

**Dwarkesh Podcast #60 - Grant Sanderson (@3blue1brown) - Past, Present, & Future of  
Mathematics**

Published - October 12, 2023

Transcribed by - [thepodtranscripts.com](https://thepodtranscripts.com)

**Dwarkesh Patel**

Today I have the pleasure of interviewing Grant Sanderson of the YouTube channel, 3blue1brown. You all know who Grant is and I'm really excited about this one.

By the time that an AI model can get gold in the International Math Olympiad, is that just AGI? Given the amount of creative problem solving and chain of thought required to do that.

**Grant Sanderson**

To be honest, I have no idea what people mean when they use the word AGI. I think if you ask 10 different people what they mean by it, you're going to get 10 slightly different answers. And it seems like what people want to get at is a discrete change that I don't think actually exists. Where you've got, AIs up to a certain point are not AGI. They might be really smart, but it's not AGI. And then after some point, that's the benchmark where now it's generally intelligent.

The reason that world model doesn't really fit is it feels a lot more continuous where GPT-4 feels general in the sense that you have one training algorithm that applies to a very, very large set of different kinds of tasks that someone might want to be able to do. And that's cool. That's an invention that people in the sixties might not have expected to be true for the nature of how artificial intelligence can be programmed.

So it's generally intelligent, but maybe what people mean by "Oh, it's not AGI." is you've got certain benchmarks where it's better than most people at some things, but it's not better at most people than others.

At this point, it's better than most people at math. It's better than most people at solving AMC problems and IMO problems. It's just not better than the best. And so maybe at the point when it's getting gold in the IMO, that's a sign that, "Okay, it's as good as the best." And we've ticked off another domain, but I don't know, is what you mean by AGI that you've enumerated all the possible domains that something could be good at and now it's better than humans at all of them?

**Dwarkesh Patel**

Or enough that it could take over a substantial fraction of human jobs. It's impressive right now but it's not going to be even 1% of GDP. But in my mind, if it's getting gold in IMO, having seen some of those problems from your channel, I'm thinking "Wow, that's really coming after podcasters and video animators."

**Grant Sanderson**

I don't know. That feels orthogonal because getting a gold in the IMO feels a lot more like being really, really good at Go or Chess. Those feel analogous. It's super creative. I don't know chess as well as the people who are into it, but everything that I hear from them, the

sort of moves that are made and choices have all of the air of creativity. I think as soon as they started generating artwork, then everyone else could appreciate, "Oh, there's something that deserves to be called creative here."

I don't know how it would look when people get them to be getting golds at the IMO but I imagine it's something that looks a little bit like how AlphaGo is trained, where you have it play with itself a whole bunch. Math lends itself to synthetic data in the ways that a lot of other domains don't. You could have it produce a lot of proofs in a proof checking language like Lean, for example, and just train on a whole bunch of those. And ask, is this a valid proof? Is this not a valid proof? And then counterbalance that with English written versions of something.

I imagine what it looks like once you get something that is solving these IMO level things, is one of two things. Either it writes a very good proof that you feel is unmotivated, because anyone who reads math papers has this feeling that there are two types. There are the ones where you morally understand why the result should be true and then there are the ones where you're like, "I can follow the steps. Why would you have come up with that? I don't know. But I guess that shows that the result is true." And you're left wanting something a little bit more.

And so you could imagine if it produces that to get a gold in the IMO, is that the same kind of ability as what is required to replace jobs? Not really. The impediments between where it is now and replacing jobs feels like a whole different set of things like having a context window that is longer than some small thing such that you can make connections over long periods of time and build relationships and understand where someone's coming from and the actual problem solving part of it. It's a sign that it would be a more helpful tool, but in the same way that Mathematica can help you solve math problems much more effectively.

### **Dwarkesh Patel**

Tell me why I should be less amazed by it or maybe put it in a different context but the reason I would be very impressed is... With chess, obviously this is not all the chess programs are doing, but there's a level of research you can do to narrow down the possibilities. And more importantly, in the math example, it seems that with some of the examples you've listed on your channel, the ability to solve the problem is so dependent on coming up with the right abstraction to think about it, coming up with ways of thinking about it that are not evident in the problem itself or in any other problem in any other test, that seems different from just a chess game where you don't have to think about what is the largest structure of this chess game in the same way as you do with the IMO problem.

### **Grant Sanderson**

I think you should ask people who know a lot about Go and Chess and I'd be curious to hear their opinions on it because I imagine what they would say is, if you're going to be as good at

Go as AlphaGo is you're also not doing tree search, at least exclusively. It's not dependent on that, because you get this combinatorial explosion, which is why people thought that game would be so much harder for so much longer. There sort of has to be something like a higher level structure in their understanding.

Don't get me wrong, I anticipate being very impressed when you get AIs that can solve these IMO problems, because you're absolutely right, there's a level of creativity involved. The only claim I'm making is that being able to do that feels distinct from the impediments between where we are now and the AIs take over all of our jobs or something. It seems like it's going to be another one of those boxes that's this historic moment analogous to chess and Go, more so than it's going to be analogous to the Industrial Revolution.

**Dwarkesh Patel**

I'm surprised you wouldn't be more compelled.

**Grant Sanderson**

I am compelled.

**Dwarkesh Patel**

Or you just don't think that skill of – this problem is isomorphic to this completely different way of thinking about what's happening in the situation and here's me going through the 50 steps to put all that together into this one proof. I'm surprised you don't think that's upstream of a lot of valuable tasks.

**Grant Sanderson**

I think it's a similar level of how impressed I was with the stable diffusion type stuff, where you ask for a landscape of beautiful mountains, but made out of quartz and gemstones. And it gives you this thing which has all of the essence of a landscape, but it's not literally a landscape. And so you realize that there's something beyond the literal that's understood here. That's very impressive.

In the same way, to solve one of these math problems that requires creativity you can't just go from the definitions. You're 100% right. You need this element of lateral thinking, which is why we find so much joy in finding the solutions ourselves or even just seeing other people get those solutions. It's exactly the kind of joy that you get out of good artistic analogies and comparisons and mixing and matching. I'm very impressed by all of that.

I think it's in the same category. And maybe I don't have the same opinions as a lot of other people with this hard line between pre-AGI and post-AGI. I just don't know what they mean by the word AGI. I don't think that you're going to have something that's this measurable discrete step, much less that a math tournament is going to be an example of what that discrete step would look like.

**Dwarkesh Patel**

Interesting.

Applied mathematicians. Where do we put them in society where they can have the biggest benefit? A lot of them go into computer science and IT and I'm sure there's been lots of benefits there. Where are there parts of society where you can just have a whole bunch of mathematicians go in and they can make things a lot better? Transportation or logistics or manufacturing? But where else do you think they might be useful?

**Grant Sanderson**

That's such a good question. In some ways, I'm like the worst person to ask about that.

This isn't going to answer your question, but instead is going to fan the flames of why I feel it's an important question.

I have actually been thinking recently about if it's worth making an out-of-typical video that's specifically addressed at inspiring people to ask that, especially students who are graduating. Because I think this thing happens when you fall in love with math or some sort of technical field, by default in school, you study that. And when you're studying that, effectively you're going through an apprenticeship to be an expert in that or a researcher in that. The structure of studying physics in a university or math in a university, even though they know that not all majors are going to go into the field. The people that you're gaining mentorship from are academics and our research is in the field. So it's hard not to be apprenticing in that.

And I also have noticed that when I go and give talks at universities or things like this and students come up after and they're saying hi, there's a lot of them like, "Grant, the videos were really inspiring. You're the reason that I studied math. You're the reason I'm going into grad school."

And there's this little bell in the back of my mind that's like, "Cool, cool. I'm amazed. I don't know if I believe that I was wholly responsible for it, but it's cool to have that impact."

But do I want that?

Is this a good thing to get more people going into math PhDs? On the one hand, I unequivocally want more people to self identify as liking math. That's very good. But those who are doing that necessarily get shuffled into the traditional outlets like math academia.

I think you highlighted it very right. Math academia, finance and computer science, data science, something in there in general are very common things to go to. And as a result, they almost certainly have an over allocation of talent. All three of those are valuable, right?

I'm not saying those are not valuable things to go into. But if you were playing God and shifting around, where do you want people to go? Again, I'm not answering your question. I'm just asking it in other words because I don't really know.

I think you should probably talk to the people who made that shift of which there aren't a huge number, but Eric Lander is maybe one good example. Jim Simons would maybe be another as people who were doing a very purely academic thing and then decided to shift to something very different.

Now I have sort of had this thought that it's very beneficial to insert some forcing function that gets the pure mathematicians to spend some of their time in a non pure math setting. NSF grants coming with a requirement that 10% of your time goes towards a collaboration with another department or something like that. The thought being these are really good problem solvers in a specific category of problems and to just distribute that talent elsewhere might be helpful.

When I run this by mathematicians, sometimes there's a mixed response where they're like, "I don't know if we'd be all that useful." There's a sense that the aesthetic of what constitutes a good math problem is by its nature rooted in the purity of it such that it's maybe a little elitist to assume that just because people are really, really good at solving that kind of problem that somehow their abilities are more generalizable than other people's abilities.

Why ask about the applied mathematicians rather than saying shouldn't the applied biologists go and work in logistics and things like that because they also have a set of problem solving abilities that are maybe generalizable.

In the back of my mind I think "No, but the mathematicians are special. There really is something general about math." So I don't have the answers. I will say I'm actually very curious to hear from people for what they think the right answers are or from people who made that switch. Let's say they were a math major or something adjacent like computer science physics. And then they decided that they wanted to pour themselves into something not because that was the academic itch that they were scratching by being good at school and getting to appreciate that. But because they stepped back and said what impact do I want to make on the world?

I'm hungry for more of those stories because I think it could be very compelling to convey those specifically to my audience who is probably on track to go into just the traditional math type fields and maybe there's room to have a little bit of influence to disperse them more effectively.

But I don't know. I don't know what more effectively looks like because at the end of the day I'm like I'm a Math YouTuber. I'm not someone who has a career in logistics or manufacturing

or all of these things in such a way that I can have an in tune feel for where there is a need for this specific kind of abstract problem solving.

### **Dwarkesh Patel**

It might be useful to speculate on how an undergrad or somebody who is a young math whiz might even begin to contemplate — here's where I can have an edge.

I'm actually remembering a former podcast guest Lars Doucet, he was a game designer and he started learning about Georgism which is this idea that you should tax land and only land. And so he got really interested in not only writing about those ideas but also with — well, if you're going to tax land you have to figure out what the value of land is. How do you figure out the value of land? There's all these algorithms of how you do this optimally based on neighboring land and how to average across land. And there's a lot of intricacies there.

He now has a startup where he just contracts with cities to implement these algorithms to help them assess the value of their land which makes property taxes much more feasible. That's another example where the motivation was more philosophical but his specialty as a technical person helped him make a contribution there.

### **Grant Sanderson**

I think that's perfect. Probably the true answer is that you're not going to give a universal thing. For any individual is going to be based on where their life circumstances connect them into something either because he had an interest in Georgism for whatever reason. But if someone I don't know their dad runs a paper mill and they're connected to the family business in that way and realize they can plug themselves in a little bit more efficiently.

You're going to have this wide diversity of the ways that people are applying themselves that does not take the form of general advice given from some podcast somewhere but instead takes the form of simply inviting people to think critically about the question rather than following the momentum of what being good at school implies about your future.

### **Dwarkesh Patel**

We were talking about this before the interview started but we have a much better grasp on reality based on our mathematical tools. I'm not talking about anything advanced. Literally being able to count in the decimal system that even the Romans didn't have.

How likely do you think it is that something that significant would be enjoyed by our descendants in hundreds of thousands of years or do you think that that kind of basic numeracy level stuff those kinds of thinking tools are basically all gone?

**Grant Sanderson**

Just so I understand the question right, you're talking about how having a system for numbers changes the way that we think that then lends itself to a better understanding of the world like we can do commerce, things like that. Or we can think in terms of orders of magnitude that would have been hard to think about. We have the word "orders of magnitude" in a way that is hard to write down, much less think about if you're doing Roman numerals. Is there something analogous to that for our descendents?

Fluency with a programming interface really can help with understanding certain problems. I think when people mess around in a notebook with something and it feels like a really good tool set. There's a way that has the same sensation as adopting a nice notation in that you write something with a small number of symbols but then you discover a lot about the implication of that. In the case of notation, it's because the rules of algebra are very constrained and so when you write something you can go through an almost game-like process to see how it reduces and expands and then see something that might be non-trivial.

And in the case of programming, of course the machine is doing the crunching and you might get a plot that reveals some extra data. I think we're maybe at a phase where there's room for that to become a much more fluid process such that rather than having these small little bits of friction like you've got to set up the environment, and you got to link it in the notebook, you've got to find the right libraries, that there's something that feels as fluid as when you are good at algebra and you're just at a whiteboard kind of noodling it out.

I think there's something to be said for the fact that there's still so much more value in paper. if you and I were going to go into some math topic right now. Let's say you ask me something that's a terrible question for a podcast but I'm like "Oh. Let's actually dig into it." The right medium to do that is still paper. I think I would break out some paper and we would scribble it out.

Whenever it becomes the case that the right medium to do that lends itself to simulation and to programming and all that, that feels like it would get to the point where it shifts the way that you even think about stuff.

**Dwarkesh Patel**

What's up with miracle years? This is something that has happened throughout science and especially with mathematicians, where they have a single year in which they make up many, if not most, of the important discoveries that they have in their career. Newton, Einstein, Gauss they all had these years. Do you have some explanation of what's going on?

**Grant Sanderson**

What's your take?



**Dwarkesh Patel**

I think there's a bunch of possible explanations. It can't just be youth because youth lasts 10 years not one year so it must have something to do with..

**Grant Sanderson**

Every 35 year old right now is like, "How dare you."

**Dwarkesh Patel**

You know what I mean. Maybe 20 years. So yeah, it can't just be that. I don't know there's a bunch of possible things you could say. One is you're in a situation in life where you have nothing else going for you or you're just really free for that one year and then you become successful after that year is over based on what you did.

But what is your take?

**Grant Sanderson**

I don't know. I agree that's probably multiple factors, not one. One thing could be that the miracle year is like the exhalation and there's been many, many years of inhalation.

The classic one is Einstein's where his miracle year were also some of the first papers springing onto the scene, and I would guess that a lot of the ideas were not bumping around his head only in that year but it's many many years of thinking about it and coalescing.

And so you might be in a position where you can build up all of this potential energy and then for whatever reason there's one time in life that lends itself to actually releasing all of that. If I try to reflect on my own history with what I'm doing now I think I didn't appreciate early on how much potential energy I had simply from being a student in college where there's just a bunch of ways of thinking about things, or empathy with new learners, or just cool concepts right? The basic concept behind a video that in fact it was many many years of like all of my time having learned math before I started putting out stuff online that I was able to eat into.

The well never runs dry there's always a long list of things that I want to cover but in some sense like I recognize that the well was at risk of running dry in a way that I never thought that it could and without being a little deliberate about devoting some of my day not just of output and producing but to stepping back and like learning new things and touching something I never would have that doesn't happen by default.

I don't know if this is all also the case for the people who have had genuine miracle years where they were like letting out all of this stuff and then it takes a decade to build up that same level of potential energy.

The other thing is you have everything to gain and nothing to lose when you are young. So even if it's not merely youth, there's a willingness to be creative and there's also none of the obligations that come from having found success before.

There's certain academics who made an extremely deliberate effort not to let the curse of success happen or there's some term for it but I think maybe James Watson had this standard reply to invitations for you know talks and interviews and things like that. It was basically like, "No to everyone because I just want to be a scientist." It was much more articulate than that and he has all these nine points but that was the gist of it.

Short of doing that I think it's very easy for someone to have a lot of other things that eat into their mind share and time and all of that such that even if it's just 20 hours a week, that really interrupts a creative flow.

**Dwarkesh Patel**

Were you a student when you started the channel?

**Grant Sanderson**

Technically, yeah. The very first video was made when I was a senior at Stanford. Basically I had been toying around with just a personal programming project in my last year of college that was the beginnings of what is now the animation tool I work with.

I didn't intend for it to be a thing that I would use as a math YouTuber. I didn't even really know what a YouTuber was. It was really just like a personal project. It was March of that year that I think that I published the first ever video. It was kind of right at that transition point.

**Dwarkesh Patel**

Would you have done it if you had become a data scientist?

**Grant Sanderson**

Data scientist and Math PhD were the two like 50-50 contenders basically.

**Dwarkesh Patel**

Is there a role in which you started doing that but then later on made Manim or do you think that that was only possible in a world where you had some time to kill in your senior year?

**Grant Sanderson**

If the goal was to make Math YouTube videos it would have been a wild thing to do it by making Manim as the method for it because it's so strikingly inefficient to do it that way. At the very least I probably would have built on top of an existing framework. There's so many

things that I would tell my past self if I could go back in time even if the goal was to make that. Certain design decisions that caused pain that could have been fixed earlier on.

But if the goal was to make videos, there's just so many good tools for making videos I probably would have started with those or if I wanted to script things, maybe I would have first learned After Effects really effectively and then learn the scripting languages around after effects that might have even been better for all I know. I really don't know.

I just kind of walked into it because the initial project was to make something that could illustrate certain ideas in math especially when it came to visualizing functions as transformations, mapping inputs to outputs, as opposed to graphing. The video output was just a way of knowing that I had completed that personal project in some sense and then it turned out to be fun because I also really enjoy teaching and tutoring.

Then again there's a lot of other people who make their own tools for math GIFs and little illustrations and things which on the one hand feels very inefficient. If people come across a math GIF on Wikipedia there's a very high probability it comes from this one individual who is just strangely prolific at producing these like Creative Commons visuals and he has his own like home baked thing for how he does it.

And then there's someone I came across on Twitter Matt Henderson who has these completely beautiful math GIFs and such and again it's a very home baked thing. It is built on top of shaders but he kind of has his own stuff there.

Maybe there's something to be said for the level of ownership that you feel once it is your own thing that just unlocks a sense of creativity and feeling like, "Hey. I can just describe whatever I want because if I can't already do it I'll just change the tool to make it able to do that". For all I know, that level of creative freedom is necessary to take on a wide variety of topics but your guess is as good as mine for those counterfactuals.

### **Dwarkesh Patel**

This is personally interesting to me because I also started the podcast in college and it was just off track of anything I was planning on doing otherwise. And this is many, many orders of magnitude away from 3Blue1Brown I don't want the audience to you know cringe in unison, but I just think it's interesting like these kinds of projects how often something later on ends up being successful is something that was started almost on a whim as a hobby when you're in college.

### **Grant Sanderson**

I will say there's a benefit to starting it in a way that is low stakes. You're not banking on it growing. I had no anticipation of much less an expectation of 3Blue1Brown growing. I think the reason I kind of kept doing it was, in the fork of life where I did the Math PhD and all that,

I thought it might be a good idea to have a little bit of a footprint on the internet for Math exposition. I was thinking of it as a very niche thing that maybe some math students and some people who are into math would like, but I could sort of show the stuff as a portfolio, not as an audience size that was meaningful.

I was surprised by what an appetite there was for the kind of things that I was making and in some ways maybe that's helpful because I see a lot of people who jump in with the goal of being a Youtuber. I think it's the most common at desired job among the youth is to be like a tiktoker or a Youtuber, which think of that what you will, but when you jump in with that as a goal you kind of aim for too large an audience and end up making the content which is best for no one because one, you're probably not that good at making videos yet and if it's a generally applicable idea, you're competing with like all of the other communicators out there. Whereas, if you do something that's almost unreasonably niche and also you're not expecting it to blow up it's like one you're not going to be disappointed, it's like outstanding when a thousand people view it as opposed to being disappointing and then two, you might be making something that is the best possible version of that content for the audience who watches it because no one else is making that for them because it's too narrow a target.

The beauty of the internet is that there's an incentive to do that and I don't know if this is the case with your podcast when you're starting out, but not thinking about how can I make this as big as possible actually made it more in depth for those who were listening to it.

### **Dwarkesh Patel**

Is it surprising to you that prehistoric humans don't seem to have had just basic arithmetic and numeracy? To us with the modern understanding that kind of stuff seems so universally useful and so fundamental that it's shocking that it just doesn't come about naturally in the course of interacting with the world. Is that surprising?

### **Grant Sanderson**

You're right that it's so in our bones that it's hard to empathize with not having numeracy. If you think, "Okay. What's the first place that most people think about numbers in their daily lives?" It's linked to commerce and money. Maybe in some ways the question is the same as, is it surprising that early humanity didn't have commerce or didn't deal with money?

Maybe when you're below Dunbar's number in your communities, a tit for tat structure just makes a lot more sense and actually works well and it would just be obnoxious to actually account for everything.

Have you come across those studies where anthropologists interview tribes of people that are removed enough from normal society that they don't have the level of numeracy that you or I do? But there's some notion of counting. You have one coconut or nine coconuts like you have a sense of that. But if you ask what number is halfway between one and nine,

those groups will answer three whereas you or I or people in our world would probably answer five and because we think on this very linear scale.

It's interesting that evidently the natural way to think about things is logarithmically, which kind of makes sense. The social dynamics of as you go from solitude to a group of 10 people to a group of 100 people have roughly equal steps in increasing complexity more so than if you go from 1 to 51 to 102 and I wonder if it's the case that by adding numeracy in some senses we've also like lost some numeracy or lost some intuition in others, where now if you ask middle school teachers what's a difficult topic to teach for students to understand they're like logarithms. But that should be deep in our bones right so somehow it got unlearned and maybe it's in the formal sense that it's harder to relearn it, but there's maybe a sense of like numeracy and a sense of quantitative thinking that humans naturally do have that is hard to appreciate when it's not expressed in the same language or in the same ways.

**Dwarkesh Patel**

Yeah, I have seen the thing from Joseph Henrich where there's still existing tribes where they're in this kind of situation. They can do numeracy and arithmetic when it's in very concrete terms, if you're talking about seeds or something but that the abstract concept of a number is not available to them.

**Grant Sanderson**

Do you think the abstract concept of a number is useful to your life?

**Dwarkesh Patel**

Oh yeah.

**Grant Sanderson**

In what ways?

**Dwarkesh Patel**

It's almost like asking — how is the concept of the alphabet useful? It comes up so often. For example, how many lights do I set up for this interview?

**Grant Sanderson**

Is that the concept of an abstract number though? Because it's like two people, two lights. One to one correspondence.

Did you leverage the abstraction of two as an object which is simultaneously a rational and a real and an integer. Is in the context of a group that has additive structure but also multiplicative. It was just there's light for you light for me.

I'm pretty sure the abstract idea of a number is important for all of us but I don't think it's immediately obvious. It's more that it shapes the way we think, I'm not sure if it actually changes the way we live. Assuming you don't work in STEM right where you literally are using it all the time.

**Dwarkesh Patel**

Yeah, I'm trying to go through my day and think through where am I using them? There's the obvious stuff like the commerce examples you mentioned where you go to a restaurant and you're figuring out what to pay or what to tip but that seems a very particular example.

Do I really use numbers that infrequently? I don't know.

**Grant Sanderson**

Many people listening are probably screaming out of their head with much more apt examples but it's hard to say.

**Dwarkesh Patel**

When a mathematician is working on a problem, what is the biggest mental constraint? Is it the working memory? Is it the processing speed? Plants are limited by nitrogen usually, what is the equivalent of nitrogen for a mathematician?

**Grant Sanderson**

That's a fun question. I'm not a research mathematician, I shouldn't pretend like I am. The right people to ask that question would be the research mathematicians. I wonder if you're going to get consistent answers as with so many things there's not one answer.

Maybe it's the number of available analogies to be able to draw connections? The more exposure you've had to disparate fields such that you could maybe see that a problem-solving approach that was used here might be useful here. Sometimes that's literally codified in the forms of connections between different fields, as functors between categories or something.

But sometimes it's a lot more intuitive. Someone's doing a combinatorics type question and they're like, "Oh. Maybe generating functions are a useful tool to bring to bear." and then in some completely different context of studying prime numbers they're like, "Oh. Maybe it could take a generating function type approach. Maybe you have to massage it to make it work.

One of the reasons I say this is that one of the tendencies that you've seen in math papers in the last 200 years is that the typical number of authors is much bigger now than before. I think people have this misconception that math is a field with lone geniuses who are

coming up with great insights alone next to a blackboard. The reality is that it's a highly collaborative field.

I remember one of the first times that I was hearing from a mathematician, I was a young kid and was in this math circles event and someone was asking this person, "What surprised you about your job?" The first thing he said was how much travel was involved. He wasn't expecting that. And it's because you know if you're studying some very specific niche field, the way that you make progress in that is by collaborating with other people in that field or maybe adjacent to that field and there's only so many that they probably aren't at your university. So you travel a lot to work with them.

These days a lot of that I think happens on Zoom but conferences are still super important and these sorts of events that bring people all under one roof like MSRI, is maybe an example of a place that's trying to do that systematically. You could say that's a social thing but I think it's maybe hitting on this idea that what you want is exposure to as many available analogies. So the short answer to your question, what is nitrogen for mathematicians, is the analogy.

### **Dwarkesh Patel**

This actually is an interesting question I wasn't planning on asking you but it just occurred to me.

Is it surprising how new a lot of mathematics is? Even mathematics that is taught at the high school level. Whereas with physics or biology, that's also new but you can tell a story where we didn't have the tools to look at the cell or to inspect an electron until very recently but we've had mathematicians for 2000-3000 years, who were doing pretty sophisticated things, even the ancient Greeks. Why is linear algebra so new given that fact?

### **Grant Sanderson**

I wouldn't have thought of math as being new in that way, especially at the high school level. I remember there's always a sensation that it's frustrating that all of the things are actually way more than a hundred years old, in terms of the names attached to the theorems that you're doing, none of them are remotely modern.

Whereas in biology, the understanding we have for how proteins are formed is relatively much more modern and you might be just a couple generations away. To some extent there's a raw manpower component to it. How many people did pure math for most of history? For most of history, no one. No one was a pure mathematician. They were a mathematician plus something else or they were a physicist or they were a natural philosopher. And in so far as you're doing natural philosophy, one component of that is developing math but it's not the full extent of what you do.

Even the ones who we think of as like very, very pure mathematicians in the sense that a lot of their most famous results are pure math like Gauss, actually a lot his output was also centered on very practical problems,

Maybe since then is when you start to get an era of something more like pure mathematicians. The raw number available that you have the man hours that are being put into developing new theorems, is probably just got this huge spike as the population grows and then also the percentage of the population that has the economic freedom to do something as indulgent as academia grows.

Maybe it's pretty reasonable that most of it is much, much more recent. That would be my guess.

### **Dwarkesh Patel**

Some of these things seem actually pretty modern like information theory. It is less than 100 years old and is pretty fundamental. Theoretically, you could have written that paper a long time ago.

### **Grant Sanderson**

That's a really good example and maybe this is a sign that the math that's developed is more in the service of the world that you live in and the adjacent problems that it's used to solve than we typically think of it. On the one hand information theory sets a good example because it's so pure that you could have asked the question, you could have defined the notion of a bit, but evidently there wasn't a strong enough need to think in that way. Whereas when you're doing error correction or you're thinking about actual information channels over a wire and you're at Bell Labs, that's what prompts it.

Another maybe really good example for that would be Chaos theory. You could easily ask why is Chaos theory so recent? You could have written the Lorenz equations since differential equations existed. Why didn't anyone do that and understand that there was this sort of sensitivity to initial conditions?

In that case it would maybe be the opposite, where it's not that you need the existence of computers as a problem to solve or the problems that they introduce are the problems to solve but instead you need them to even discover the phenomenon in the first place.

A lot of original concepts in chaos theory came from basically running simulations or doing things that required a massive amount of computation that simply wouldn't be done by hand. Someone could ask the question but they wouldn't have observed the unexpected phenomenon and there, even if it's questions that are as relevant to a pre-computer world as to a post-computer world like the nature of weather modeling, or just the nature of



three-body problem, all of that kind of stuff, somehow without the right tools for thought it just didn't come into the mind.

So yeah maybe there's other things like that where those questions or pieces of technology that start to fundamentally shape everyone's life will then invariably also shift the mathematician's focus.

### **Dwarkesh Patel**

This actually reminds me of the first day of Scott Aaronson's quantum information class. He said, "What I'm about to describe to you could have been discovered by a mathematician before quantum physics existed. If only they had asked the question of we're going to do probabilities but we're only allowed to use unitaries." They could have just discovered quantum mechanics or quantum information from there.

### **Grant Sanderson**

The thing about math, especially if you're talking about pure axiomatized math, the experience as an undergrad is that you are going through a textbook and it starts with saying here's the axioms of this field and then we're going to deduce from those axioms various different lemmas and theorems and proceed from that.

With that as the framing you get the impression that you could have just come up with any axioms. Just make up some pile of axioms, deduce what follows from them and the space of possible math is unfathomably huge. So you need some process that culls down what are the useful things to maybe pursue.

So one of the things that I think is all too often missing in those pure math textbooks is the motivating problem. Why is it that this was the set of axioms people found to be useful and not something else? The framework for quantum information theory, you married together linear algebra and probability that's great, but there's all sorts of other things where you could kind of try to cram them together and maybe get some sort of math out.

The question becomes is it worth your time to do that?

Knot theory is something that emerged because Lord Kelvin had a theory that all of the elements on the periodic table had structures which were related to a knot. A knot being if you have a closed loop in 3D space but if you wanted to continuously deform it without it ever crossing itself, you ask the question Could you get back to say an open loop? Or if you can't get back to an open loop, what are the set of all other loops in 3d space that could be deformed into that? And you end up categorizing what all the different knots are.

This was started with a completely incorrect theory for what's going on at the atomic level that gives atoms this very stable structure because I think he found with smoke rings like if

you're somehow very dexterous, you can get them to form knots in 3D and they're very stable. In that it'll never cross over itself. So it has all those properties now that was irrelevant for understanding the periodic table but it was an interesting mathematical question and people kind of ran with it and in that case it was an arbitrary reason that someone thought to ask the question and then some people ran with it and frankly it's probably fewer people who run with it than would if it turned out to be a more useful question.

So really, you want to ask what are the things that prompt people to ask what turns out to be a mathematical question given that the space of what would be mathematical questions is so unfathomably huge that it's just impossible to explore it through a random walk.

**Dwarkesh Patel**

Wait, are you saying that Lord Kelvin's apple story was that he was smoking a lot of pipe and he categorized his puffs. [Laughter]

You and other creators have changed how pedagogy happens via animated videos. What would it take to do something similar for video games, text, and all these other mediums? Why hasn't there been a similar sort of broad-scale adoption and transformation of how teaching happens there?

**Grant Sanderson**

I'm not sure I understand the question. You're saying where there's been a rise of explanatory videos, why is there not a similar rise in pedagogical video games?

I don't play enough games so I can't really speak to it in the way that well-versed game designers can but one thing to understand is that games are very hard to make. It takes a lot of resources for a given game and whenever people seem to try to do it with pedagogy as a motive it seems to be the case that they are not fun in the way that people would want them to be fun and then the ones that are actually most effective are not as directly educational.

The one game that I actually have played because enough of my friends told me hey you should really do this it seems relevant to Math explanation in the last like decade was The Witness. Have you played it?

**Dwarkesh Patel**

I've heard about it.

**Grant Sanderson**

As someone who doesn't play games and then did play it, it's fantastic. It's absolutely well done in every possible way that you could want something to be well done. Critical on that is the nature of how problems are solved.

The reason people are recommending it to me is because the feeling of playing the witness is a lot like the feeling of doing math. It's non-verbal, you come across these little puzzles where the simple mechanics of one puzzle inform you about the fundamental mechanics that become relevant to much much harder ones such that if you do it with the right sequence you have the feeling of epiphany in ways that are very self-satisfying.

You come away feeling like you should be able to do something like this for math and maybe you can. It's just that it's so hard to make a game at all that there's just not the rate of production that you would need to explore to get enough games out there that one of them hits.

There's a lot of math videos on YouTube. It's okay that most of them suck. It's okay because you just need enough that when someone searches for the term that they want they get one that is good and scratches that itch. Or that you know they might get recommended something that is bringing a question to their mind that they wouldn't have thought about but they become really interested once it's there. Whereas with video games, you're also spending a lot more time as a user on each one. Rather than a five minute average experience it's a many, many hour average experience.

You ask the same question on text. I don't know if I accept the premise that there's not the same advances and innovation in the world of textual explanations. Mathagon is a really, really good example of this. It's like the textbook of the future. It's basically an interactive textbook. The explanations are really good. In so far as it doesn't have more of an impact or more of a reach it's maybe just because people don't know about it or don't have an easy means of accessing something that recommends to them like the really good innovations happening in the world of textual explanations in the way that youtube has this recommending engine that tries its hardest to get more of these things in front of people.

In the world of actual written textbooks, there's so many that I like so much that I think it would be a disservice to talk about that medium as not making advances in terms of more and more thought put towards empathy to the learner and things like that.

### **Dwarkesh Patel**

Should the top 0.1% of educators exclusively be on the internet because it seems like a waste if you were just a college professor or a high school professor and you were teaching 50 kids a year or something. Given the greater scale available should more of them be trying to see if they can reach more people?

### **Grant Sanderson**

I think it's not a bad thing for more educators who are good at what they're doing to put their stuff online for sure. I highly encourage that even if it's as simple as getting someone to put

a camera in the back of the classroom. I don't think it would be a good idea to get those people out of the classroom.

If anything I think one of the best things that I could do for my career would be to put myself into more classrooms. Actually I'm quite determined at some point to be a high school math teacher for some number of years. There's such an opportunity cost that it's probably something I would plan on notably later as long as there's not other life logistics that occupy a lot of mind share because everything I know about high school teaching is like it just kicks your ass for the first two years.

One of the most valuable things that you can have if you're trying to explain stuff online is a sense of empathy for what possible viewers that are out there. The more distance that you put between yourself and them in terms of life circumstances. I'm not a college student so I don't have the same empathy with college students. Certainly not a high school student, so I've lost that empathy. That distance just makes it more and more of an uphill battle to make the content good for them and I think keeping people in regular touch with just what people in the classroom actively need is necessary for them to remain as good and as sharp as they are.

So yes, get more of those top 0.1% to put their stuff online but I would absolutely disagree with the idea of taking them out of their existing circumstances. Maybe for a year or two so they don't lose that sharpness but then put them right back in because it makes them better at the online exposition.

The other thing I might disagree with is the idea that the reach is lower. Yes, it's a smaller number of people but you're with them for much, much more time and you actually have the chance of influencing their trajectory through a social connection in a way that you just don't over Youtube.

You're using the word education in a way that I would maybe sub out for the word explanation. You want explanations to be online but the word education derives from the same root as the word educe, to bring out, and I really like that as a bit of etymology because it reminds you that the job of an educator is not to like take their knowledge and shove it into the heads of someone else the job is to bring it out. That's very, very hard to do in a video and in fact even if you can kind of get at it by asking intriguing questions for the most part the video is there to answer something once someone has a question.

The teacher's job, or the educator's job, should be to provide the environment such that you're bringing out from your students as much as you can through inspiration through projects through little bits of mentorship and encouragement along the way that requires you know eye contact and being there in person and being the true figure in their life rather than just an abstract voice behind a screen.

**Dwarkesh Patel**

Then should we think of educators more as motivational speakers? As in the actual job of getting the content in your head is maybe for the textbooks or for Youtube but why we have college classes or high school classes is that we have somebody who approximates Tony Robbins to get you to do the thing.

**Grant Sanderson**

That would be a subset of it but there's more than just motivational speech that goes into it. There's um facilitation of projects or even coming up with what the projects are or recognizing what a student is interested in so that you can try to tailor a question to their specific set of interests or you can maybe act as the curator. Where, "Hey, there's a lot of online explanations for what a Poisson distribution is. Which of these is the right one that I could serve?" and based on knowing you as a particular student what might resonate. You might be in a better position to do that. All of that goes beyond being a Tony Robbins saying, "Be the best person that you can be." and all of that.

One thing I might say is that anytime that I'll chat with mathematicians and try to get a sense for how they got into it and what got them started, so often they start by saying there was this one teacher and that teacher did something very small – like they pulled them aside and just said, "Hey. You're really good at this. Have you considered studying more?" or they give them an interesting problem.

And the thing that takes at most 30 minutes of the teacher's time, maybe even 30 seconds, has these completely monumental rippling effects for the life of the student they were talking to that then sets them on this whole different trajectory.

Two examples of this come to mind. One is this woman who was saying she had this moment when she got pulled aside by the teacher and he just said, "Hey, I think you're really good at math. You should consider being a math major." which had been completely outside of her purview at that time. That changed the way she thought about it. And then later she said she learned that he did that for a large number of people. He just pulled them and was like, "Hey, you're really good at math." So that's a level of impact that you can have as a figure in their lives in a way that you can't over screen.

Another one which was very funny. I was asking this guy why he went into the specific field that he did. It was a seemingly arbitrary thing in my mind but I guess all pure math seems to be. He said that in his first year of grad school he was sitting in this seminar and at the end of the seminar the professor, who was this old professor who he had never met him before, they didn't have any kind of connection. He seeks this guy out and comes up and he says, "You. I have a problem for you. A good research problem that I think I think might be a good place for you to start in the next couple months" and this guy was like "Oh, okay" and he gets this research problem and he spends some months thinking about it and he comes back

and then it later came to light that the professor mistook him for someone else that was someone he was supposed to be mentoring. He was just the stereotypical image of like a doddering old math professor who's not very in tune with the people in his life that was the actual situation but nevertheless that moment of accidentally giving someone a problem completely shifted the research path for him, which if nothing else, shows you the sensitivity to initial conditions that takes place when you are a student and how the educator is is right on that nexus of sensitivity who can completely swing the fences one way or another for what you do.

For every one of those stories there's going to be an unfortunate counter balancing story about people who are demotivated from math. I think this was seventh grade. There was this math class that I was in and I was one of the people who was good at math and enjoyed it and would often help the people in the class understand it. I had enough ego built up to have a strong shell around things. For context, I also really liked music and there was this concert that had happened where I had a certain solo or something earlier in that week.

There was a substitute teacher one day who didn't have any of the context and she gave some lesson and had us spend the second half of the class going over the homework for it. All of the other students in the class were very confused and I think I remember like they would come to me and I would try to offer to help them and the substitute was going around the class in these circles and basically marking off a little star for how far down the homework people were just to get a sense are they progressing. That was kind of her way of measuring how far they were. When she got to me I had done none of them because I was spending my whole time trying to help all of the others and after having written a little star next to the same problem like three different times she said to me like, "Sometimes music people just aren't math people." and then keeps walking on.

I was in the best possible circumstance to not let that hit hard because one, I had the moral high ground of "Hey, I've just been helping all these people. I understand it and I've been doing your job for you." This was my little egotistical seventh grade brain. I knew that I knew the stuff. Even with all of the armor that was put up, I remember it was just this shock to my system, she says this thing and it just made me strangely teary-eyed or something.

I can only imagine if you're in a position where you're not confident in math and the thing that you know deep in your heart is actually you are kind of struggling with it, just a little throwaway comment like that could completely derail the whole system in terms of your relationship with the subject.

So it's another example to illustrate the sensitivity to initial conditions. I was in a robust position and wasn't as sensitive. I was gonna love math no matter what but you envision someone who's a little bit more on that teetering edge and the comment, one way or another, either saying you're good at this you should consider majoring in it or saying,

"Sometimes music people aren't math people" which isn't even true. That was the other thing about it that niggled at my brain when she said it.

All of that is just so important for people's development that when people talk about online education as being valuable or revolutionary or anything like that, there's a part of me that sort of rolls my eyes because it just doesn't get at the truth that online explanations have nothing to do with all of that important stuff that's actually happening and at best it should be like in the service of helping that side of things where the rubber meets the road.

### **Dwarkesh Patel**

I had Tyler Cowen on the podcast and he obviously has Marginal Revolution and these Youtube videos where he explains economics and he had a similar answer to give. I asked him, should we think of you as a substitute for all these economics teachers? And in his mind as well he was more a complement to the functions that happen in the class.

And to your point about the initial conditions, I'm sure you remember the details of the story but I just vaguely remember hearing this, wasn't there a case where a mathematician who later ended up becoming famous? He arrives late to a lecture.... Do you want to tell the story

### **Grant Sanderson**

I don't remember it beat for beat but I think it was a statistics class and he was a grad student and he comes in late and there's two problems on the board that the professor had written. He assumed that those two problems were homework and so he goes home and works on them and after a couple weeks he goes to the professor's office and turns in his homework.

He's like, "I'm sorry. I'm so late. This one just took me a lot longer than some of the others." And the professor's like "Oh, okay." and just shuffles it away. Then a couple days later when the prophet had the time to like go through and see them, he realized that the student had fully answered these questions, what the student didn't know is that they were not homework problems written on the chalkboard they were two unsolved problems in the field that the prof put up as examples of what the field was striving.

I don't remember what problems they were so that would be more fun color to add to the story but then as the anecdote told to me however many years ago goes, the prof then finds the students' housing and knocks on the door "Do you realize that these were actually unsolved problems?" and then he gets to basically make those his thesis. So yeah, that idea of just being given something for completely random reasons and it shifts the course of what you do.

**Dwarkesh Patel**

It's the thing where if you know a crossword is solvable, you just keep going at it until you solve it.

**Grant Sanderson**

Or the four-minute mile, right?

**Dwarkesh Patel**

Exactly. That's a great example.

Another valuable experience, at least one I had, was taking Aaronson's classes in college and realizing I am at least two standard deviations below him and that was actually a really valuable experience for me not because it increased my confidence in I didn't have a moment where I was like, "Oh wow. I'm good at this" but it was useful to know. Podcasting is an easier thing to do right so then it's good to know that there are actual technical things out there where knowing that you can get really deep into something and people are just gonna be way above you having that sort of awareness.

**Grant Sanderson**

Do you think it's fair to have a mental model that has a static g-factor type quality here such that your two standard deviations below and that is forever the state of things? Or do you think that the right mental model is something that allows for flexibility on where contributions actually come from, or where intuitions come from. That through many years of experience in certain kinds of problem solving maybe what seemed like a flash of insight was actually like the residue of just years of thinking about certain kinds of puzzles that he had, that you maybe didn't.

**Dwarkesh Patel**

Can I tell you a story from that class actually?

**Grant Sanderson**

Yeah, go for it.

**Dwarkesh Patel**

He was giving a proof of a very important method in complexity theory that helped to prove the bounds of the complexity of different problems and he explains it and he says, "You know, in 1999, I approved this myself but I realized that six months before somebody had already published a paper with this method and I realized I'm catching up to the frontier now. But when I was a kid I was doing Euler, that's 2000 years in regress. Now I'm six months behind."



And then so later on in the day I'm like, "Wait, 1999. How old was Scott Aaronson in 1999?" and I think he was 18 or 19 and he was basically proving frontier results in complexity theory. At that point you're like, "All right. Aaronson's a special animal here."

**Grant Sanderson**

You are right. He's probably a special animal.

**Dwarkesh Patel**

But it's just broadly good to have that sort of upper constraint on your Dunning-Kruger that this exists in the world.

**Grant Sanderson**

Maybe the thing that I would want to say is that whatever the scale is on which he's two standard deviations above you, that might not be the one scale that matters and that contributions to these fields don't always look like genius insights and that sometimes there's fruit to be born from say becoming kind of an expert in two different things and then finding connections between them. The people who make contributions are not necessarily the Scott Aaronson's of the world. Still. You are probably right it is true that there are people like that. Von Neuman's another example of one of these, right?

**Dwarkesh Patel**

How much does Godel's Incompleteness theorem practically matter? Is it something that comes up a lot or is it just an interesting thing to know about the bounds that isn't applicable day to day?

**Grant Sanderson**

You've asked me another question where I'm not the best one to answer and I should throw that as a caveat to begin. From what I understand, it really doesn't come up.

The paradoxical fact that it's conveying, the idea that you can't have an axiom system that is both that will basically prove all of the things that are true and which is also self consistent. The contradiction that you construct out of that has the same feeling as the sentence. "This statement is a lie." We think about the statement. If it's false, then it must be true. If it's true, it must be false. It's that same flavor. And you might ask, does the existence of that paradox mean that it's hard to speak English. [Laughter] It's so rare that you would come up with something that happens to have a bit of self reference in it.

One of the first times that there was something that came up that didn't feel quite as pathological in that way, if the curious listener wants to go into it, that search term would be Paris Harrington theorem .

It's a little pathological. wasn't the really question that came up that didn't seem like it was deliberately constructed to be one of these self-referential things where, you know, it shows itself to be outside the bounds of whatever axiom system you were starting with. It was shown to be unresolvable in a certain sense. But it was asking a... I don't want to say natural because a lot of these math questions aren't natural. It was asking a question where you wouldn't expect that to be true.

So maybe at the edges of theory, there are sometimes when the paradoxes that are possible, show. The impression I get is that no mathematician is thinking about it. They're not actively worrying about it. It's not like "Oh god, can I be sure that the stuff that I'm going to show is true."

For all the practical problems like the Riemann hypothesis or twin primes, almost everyone's like, "No, there's going to be an answer." It may be that they turn out to be unresolvable in one of these ways but there's just a strong sense that that theorem came from a pathology in a way that natural questions that people actually care about don't.

### **Dwarkesh Patel**

That's really interesting that something from the outside and in popularizations seems to be a very fundamental thing where people have definitely heard about this.

A good analogy here is the halting problem in computer science. One of the first things you learn in a computer science course is the proof of the halting problem and it's another one of those things where you don't really need to be able to prove that you have that sort of program available.

### **Grant Sanderson**

No more comments.

### **Dwarkesh Patel**

Why are good explanations so hard to find, despite how useful they are? Obviously, other than you, there's many other cases of good explanations. But generally, it just seems like there aren't as many as there should be. Is it just a story of economics where it's nobody's incentive to spend a lot of time making good explanations? Is it just a really hard skill that isn't correlated with being able to come up with a discovery itself? Why are good explanations scarce?

### **Grant Sanderson**

I think there's maybe two explanations.

The first less important one is going to be that there's a difference between knowing something and then remembering what it's like not to know it. And the characteristic of a

good explanation is that you're walking someone on a path from the feeling of not understanding up to the feeling of understanding.

Earlier, you were asking about societies that lack numeracy. That's such a hard brain state to put yourself in, like what's it like to not even know numbers? How would you start to explain what numbers are? Maybe you should go from a bunch of concrete examples. But like the way that you think about numbers and adding things, it's just you have to really unpack a lot before you even start there.

And I think at higher levels of abstraction, that becomes even harder because it shapes the way that you think so much that remembering what it's like not to understand it. You're teaching some kid algebra and the premise of like a variable. They're like, "What is X?" It's not necessarily anything but it's what we're solving for. Like, yeah, but what is it? Trying to answer "What is X?" is a weirdly hard thing because it is the premise that you're even starting from.

The more important explanation probably is that the best explanation depends heavily on the individual who's learning. And the perfect explanation for you often might be very different from the perfect explanation for someone else. So there's a lot of very good domain specific explanations. Pull up in any textbook and like chapter 12 of it is probably explaining the content in there quite well, assuming that you've read chapters one through 11, but if you're coming in from a cold start, it's a little bit hard.

So the real golden egg is like, how do you construct explanations which are as generally useful as possible as generally appealing as possible? And that because you can't assume shared context, it becomes this challenge. And I think there's like tips and tricks along the way, but because the people that are often making explanations have a specific enough audience, it is this classroom of 30 people. Or it's this discipline of majors who are in their third year. All the explanations from the people who are professional explainers in some sense are so targeted that maybe it's the economic thing you're talking about. There's not, or at least until recently in history, there hasn't been the need to or the incentive to come up with something that would be motivating and approachable and clear to an extremely wide variety of different backgrounds.

### **Dwarkesh Patel**

Is the process of making your videos, is that mostly you?

### **Grant Sanderson**

Yes.

**Dwarkesh Patel**

Given the scale you're reaching, it seems that if it was possible, a small increase in productivity would be worth an entire production studio. And it's surprising to me that the transaction cost of having a production setup are high enough that it's better to literally do the mundane details yourself.

**Dwarkesh Patel**

I mean, this could honestly just be a personal flaw. I'm not good at pulling people in and then I've struggled to do this effectively in the past. But a part of it is that the seemingly mundane details are sometimes just how I even think about constructing it in the first place.

The first thing that a lot of YouTubers will do if they can hire is hire an editor. And this will be because they film a lot of things. And so a lot of the editing process is removing the stuff that was filmed that shouldn't be in the video and just leaving the stuff that should be in the video. And that's time consuming and it's kind of mundane. And it's probably not that relevant to what the creator should be thinking about.

The editing process for me, I start by laying out all of the animations and stuff that I want in a timeline and then once I record the voiceover, the actual editing is like a day. I guess I could hire someone and gain a day back of my life but the communication back and forth for saying what specifically I want, all of the little cuts that I'm making along the way are my way of even thinking about what I want the final piece to be and are such that it would be hard to put it into words.

It's similar for why I maybe find it quite hard to use Co-pilot and some of these LLM tools for the animation code. It can be super great if you're learning some new library and it knows about that library that you don't. But for my library that I know inside it out, if I'm just using it, it feels like, "Oh. This should be the most automatable thing ever. It's just text." I should be the first YouTuber who can actually do this better because the substance behind each animation is text, it's not like an editing workflow in quite the same way.

But it doesn't work. And I think it's because maybe it's just because you need a multimodal thing that actually understands the look of the output. Like the output isn't something that is consumable in text. It's something about how it looks.

But at a deeper level, I can't even put into words what I want to put on the screen, except to do so in code. That's just the way that I'm thinking about it. And if I were to try to put into English the thing that I want as a comment that then gets expanded, that task is actually harder than writing it in the code. And if it's clunky to write in code, that's a sign that I should change the interface of the library such that it's less clunky to be expressive in the way that I want.

And it's in that same way where a lot of the creative process that feels mundane, those are just the cogs of thought slowly turning in a way that if they weren't turning for that part, they would have to be turning during the interface of communication with a collaborator.

### **Dwarkesh Patel**

On the point of working with Co-pilot where we can visualize the changes you wanted to make. The Sparks of AGI paper from Microsoft Research had an actually really interesting example where it was generating LaTeX and they generated some output and they say "Change this so that the visual that comes up in the rendering is different in this way." And it was actually able to do that, which was their evidence that it can understand the higher level visual abstraction. I guess it can't do that for Manim.

### **Grant Sanderson**

There's a couple reasons why it might not be as fair a comparison. There are two versions of Manim. There's a community version that is by the community for the community and then mine, the interfaces are largely similar. The rendering engines are quite different, but because of slight differences in that and it might have a tendency to learn from one or its examples from one and it's intermixing them. So stuff just doesn't quite run when there's discrepancy.

Maybe I shot myself in the foot because I don't really comment my code that much for my videos. It's like a one and done deal. The way that I'm making it feels much more like the editing flow. If you were to look at the operation history of someone in After Effects.

It's a little bit more like that where there's not a perfect description in English of the thing that I want to do and then the execution of that. It's just the execution of that.

It's not meant to be editable in hindsight as much because I'm just in the flow of making the scene for the one video. Maybe I could have given it a better chance to learn what it's supposed to be happening by having a really well documented set of — This is the input. This is the output. This is the comment describing it in English. But even then that wouldn't hit the problem. I would have to articulate what the thing I want is in the first place. And the program language is just the right mode of articulation in the first place.

### **Dwarkesh Patel**

This is something I was really curious about ever since I learned about it. I watched many of the Summer of Math Exposition prize videos and it was shocking to me how good they were. Many of them looked like entire production studios were dedicated to making them. And it was shocking to me that you could motivate and elicit this quality of contribution given the relatively modest prize pool, which was like five winners, \$1,000 each.

What is your explanation of just running prizes like this? Why were you able to get such high quality contributions? Is the prize pool irrelevant? Is it just about your reputation and reach?

### **Grant Sanderson**

I do wonder how relevant the prize pool is. We've been thinking about this because we did it first in 2021 and then we plan to continue doing it annually. If I was a mover and a shaker, I probably could raise much more if I wanted to get a big prize pool there. I don't think it would change the quality of the content because the impression I get is that people aren't fundamentally motivated by winning some cash prize.

Certainly, they're not investing that time with an expected value calculation. If they are, that's a terrible, terrible plan. And if anything, a higher prize pool might be a problem. Let's say it was a hundred thousand dollar prize for each of the winners, then it would be a real problem where someone would, and people do, delusionally think that they're very likely going to be the winner and they might actually pour a lot of their own resources into it with the expectation of gaining it. And then that's just a messy situation. I don't want to be in a situation where someone asks "Why wasn't mine chosen as a winner?!" Because the whole event is not supposed to be about winners.

Maybe for the listeners who don't know, I should describe the summer of math exposition.

Actually, the history is a little bit funny because it started with an intern application where in 2021 I wanted a couple interns to do a certain thing on my website basically and I put out a call for people to apply. I got 2,500 applicants and somewhere in the application I mentioned that during the summer, in addition to the main task I wanted them to do, I'd give them freedom to do something relevant to math exposition online that was their own thing and that I'd be happy to provide some mentorship or just give them the freedom to do that one day a week. And I asked them to give me a little pitch on what their idea would be.

As I went through all of the applications, which was a lot, I felt so bad because so often the person would have a little pitch and like what they would want to make. And in my mind, I think, "Cool. You should make that! You don't need me to do that. Just spend your summer making that." Why not?

And people were clearly inspired by the thought of adding something and like I said earlier, being a youtuber is the most common job aspiration among the youth these days. And so as a consolation of sorts to those 99% that I had to reject for the internship, I said we're going to host this thing called the Summer of Math Exposition where we'll give you a deadline. I'll promise to feature five of you in a video. And if you feel like the thing that you were going to do, like with me as your 20 project as an intern is something you're excited about, make it a hundred percent project. Just do it anyway and like I can give you this little carrot in the form of featuring it in a video and give you a deadline, which let's be honest is what actually

makes the difference between people doing something and procrastinating on it sometimes.

Brilliant.org said they would be happy to put some cash prizes in. So I said, sure, why not? I don't think the cash prize is super important, but it's nice. It shows that someone actually cared and put some real thought into doing something that wasn't just a made up gold star, but they put some material behind saying that you were selected as a winner of this thing.

But all in all, it was never supposed to be about choosing winners. It was just to get more people to make stuff. And if anything, I'd actually I love it when I see stuff from existing educators and teachers where it's maybe not the youth who want to be youtubers pouring their hearts and souls into it, but it's the educator who built a lot of intuition over the course of their career for what constitutes a clear explanation and they're just sharing it more broadly.

So, to your question on – What is it that caused there to be such high production quality in some of the entries there? Part of the answer might just be that like tooling is so good now that individuals can actually make pretty incredible things sometimes.

### **Dwarkesh Patel**

I misphrased if I said production quality, I just meant the whole composition as a whole.

### **Grant Sanderson**

Yeah, well there's a selection filter too, right? In that first year, there were 1200 submissions and I featured five of them in the winning video. So of course, they're necessarily unrepresentative of the norm by the very nature of who I was choosing to feature.

### **Dwarkesh Patel**

But the fact that something that high quality was even in the pool.

### **Grant Sanderson**

I think it hits a little bit to your miracle year point where I think what might be happening is you have people with a ton of potential energy for something that they've kind of been thinking about making for a long time. And the hope was to give people a little push. Here's a deadline. Here's a little prize. Here's a promise that maybe if you make it, it won't just go into the void, but there's a chance that it could get exposed to more people, which I think is absolutely played out.

And not for the reason that someone might expect where I choose winners and I feature those winners and people watch them. A huge amount of viewership happens before I even begin the process of looking at them. And this was an accident too, where in this first year, we got 1200 submissions. I said expect judges who are reviewing it to spend at most 10

minutes on each piece. So it could be longer, but don't rely on someone watching it for more.

But realistically, when I'm reviewing something, I want to watch the whole piece. I absolutely do not have time to watch that many. I've learned it takes me about two weeks of just full time work to watch 100 of these pieces and give the kind of feedback that I want.

To manage that problem of more than we could manually review, we put together this peer review system that would basically have an algorithm feed people pairs of videos.

And they would just say which one is better and then it would feed them another one. And in the first two years, we just used a tool that was common for hackathons that did this. And what that did is one, it gave us a partially ordered list of content by quality loosely. We didn't need it to be perfect. We just needed there to be a very high chance that the five most deserving videos were visible somewhere in that top 100.

So there the algorithm doesn't have to be perfect.

A thing I've learned about the YouTube algorithm is — in theory, you would want to just use machine learning for everything. You have some massive neural network where on the input of it, it's got five billion videos or however many exist. And the output decides what seven are best to recommend to you. That is completely computationally infeasible.

I think this is all public knowledge. What you have to do instead is use some sort of proxies as a first pass to nominate a video to even be fed into the machine learning driven algorithm. So that you're only feeding in like a thousand nominees.

So the real difference that it can make if you've made a really good video, between it getting to the people who would like it and not getting there. It's not the flaws in the algorithm. The algorithm is probably quite good. It's the mismatch between the proxies being used to nominate stuff to see whether it's even in the running.

One of the things used for nomination is understanding the co-watch graph where if you've watched video A and you've also watched video B and then I watch video A. Your watching both of those gives a little link between them, or maybe you and a ton of other people watching both of them gives a little link between such that once I watch video A, B is potentially nominated in that phase because it's recognized that there's a lot of co-watching.

That's something that I'm sure is still quite challenging to do scale but it's more plausible to do at scale than like running some massive neural network. And so I think what might have happened is that by having a bunch of co-watching happening on this same pool of videos,



all you need is for some of them to have decent reach and get recommended, right? Because then that's like igniting a pile of kindling where then if others are good, if they're going to give people good experiences, they get not only nominated but then recommended which then kicks back in the feedback loop there.

That turns out to be as close to a guarantee as you can get of saying if you make something that's good, it's a good piece that will satisfy someone, they come away feeling like they learned something that they otherwise didn't know and it was well presented, if you can get it into this peer review process, it will reach people. It's not just going to be shouting into the void

And in this case, last year there were over a hundred videos where after the first two weeks they had more than 10,000 views. Which I know is small in the grand scheme but for a fresh channel, talking about a niche mathematical topic, to be able to put it out and get 10,000 people to watch it is amazing. And the idea that that it happen for over a hundred people is amazing

That had nothing to do with the prize pool, right? In that the motive might have been a hope of actually getting some reach and having some sense of a guarantee of there being some reach

Ironically the reason to do the whole peer review system in the first place is in the service of selecting winners. If you just said "Hey, we're having a watch fest where everyone watches each other's things." Somehow it wouldn't quite have the same pull that gets people into it. So I think it still makes sense to have winners and to have some material behind those winners. It doesn't have to be much though. And if anything, I think it might ruin it to make it too much. I will also say it's \$15,000 actually because we give \$500 to 20 different honorable mentions, at least this year. Still pretty modest in the scheme of how much money you can invest to try to get more math lessons in the world.

### **Dwarkesh Patel**

I watched many of the honorable mentions as well because they were just topics that were interesting to me. It's like the thing that the president of Chicago University said. He said we could discard the people we admitted and select the next thousand for our class and there would be no difference.

By the way, I really admire not only the education that you have provided directly with your videos which have reached millions of people, but the fact that you're also setting up this way of getting more people to contribute and get to topics that you wouldn't have time to get to yourself. I really admire that you're doing that.

If you're self teaching yourself a field that involves mathematics, let's say it's Physics or some other thing like that, there's problems where you have to understand how do I put this in terms of a derivative or an integral and from there, can I solve this integral? What would you recommend to somebody who is teaching themselves quantum mechanics and they figured out how to put how to get the right mathematical equation here. Is it important for their understanding to be able to go from there to getting it to the end result or can they just say well, I can just abstract that out. I understand the broader way to set up the problem in terms of the physics itself.

### **Grant Sanderson**

I think where a lot of self learners shoot themselves in the foot is by skipping calculations by thinking that that's incidental to the core understanding. But actually, I do think you build a lot of intuition just by putting in the reps of certain calculations. Some of them maybe turn out not to be all that important and in that case, so be it, but sometimes that's what maybe shapes your sense of where the substance of a result really came from.

I don't know it might be something you realize like "Oh, it's because of the square root that you get this decay." And if you didn't really go through the exercise, you would just come away thinking like instead of coming away thinking like such and such decays but with other circumstances, it doesn't decay and not really understanding what was the core part of this high level result that is the thing you actually want to come out remembering.

Putting in the work with the calculations is where you solidify all of those underlying intuitions. And without the forcing function of homework, People just don't do it. So I think that's one thing that I learned as a big difference post college versus during college.

Post-college, it's very easy to just accidentally skip that while learning stuff and then it doesn't sink in as well. So I think when you're reading something, having a notebook and pencil next to you should be considered part of the actual reading process.

And if you are relying too much on reading and looking up and thinking in your head, maybe that's going to get you something but it's not going to be as highly leveraged as it could be

### **Dwarkesh Patel**

What would be the impact of more self teaching in terms of what kinds of personalities benefit most? There's obviously a difference in the kind of person who benefits most. In a situation where it's a college course and everybody has to do the homework, but maybe some people are better tuned for the kind of work that's placed there versus all this stuff is available for you on youtube and then textbooks for exercises and so on but you have to have the conscientiousness to actually go ahead and pursue it.

How do you see the distribution of who will benefit from the more modern way in which you can get whatever you want but you have to push yourself to get it.

### **Grant Sanderson**

There's a really good book that's actually kind of relevant to some of your early questions called *Failure to Disrupt* that goes over the history of educational technology. It tries to answer the question of why you have these repeated cycles of people saying such and such technology that almost always is getting more explanations to more people, promises that it'll disrupt the existing university system or disrupt the existing school system and just kind of never does.

One of the things that it highlights is how stratifying these technologies will be in that they actually are very very good for those who are already motivated or kind of already on the top in some way and they end up struggling the most just for those who are performing more poorly.

And maybe it's because of confounding causation where the same thing that causes someone to not do poorly in the traditional system also means that they're not going to engage as well with the plethora of tools available.

I don't know if this answers your question, but I would reemphasize that what's probably most important to getting people to actually learn something is not the explanation or the quality of explanations available because since the printing press that has not been true. Not literally true because maybe access to libraries it's not as universal as you would want. But people had access to the explanation once they were motivated.

But instead, it's going to be the social factors. Are the five best friends you have also interested in this stuff and do they tend to push you up or they tend to pull you down when it comes to learning more things? Or do you have a reason to? There's a job that you want to get or a domain that you want to enter where you just have to understand something or is there a personal project that you're doing?

The existence of compelling personal projects and encouraging friend groups probably does way way more than the average quality of explanation online ever could because once you get someone motivated, they're just they're going to learn it and it maybe makes it a more fluid process if there's good explanations versus bad ones and it keeps you from having some people drop out of that process, which is important.

But if you're not motivating them into it in the first place, it doesn't matter if you have the most world-class explanations on every possible topic out there. It's screaming into a void effectively.

And I don't know the best way to get more people into things. I have had a thought and this is the kind of thing that could never be done in practice but instead it's something you would like write some kind of novel about, where if you want the perfect school, something where you can insert some students and then you want them to get the best education that you can, what you need to do is – Let's say it's a high school. You insert a lot of really attractive high schooler plants as actors that you get the students to develop crushes on. And then anything that you want to learn, the plant has to express a certain interest in it. They're like, "Oh, they're really interested in Charles Dickens." And they express this interest and then they suggest that they would become more interested in whoever your target student is if they also read the Dickens with them.

If you socially engineer the setting in that way, the effectiveness that would have to get students to actually learn stuff is probably so many miles above anything else that we could do. Nothing like that in practice could ever actually literally work but at least viewing that as this end point of "Okay, this mode of interaction would be hyper effective at education. Is there anything that kind of gets at that?"

And the kind of things that get at that would be – being cognizant of your child's peer group or something which is something that parents very naturally do or okay, it doesn't have to be a romantic crush, but it could be that there's respect for the teacher. It's someone that they genuinely respect and look up to such that when they say there's an edification to come from reading Dickens, that actually lands in a way.

Taking that as a paragon and then letting everything else approximate that has, I would emphasize, nothing to do with the quality of online explanations that there are out there that at best just makes it such that you know, you can lubricate the process once someone is sufficiently interested.

**Dwarkesh Patel**

You found a new replica use case.

**Grant Sanderson**

Yes. I mean, I'm not saying we should do it, but think of how effective that would be.

**Dwarkesh Patel**

Final question. This is something I should have followed up on earlier, but your plans to become a high school teacher for some amount of years. When are you planning on doing that and what do you hope to get out of that?

**Grant Sanderson**

I would say no concrete plans. I would want to do it in a period where I also have young children and therefore it would make sense to. Maybe a lot of people will say this kind of

thing but there's friends of mine who think when their child is in high school, that's when they would want to be a high school teacher.

I think there are two things I would want to get out of it. One of them, as I was emphasizing, I think you just lose touch with what it's like not to know stuff or what it's like to be a student and so maintaining that kind of connection so that I don't become duller and duller over time feels important.

The other, I would like to live in a world where more people who are savvy with STEM spend some of their time teaching. I just think that's one of the highest leverage ways that you can think of to actually get more people to engage with math

And so I would like to encourage people to do that and call for action. Some notion of spending, maybe not your whole career, a little bit of time. In teaching, there's not as fluid a system for doing that as a going through a tour of service in certain countries where everyone spends two years in the military

Shy of having a system like that for education, there's all these kind of ad hoc things where charter schools might have an emergency credential system to get a science teacher in. Teach for America is something out there.

There's enough ways that someone could spend a little bit of time that's probably not fully saturated at this point that the world would be better if more people did that and it would be hypocritical for me to suggest that and then not to actually put my feet where my words are.

### **Dwarkesh Patel**

I think that's a great note to leave it on Grant. Thanks so much for coming on the podcast and genuinely, you're one of the people I really, really admire but what you've done for the landscape of Math education is really remarkable. So this is a pleasure to talk to you.

### **Grant Sanderson**

Thanks for saying that. I had a lot of fun.