone knows how to read by, say, age three or four or five or whenever they teach reading in school these days.

But what I mean is that undergraduates-- generally, they have no experience on how to read a textbook because they've had so much experience with us telling them stuff everything on the board. So they have no incentive to actually read the textbook. And they don't learn how to do that.

They think textbooks are read the way you read Jane Austen novels. You just read for plot. Something happened to some equation, then something else happened to some equation. And you just carry on, paragraph by paragraph, as you would a novel.

That way of reading is completely hopeless for technical material. But graduate students-- not always, but generally-- have much more maturity about this. So graduate students actually can or often are closer to being able to read technical material with skill. For example, graduate students often read papers in their own field. And you know you have to read a paper differently than you would a Jane Austen novel.

So because of that, you can actually teach equations very differently. What you do is you give all the long, messy, yucky parts-- you leave that for the notes, for the book, somewhere where everything is printed in a very easy to read format, rather than copying long, long, long, long strings of symbols off the board.

So that connects back so what we talked about last time, which is chunks. So if you put long equations up on the board, generally, you overflow the chunk system. And once that happens, people start making mistakes. So you want to avoid doing that as much as possible.

And with graduate students, it's even easier to do, because you leave all that for a type set, professionally published book or type set by yourself, but somehow printed in a clean way with no mistakes. And you can then, in class, discuss the meaning of the terms. What are the terms? Where do they come from? Why would you expect that kind of term?

So I'll give you an example of doing that with an equation today. But generally speaking, all of what I was talking about last time applies perfectly well to graduate classes, even if people at first think that it doesn't.

So I should say if any questions occur to you as I'm answering questions-- basically, questioning beginning questioning-- please raise your hand and ask them right now.

OK. I found that the square diagram muddled the development of the Hardy-Weinberg equation. So the square diagram was this one.

So the conclusion from that was the question-- which is shouldn't college students be comfortable with expanding p plus q squared? Why do they need a diagram? And the answer isn't that people aren't comfortable with expanding p plus q squared, although you will find people for whom this and this is just a symbol replacement strategy. In other words, it's something like you program in a computer whenever you see this pattern do this. But the terms don't actually have meaning for people. They don't know why those terms are that way. And if they'd misremembered and put cubes here, they would write that down too.

So the picture actually makes it clear what the meaning of the term is. So it gives, actually, meaning to what people might be otherwise comfortable with or having done from lots of practice.

So I should say there's a difference between just rote matching between here and here versus this kind of understanding. And that's illustrated by the following research, which is about people's ability to multiply and add when people have brain damage.

So the technical term is brain lesions. The people with brain lesions in the-- let's see if I can say this right. Some people with brain lesions in the arithmetic areas lose their ability to add, but they can multiply fine. And people with brain lesions in the verbal areas, they lose their ability to multiply, but they can add fine.

Now, this seems kind of strange, right? So why would multiplication and addition to not go together when you get brain damage in the arithmetic area? So if the damage is either an arithmetic or verbal, here's what you lose. So arithmetic area damage, you lose addition. Here, you lose multiplication.

And that is very bizarre, because you'd think you should lose multiplication here, too. And the reason that it doesn't work that way is because of the way multiplication is generally learned and taught. So for example, in England the way it's taught is-- this is research from Brian Butterworth in England.

The way multiplication is taught in England is in the multiplication table, same is here. And people memorize it as six 9's is 54. So six 9's is 54. Here, I think I learned multiplication both places. So 6 times 9 is 54.

Either way, that's a purely linguistic string. It's no surprise that when you lose verbal ability, you lose the ability to memorize linguistic strings. So the multiplication table went.

So what that tells you is that most people-- and I don't know if this is true for everyone, but my guess is it depends on how you learned the multiplication table. For most people, they learned the multiplication table in a way that is not meaningful. They've learned the multiplication table purely linguistically.

And that is the problem with just going down this path only, which is that-- again, this is where it's important to know where the students come from. If you know students basically have just linguistically memorized this transformation to that transformation, like the rock bands in other countries that sing in English. They know when they see English words. They know what to say. But they don't necessarily understand any of the words.

That is actually very common. And this is another linguistic transformation. There's no meaning underneath it. So here to give it some meaning underneath it, some picture, it actually incorporates another brain area into the understanding, which is where addition actually has some meaning to people.

So that brings up a related question, which is if people have learned multiplication in this linguistic way, how could you teach it in a non linguistic way so that people actually understand it?

There are several ways. I don't know of any studies that show that after doing it this way and then brain damage, it doesn't get lost. But my speculation is that if you learn it, actually, pretty much the way I learned it, you wouldn't actually lose it.

And the way is, for example, suppose you have to multiply 6 times 9. Rather than memorizing that as 54, you think about what it should be. And you reason your way to it. And we think, oh, well, that's slower. Yeah, it's slower in the beginning.

But in the end, you get to the same place. But you've put it in a different part of the brain. So the way you could do this one, for example, you say 6 times 9. That's 6 times 10, which is really easy, minus 6. So 6 times 10 you know not by memorizing but from understanding the number system.

So you know that that's 60 because of the way the number system works. You don't have to memorize that. That's the kind of example where I think calculators should be programmed to self destruct or at least not work for about a week if you type in a problem like that. They should just freeze as an incentive to actually think about these things before you put them into the calculator.

So minus 6, and you get 54. So now, that's a bit longer the first time you do it. It's longer than memorizing it. But after you do multiplication a bunch of times like this, you actually reinforce the meaning of the number system, and you come up with the same answer. And eventually, you will memorize it, but you've memorized it in a different way.

Here's another example. From the 12 times tables-- 8 times 12. Now, should you

memorize that as 96? No. You should write that as $10 \text{ minus } 2 \text{ times } 10 \text{ plus } 2$. So that's equal to $10 \text{ squared minus } 2 \text{ squared equals } 96$. And you can even draw a picture for this and show what happens to squares and areas.

So again, you've come up with the same answer, but you've done it in a meaningful way. So that's one answer to why it's worth showing pictures, even if people can do the algebra. It reinforces the algebra.

Is determining what constitutes a chunk simply a matter of my intuition about students' level? Well, to some extent, it is. But in the chess playing research I talked about, they actually had more objective ways of determining what a chunk was.

And what they did is they put eye trackers on people. And they had the chess masters and the non chess masters and the experts look at the board, and they tracked their eyes to see what they did. So the non chess masters, their eyes wandered all over. But the chess masters looked at pieces in groups. They would go here, here, here, here, and then here, here, here, here.

And the assumption was that that's a chunk and then that's a chunk. Or if they went here then to there, that somehow, these two chunks were related. So there was ways in the chess playing problem of measuring chunks.

And then how do you apply that to, say, teaching physics? Well, that's a matter, partly, of intuition. And the idea is to help the students build up the chunks. So you have to be on the watch for what chunks you use. So you have to introspect. OK.

How do you keep things interesting and reveal material in a time appropriate way? So the comment was that it took a while to do the Hardy-Weinberg with the story and then building up to it. So this is a question of how do you plan lecture time. What's worth doing in lecture? Is it worth spending a bunch of time on understanding the concepts?

And there's two extremes to this. I'm basically towards one extreme, which is that if you don't understand the concepts, it's not even worth learning the thing, because

as a student, if you don't understand the concept, you might as well just forget the material right now, because you're going to forget it soon anyway. So the only benefit from taking the course if you don't understand the material, is you just get a great on a final exam and you pass a requirement. But in terms of actually changing how you see the world, it has no value.

So there was a study done, I think, at Carnegie Mellon. Yeah, this was at Carnegie Mellon. They studied freshman physics students. So they had students who took freshman physics and students who didn't take freshman physics. And then a year later, they gave them the freshman physics final exam-- I think it was the same one, pretty much, as the students who took freshman physics took, except with numbers changed, but otherwise, it was same problems-- to see whether taking freshman physics had any effect on whether you did well on a freshman physics final a year later.

And the conclusion was that it had no effect a year later. So yeah, sure, if you're give them the final the next day after the final exam that they took, they would have done pretty well, maybe 50%. Who knows what the time constant is. But certainly, by a year, it was gone.

So what that tells me is that that regular way of teaching material actually did no good to the students except for passing a requirement, but no intrinsic good for their way of analyzing the world. So what that says is then you really need to find something different.

And yeah, if it means it takes extra time in lecture, a bunch of time in lecture, so be it. At least people will understand something and they will change how they see the world. And that's the motive for today, which is to really understand misconceptions, because if you don't understand the misconceptions, you're not going to be able to teach in a way that produces long lasting learning.

OK. So another one was how do I apply this to something really abstract? The way of approaching equations, like the proof of-- oh, this is actually one of my favorite infinite series. So the question is how do you apply it to something really abstract

like this?

So this is a famous infinite series. What's the sum of that? And it turns out to be $\pi^2/6$. Well, even that, you think, well, how can I apply it to that? And it turns out there's great stories about that one, too. I think this is a story about this problem, which is that no one knew how to do this sum from 1 to infinity. It's quite a hard sum. And so it was set as a problem, basically, for the mathematicians and physicists of Europe.

And then someone produced a solution. I think someone produced a solution anonymously. And everyone basically figured out who had done it, because it had their handy mark. So the solution was by Euler. And it involved a whole bunch of trickery with polynomials and infinite degree polynomials. And it was a really sly method.

So if I was going to teach this equation, I would actually teach the history of it-- how was really hard, how you could actually guess this. What are ways you could guess this? Well, you could actually approximate the sum, get a number that's one or two or three digits accurate.

And then you feed it into-- does everyone know this guy? If you Google for that, you should probably find it. It's called the inverse symbolic calculator. It's a fantastic thing. I do not know how it works. And I would love to know how it works.

But what it is is you feed in a number, and it will tell you all the ways of producing something really close to that. So for example, if you put in 3.141, it'll say a bunch of numbers that get near here. And one of them is π .

If you put in this to one or two or three digits, it'll probably guess for you $\pi^2/6$. So that's one way of getting at an answer.

So part of the way of teaching it is to say, well, let's somehow get an answer with ways we can do that aren't too abstract. And then let's see if we can justify that answer. So even there, you're not lost. There's always stuff you can do.

Oh, another question, which was readings-- how do you incorporate readings into a course so that students do it? So that was asked twice, actually, because I didn't answer it the first time. So one way to do readings is something called reading memos.

And it's an MIT invention by Edwin Taylor, who's recently retired from the physics department. And what a reading memo is-- I'll put up the handouts. So I have often done this in my classes. I'll put up the handout for you to use that I give out and you can just copy it or do whatever with it.

So a reading memo is a request to the students to write you a short memo about something that you've asked them to read. It could be the draft notes for your textbook that you're working on, which is what I often do it with. Or it could be the textbook someone else wrote and you ask the students to read chapter.

And what it's not is it's not a summary of the text, because you already know what the text says. There's no point in getting a summary. What it is is students' reactions to it.

So anything that questions, things that puzzle them-- oh, I didn't understand why you did this or the author did this. And then maybe three pages later, oh, now, I see-- which, if you're the author of those notes, you know that you explain the two things out of order and you should connect them.

But what that does is it teaches students how to read actively, because again, I like I talked about, people just do Jane Austen's approach to reading technical material. And by getting students to read actively and formulate questions, by doing that, students learn a different way of reading, a way necessary for reading technical material.

And also by writing their questions down and you seeing the questions, you actually get a view into how the students are thinking. So it's actually a way of understanding what their misconceptions are, their conceptions of the field are, and tuning your teaching. Just automatically, you'll find your teaching will impedance match to where

the students are just by reading the reading memos.

It has a further benefit, which is that it inverts the normal power relationship between teacher and student. So for example, most assignments, problem sets, there's the correct answer, which you know. And you're seeing whether they know the correct answer. So they're now writing an answer, worried whether they're correct or not. And then you're judging them.

So normally, p set, the power hierarchy is you and then the student down here. And the student is looking up to you for validation. So now, this is not a good thing to teach. And maybe it's hard to avoid with problem sets. You have to do problems sets somehow when you teach people to do problems.

But you want to minimize this as much as possible, because it's not a transferable way of dealing with the world. They can't use that when they go elsewhere. And it teaches bad habits of deference to authority.

So that's normal. How does a reading memo work? Well, it's the other way around. If the student says, this is confusing, by definition, they are correct. It's confusing. They are the expert on what's confusing or not.

So it inverts the hierarchy to this. And you become very interested in what the students are saying. They are the experts now. And I've had very good results with doing reading memos. And my explanation is that it's because of this inversion of power hierarchy.

Now, what I mean also by good results-- two things. One is that I get fantastic feedback on the things I'm writing. The other is that I find students actually want to do reading memos after the class finishes. They say, oh, if you have more notes, can we just do some more reading memos? Great. Let's do that.

And it's because it's actually-- if you write the problem sets, they'll often say, can we do more problem sets? But that requires a fair amount of work to minimize the hierarchy. But it's automatically here in the correct hierarchy. So students actually enjoy doing that and want to continue.

So that's one excellent way of incorporating reading into class. So now, the problem is, what you do when you have 50 people in a class and you get 50 memos? So I've had this problem. And one thing I do is I just feel overwhelmed and I just flip through them but I don't know what to do.

But another is I revise my notes based on it. But the I think correct solution is an online system. So what you want is an online system where you can post a PDF file, and then people make comments on the PDF file. So everyone gets to see an image of the page, and they can just click and make a comment. And everyone gets to see their own comments, and then when they submit them, they get to see everyone else's comments.

So I actually wrote half of that system. And there's a graduate student in EECS who I think has now written a whole system independently of me. So I'm going to try it out and see how it works and try it in some of my classes this semester.

So the benefit of that is that you can then see all the comments at once, rather than flipping through 50 sets of reading memos with page numbers on them. OK.

How do I come up with intuition examples? How do I know if what builds intuition for me will also build intuition for the students? It's a very good question.

Is it just my personal opinion or is it just the teacher's personal opinion? One of the whole themes about this class is yes, teaching does have a fair amount of art and there is a fair amount of personal opinion in it. But there's also a fair amount of science and things you can do to make it more objective.

And one of them is actually to do reading memos. Any way you can to learn how students think will make it so that your intuition about the students actually matches how the students really think. That's the whole purpose of today about talking about misconceptions.

Reading memos are way of understanding what students think. Oh, there they are-- the reading memos. So once you understand what students think, it's much easier

to realize, just intuitively choose things that you know are going to work for them.

The other way is the feedback sheet. So every time the students tell you, oh, this really helps me or this really didn't help me at all, you now have one more piece of feedback about what works for them and what doesn't work for them. So then, you can actually choose intuition examples.

How do you invent them from scratch? Well, there are some general principles. One is use pictures whenever you can. Generally, that speaks to people's intuition, just because people have much more hardware for pictures than they have for equations.

So I try to put myself in the position of the student, and I say, well, for example, here. I say, yeah, all these equations may well be true, but I want a way that makes me see it instantly. And that just forces me to start looking for pictures. And that tunes me, actually, towards what students need. So you can do the same.

I think that was most of the questions. The other questions were similar to that. And I'll answer any that were new that I haven't answered-- but I think most of them I have answered-- at the beginning of the next lecture. So any questions that were generated by the questions? Yes, question.

AUDIENCE: So your story about the physics retention issue made me question the idea of the survey courses or topics in class, where it's like, I don't really need to learn this. I just want you to-- been exposed to it so that they remember enough to know to go back to it if you need it someday. That research, that that has any prayer of working.

PROFESSOR: Right. So the comment was what I'd said about even a year of intense freshman physics-- maybe it was a semester. I forget if it was just mechanics. But it was either one semester or a whole year freshman physics did no good towards long term understanding and change of understanding of freshman physics.

Well, what does that say about these big survey classes, where you're not expected to understand? So in freshman physics, at least they have the expectation that

you're supposed to understand everything. Now, what about the courses where they start with the expectation that you're not going to understand most of it?

That one is going to be totally hopeless. Then I think that is basically true. And maybe you could justify back in the day-- let's say, 400 years ago, even when books were around but there was no web. People need to know what books are out there. So this is a traditional thing in Cambridge, for example. They just give people big huge reading lists.

So then you go back. Later, you're like OK, these are the key books in the area. So you know one place to go for a reading list. That's kind of obsolete now with the web. If you want to find out something-- for example, suppose there was an equation I didn't know about. Let's say the Black-Scholes equation.

Would I say I wonder if I took any classes about Black-Sholes-- maybe, let me go flip through all my course notes? No. You just type it into some web search and see what shows up. And that is much more likely to be more relevant than some notes that you might have had or not had.

So yeah, I think the survey courses are completely pointless. Now, that doesn't mean that introduction to a field is pointless. But it means that the way to do the introduction has to be very different. You can't just scatter a bunch of topics at people.

What you have to do is figure out-- and we're going to talk about this when we talk about course design. You have to figure out what are the core reasoning ideas special that that field has to offer to the world? For example, history-- what's special about history? Well, historians have a sense of how to evaluate the validity and reliability of evidence and contradictory evidence. That's something you don't get in many other fields.

For example, it's in between a science and a pure literature field. In just straight literature, reading novels, there's historical evidence and things, but generally, you're reading in a different way. In math, it's hard to know where contradictory

evidence comes in, although there are ways of teaching math which I like which do that.

But generally, history has something new to offer, which is it's a messy world. You have noisy evidence. What do you do? Well, that's something that an intro survey course could actually teach. And that, somebody could transfer, even if they don't remember when did the Magyars invade Europe and all the random stuff that would be in a survey course. So those would be grist for the mill hung off big ideas.

So I'm going to talk about that when we talk about course design. But yeah, you're right. That course design is completely hopeless the general big survey. Other questions? Yes. Yes. Could you tell me your name?

AUDIENCE: Meg.

PROFESSOR: Meg. And what was your name?

AUDIENCE: Amy.

PROFESSOR: Amy. Thank you.

AUDIENCE: I was [INAUDIBLE]. I was just thinking about how for me, even if I understand something really well at the time, and I know that I'll use it again, it takes me-- having a test in front of me will not actually reflect whether or not I'm going to understand it, given the short period of time to remind myself. And so I'm wondering, if testing people out of the blue a year later is actually capturing whether the people who take this and perform faster [INAUDIBLE].

PROFESSOR: Yeah. That's a good question. So maybe it was a slightly unfair test, because they were just tested out of the blue.

AUDIENCE: If they had been using it all along-- if, in that period of time, they had taken courses that built on that material-- then, they would be reinforcing it all the time if they can.

PROFESSOR: Right. So if they had taken courses that used freshman physics throughout, maybe they would have remembered the freshman physics better. And I'm sure that that's

true. So I can tell you one story from my graduate time.

So I did a Ph.D. In physics. And I had to do the qualifying exam. And to do the qualifying exam, you have to study a whole bunch of undergraduate physics and then take the exams on it. And then you take a bunch of courses in various fields.

Now, the only thing, basically, I remember from all the electromagnetism is one thing, which is I understand pretty well the index of refraction. And why is that? That's because I was really pissed off-- sorry for the camera. I was really annoyed about the following thing.

So this actually goes back to what I was talking about about contradictions, which is that you're always told in relativity that speed of light equals C . And that's the great postulate of relativity-- that the speed of light does not change, dammit. And that's what Einstein said. And there's all these thought experiments with trains and lightning bolts and people throwing rocks from the train at different speeds. And it's C . It's C . It's C .

And then, somewhere later in an electromagnetism course, they say, the speed of light in a medium with an index of refraction N is C over N , where N is typically around 1, maybe a little bigger like 1.001 for air maybe, 1.33 for glass.

And how do those fit together? So that really annoyed me. And I wanted to get to the root of it and say, well, how could it be that you could have a speed of light that's always C , yet it looks like the speed is some lower number, V ?

So I worked out a whole bunch of stuff about how electrons scatter radiation and all the scattered radiation adds up and makes it seem like it's slowing down the light. And because of that, I actually understand this. And also because of that, which may be not so good, every time there's an electromagnetism problem I have to do, I always try to fit into a scattering problem. And if I can't do that, then I just can't do it at all.

And that's despite taking two years of electromagnetism-- two years as an undergrad and one year as reviewing as a grad student. So the point I'm trying to

make by that is that most stuff disappears. And the way to really make stuff stay is you really have to struggle with something.

And that's the most efficient way to make something stay. And that's not what happens in your traditional class, and even less in a survey class.

AUDIENCE: I'm just also asking-- it's almost impossible sometimes when you see material for the first time-- you might have to see it three or four times before anything comes, you're really going to understand it. And if the class is set up so that you will only see it one time because that's all the time you have, then you have to take more classes on it.

PROFESSOR: Right. So maybe you need to see something a few times to really understand it. So if a class doesn't give you that chance and then you have to take Thermo Two and then Thermo Three.

AUDIENCE: Hopefully the class gives you a chance. Say that it's building on this.

PROFESSOR: Right. So you should try to build that into the class. So there's a name for that, which is called the spiral curriculum. And there's a lot of sense to that.

So I'll just put a quick-- so you show the idea in its crude form. And then you spiral back to it in a more sophisticated way. But you'd like to do that soon, before the connection is gone, before this is actually wafted away. You want to spiral back to it. And so you'd like your class to do that.

So what that shows is that you should start with the big ideas and then you should refine them, because if you start with all the little details here, you'll just flood the chunking system. And actually, there will be no memory of it here and then you have to start over from scratch here.

You had a question. Could you tell me your name?

AUDIENCE: Roderigo.

PROFESSOR: Rodrigo, yeah.

AUDIENCE: Regarding the reading memos, do you think good to also ask the students questions about the readings? The reason why I am asking is because I know of a class that's actually implemented that PDF annotation system. And a lot of the comments that they get are too trivial, so to speak. They don't ask those type of questions. [INAUDIBLE].

PROFESSOR: So the comment is that a class that actually implemented the online annotation system, they found that the comments are too micro level and not broad enough about what's really confusing or interesting.

And so I haven't tried the online system myself yet. What I do know is that on the paper, if you do it on paper, you get really insightful, detailed comments. Now, I don't know what variables are different. One might be that the online system just encourages-- because everyone does things quick online-- it encourages quick clicking.

So it may be that it encourages less depth of thinking, whereas writing it on paper, I actually find I get a mix of a whole bunch of not trivial comments, but small comments like, some typos, that equation isn't quite right.

But then I get things like, I don't see the picture here or I don't understand why you did this now. And that, to my mind, is it useful comment-- or a question like that can't be right because of the following counter argument.

So on paper, I get a lot of interesting things. So it may be that online isn't as good. And that's one reason I want to try it and see. So it may be that they need to go back to paper despite it being harder.

Another is also what is the material? If the material is really boring, you're going to get really micro comments. So it helps to actually have written interesting material or give people interesting stuff to read that people are likely to think about and make comments. Question. Can you tell me your name?

AUDIENCE: Brian.

PROFESSOR: Brian. Yes.

AUDIENCE: I notice when you talk about equations, you like to give them the name that they're most commonly [INAUDIBLE]. Do you find that students get more understanding of concepts like this when they're given a name based on discovery versus terminology based on use of the equation? I think, from my mind, if you want to describe the relationship between stress and strain in material, do you want to call it Hooke's law? Or do you just want to call it the constitutive equation for solid material?

PROFESSOR: Right. OK. So the question is what about naming equations? Should you name them by who made them or by how they're used or what they are? So another example of that is, for example, the fluid mechanics equation. Should you call them Navier-Stokes or should you call them fundamental fluid equations?

And there's a tension there. First of all, I'll say that you do want to give a name. The big win is giving the thing a name, because that makes a unit of thought for the students. So that's the first order bit, the first order term.

The second order term is how should you name it? And there, there's something which is that you want a name that's common. So if, for example, people look it up elsewhere, they're likely to find more stuff about it. On the other hand, you want a name that's intuitively meaningful. So there's a tension. There's not often a right answer to that. And you can go either way.

So for example, I wouldn't probably use constitutive equation for the solid, because I have to think, what the hell does constitutive mean? So myself, it doesn't mean anything to me. There's another word. I'm trying to remember what it is.

Epistemology-- people just use it like it's just a plain, obvious word. But every time I hear it, I have think, what the hell does that mean? And then I put the translation into the sentence they're saying. And then I can sort of parse what they're saying.

So constitutive equation, to my mind, that's a word that is meaningful to the experts

and not so much to the students. So I would maybe call it the ideal spring equation, because it is the ideal spring equation. It's just the proportionality is slightly general because you have tensors. But otherwise, it is the ideal spring equation. So it connects to something.

And you can say, OK, who discovered? Hooke. So we often called it Hooke's Law. So then they have both And there's no harm in doing that. Question.

AUDIENCE: Because you mentioned a little bit before [INAUDIBLE]. If you're trying to teach something many times, then maybe it's better to tell the end, and then--

PROFESSOR: Yeah, the spiral.

AUDIENCE: Yeah. And I was just thinking because of the whole work-- so there are these really topics like the ideal gas law. And is it OK if I were to lie or say to students just for the sake of simplicity, for example, do you want to say whether you can actually use that equation in some cases or is it not logical always, and then you go--

PROFESSOR: Could you tell me your name?

AUDIENCE: Cecilia.

PROFESSOR: Cecilia. Yeah, thank you. That's an excellent question as well. So the question is, how much should you lie, if at all? For example, if I'm recommending teaching the big ideas in the overall approach first, that's almost necessarily going to involve some amount of lying, because the truth is complex and messy, the full truth.

And basically, you do want to lie. Some people hate it. But there's actually a really good book that follows this principle. And even if you're not interested in the typesetting system, you can see how it's played out in this book. It's called *The Tech Book*.

So that's the manual for the tech typesetting system, which I use and many people in math and physics use. Now, the reason it's interesting is that Knuth actually tells you in the preface, I'm going to lie to you.

So what he does is he has three levels of statements. There's statements that aren't marked with a-- so there's that sign. On the road, it means slippery, icy, something like your car might fish tail basically, danger. So there are statements without one of these, which may have some lies in it, not the full truth so that you just get the idea-- what are the fundamental concepts that tech uses.

When he starts to get into some gory details, but not super gory. He puts one of those bends. And when he has a super gory details, there's two of those bends. And then he says, look, don't read any of these things unless you've been working with tech for a year and are pretty competent with it. Don't worry about that. You'll be able to do what you need to do just by reading this and maybe the single bend sections.

So that's an example where the lying was put to really good use. And yeah, you should lie. In fact, you have to lie. There's no way to avoid it. And in fact, everything is a lie. It has to be, because to understand the universe, our brains are a constituent of the universe. So there's no way to understand the full universe, because that would involved packing more than our brain capacity into our brain. So just there is a pigeon hole principle proof that you have to lie to understand the universe.

So you have to say some stuff that isn't quite true. And where the art is is in choosing what is a useful lie. So what you want to do is develop the art of skillful lying. And that's a mark of a really, really good teacher-- skillful lying. Yes.

AUDIENCE: How do we know that you're not lying to us now?

PROFESSOR: I probably am. The question is how do you know that I'm not lying to you now? And I probably am, the reason being that I've now practiced lying so much, I don't even know when I'm lying and when I'm not. So no, I'll give you an example. I am definitely lying to some extent.

For example, there probably are situations where you don't want to any lies at all-- for example, teaching people how to manipulate the machines in the intensive care

unit. Maybe the first thing you need-- if you have a half an hour to teach them, you'd better teach them, memorize these damn things and don't mess it up or you'll kill somebody.

Maybe there are situations where lying is less important and lying is more important. And I haven't talked about those. So right away, I have lied to you, just because that. And I've skipped details because I wanted to get the big idea across.

So lying actually comes very naturally to me because I think the most important thing is the big details. So just by saying big details first, you're automatically lying. And I'm recommending it highly to you, too. Rodrigo.

AUDIENCE:

I have a comment on the lying. As a student, I have students, a lot of friends, who told me that they don't like it when teachers lie. But what I think the differentiating factor is whether they tell you that they're actually lying. And if they do, then it's totally fine. And if they don't, and a couple months from there, they just told you that everything was not really true, the students might get upset.

PROFESSOR:

Right. So the comment is that students often don't like when they find out that they've been lied to. But it's OK if you tell them that you're lying to them. And there's a general principle there which you can use in all your teaching, which is that whenever you do anything slightly nontraditional-- whenever you do anything like that-- it's really important to tell the students what you're doing and why, because they will go with you.

They'll go along with you if you explain to them the motive for it. So you tell them, look, this is a really complex subject. There is no way to understand the whole subject at its first glance. You need to develop high level structures first. And then you can put in the details underneath that.

So I'm going to tell you just the high level structures first. And there will be some untruths in that, but I'm not going to tell you things that are completely false that you have unlearn. I'm going to tell you things that you have to refine your understanding of or that aren't the whole truth.

So together with the idea of lying, you want to minimize stuff that you tell them that they have to unlearn and make it so that you're telling them stuff that they can keep using. It's just not the full story. OK.

So what I'm going to do is I'm going to give you just a quick example of another equation and how you could teach it. Then we'll take a short break and then we'll do misconceptions. OK.

So the other equation that I want to explain is this one. So I did a biology example before, so I chose a physics equation this time, which is the wave equation.

OK. Now, also to vary it, with the biology example, I introduced it with a bit of history. With the wave equation, I'm going to actually introduce it with a different approach, which is not to talk about the history but actually to get the students to try to construct the equation.

OK. So the question is what the hell is the wave equation? Well, the wave equation describes-- so here's some string. And here's your coordinate x . And you want to know how does the height of this piece here, the height being f of x as a function of time as well, change with position and with time?

So you want to figure out an equation for that behavior of that string, that stretch between two points. For example, this might be a guitar string, and these are the two ends of the guitar string. And it's under tension. And you want to know, how does it move?

OK. So we're going to construct the equation. So let's say this is a differential equations class for people in physics who are learning mathematical methods. And they want to learn how to construct differential equations. So maybe they're engineers, physicists. And they're, say, juniors.

So they have some mathematical sophistication and some knowledge of forces in physics and some knowledge of differential equations. So I'm going to write down the rough form of the equation, and we're going to try to figure out all the pieces and fill in the missing pieces.

So I'm going to give the general form of the thing. So there's some derivative of f with respect to time-- one or two derivatives, we're not sure-- is equal to something here. And then there's some constant here or maybe here. So there's a whole bunch of question marks to fill in. And what we're going to do is reason about what they are.

OK. So the first question is to figure out this guy. How many time derivatives do we need? So what this equation describes is the motion of this point.

Now, why does the point move? Well, you could ask the class, and eventually, they'll come up with, well, there's some forces on it, because the string's under tension, so there's forces on the point. So here's our point.

And here's, say, the string going through it. This is a blow up now of this region. OK. So now, what are the forces on this guy? Well, there's a force from that piece of the string and a force from that piece of the string.

OK. Because there's force, what's going to happen to the thing? Is it going to have a velocity or an acceleration? Anyone?

AUDIENCE: Acceleration.

PROFESSOR: Acceleration. OK. How many derivatives does acceleration have? Two. OK. So we're going to put two derivatives here. So you need two derivatives. So this, and now, we're going to say this is some kind of acceleration.

And maybe there's masses in there and stuff, which would all be slurped into these constants. So we have an acceleration. And now, we have to decide what generates acceleration. This side is generating acceleration or the force.

So we want to decide, for example, one or two derivatives. Generally, most equations either have one derivative or two. Some really nasty ones have four. But generally, it's one or two. So we'll just choose. Is this one or two? OK. And then the last thing we're going to do is to figure out the constant.

OK. So now, one or two derivatives-- one way to decide that is to make something that has just one derivative in it. So if here is my string and here is the point-- so f has a non-zero-- so it has a non-zero df/dx . But the second derivative is 0, because it's a straight line.

OK. So we know the second derivative is 0. Let's see what we can figure out what the force or the acceleration should be. OK. Well, here is a point. There's going to be a force on it from that end and that end.

And what's the net result of these two forces? 0. So when the second derivative is 0, in this case at least, we would like the force to be 0, which means that force has to be connected to the second derivative of position.

So we got that. Now, the next problem is to work out what goes here. The first thing is what's the sine. Should it be plus or minus? OK. So I'm going to ask you that. So find a reason whether it should be plus or minus here and a reason. So find a neighbor or two, and we'll take a vote-- plus or minus, intuitive reason for it.

I can't use this example to decide. Let's vote. Everyone have their votes ready? Who votes for plus? So about-- who votes for minus? OK. That's great. So we have a diversity of opinion.

So right away, what that shows you is that it's worth actually discussing that point in class, because if you just tell people, they'll write down something you tell them, but they won't have actually internalized it. It'll just be something that maybe contradicted what they said, thought, or not. And then they have to remember, was it what I thought or was what I not thought? Or was it what I not thought or what I thought?

So they don't actually understand why. So it's actually worth going through the discussion. So what example could you use to decide? Who haven't I heard from? Yeah. Could you tell me your name?

AUDIENCE: Mike.

PROFESSOR: Mike.

AUDIENCE: Well, if you're at the top of the arc, and you don't it's going to want to be pulled down. It's coming down.

PROFESSOR: OK. So let's use an arc like this. We can't use this one, because it's 0, and 0 doesn't have a sign. So it doesn't help us. So the next most complicated thing is this-- so an arc like that. Here's my point. So the force is downwards, or the acceleration is downwards. So we should have negative acceleration.

And what's the second derivative with respect to position here? That's also downwards, because the arc is like that. So the derivative has the same direction-- the space derivative and the acceleration. So it should be plus.

So we got the sign right. And to emphasize the importance of getting the signs, there's an interesting comment from Feynman back in the '60s at some conference about quantum electrodynamics. He was commenting about how crazy the whole process that he invented for solving quantum electrodynamics is.

And his reason that is so crazy-- he said, well, we do the first order term, and then we calculate the second order term and add it to the first order term. But it's very worrying that when we calculate the second order term after the first order term, we don't know whether the second order term is positive or negative. We just have to calculate it and see what it's going to be. But we can't predict ahead of time whether it's positive or negative.

And what that speaks to is the importance that physicists attach to knowing the sign of an effect. Is a plus or is it a minus? And so that's fundamentally important. You want to make sure you get that right.

Now, why did I do the sign of the effect after this, after the order of the derivatives? Because the most important thing is what do the terms mean. So this is a curvature. This is an acceleration. So here, let me write that-- curvature.

Until you know what the terms mean, there's no hope of figuring out what the sign

that connects them is. But the sign is the very next thing you do. And it's really important. It's a plus.

And now, we need to put in one more thing, because the dimensions are totally bogus. This is a position divided by a time squared. This is a position divided by a length squared. So we need to actually multiply by something to make the units come out correct.

So we need something that puts a time squared here and a position squared over here. So then the units will work out-- or a position squared over time squared, which is C squared, which is some speed squared. So speed squared is position squared over time squared. And that makes all the units work out. So there you go. There you have the wave equation.

So now, if you're going to actually formally derive it, that's all fine. But I would do this first so that people know where every single term in the equation comes from. And now, I've been a bit sloppy. These are really partial derivatives. But that's, again, an example of lying. I wouldn't worry about whether it's a partial or total derivative at the beginning, because that's not the fundamental idea.

The fundamental idea is that it's curvature on this side connecting to acceleration on this side, and they're connected by a positive sign and by something that has dimensions of speed squared, which turns out to be the speed at which the wave moves. So that's a way to introduce yet another equation, not related to the history, but actually connecting to the intuitive approach. Any questions about that?