

▶ SAM HARRIS / ANGELA DUCKWORTH / RAPHAEL MILLIERE /
JOHN O'DONNELL / YASHENG HUANG / ALAN RUSBRIDGER /
HUW VAN STEENIS / NATHAN MYHRVOLD ■

THE PODCAST

WISDOM AND INSIGHTS FROM OUTSTANDING LONGFORM PODCASTS **READER**



ISSUE 11

WWW.PODREAD.ORG

AUD	GBP	USD	EUR
\$14.99	£8.99	\$10.99	€11.99



ISSN 2653-1213

TPR IS A PROUD SUPPORTER OF



On Meditation, Mindfulness and Morality

On the Importance of Grit

How Artificial Intelligence Thinks

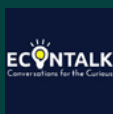
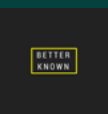
Why Electrifying Industrial Heat is Such a Big Deal

How China's Economic and Political History Can Help Us Understand its Future

A Conversation for the Curious

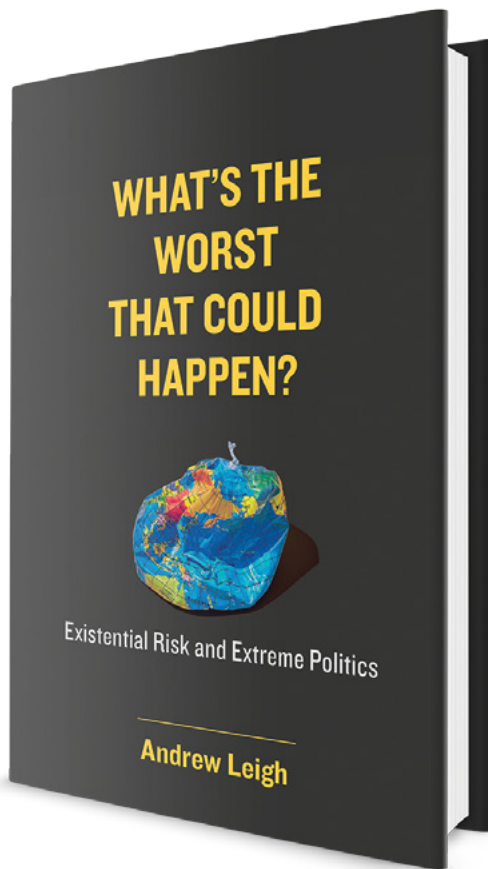
On the History and Cause of Banking Crisis

On Technology, Climate Change and Geoengineering





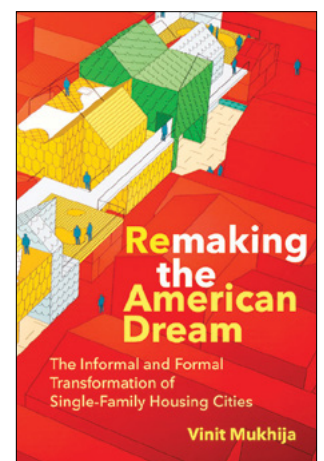
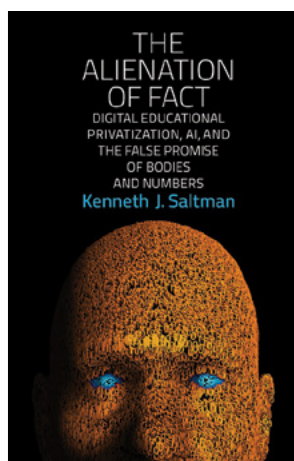
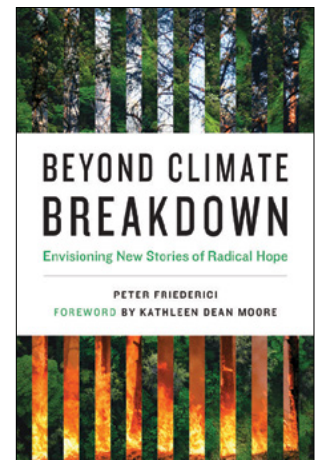
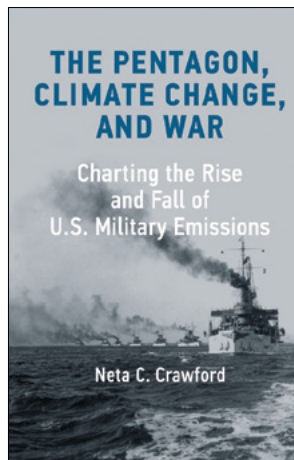
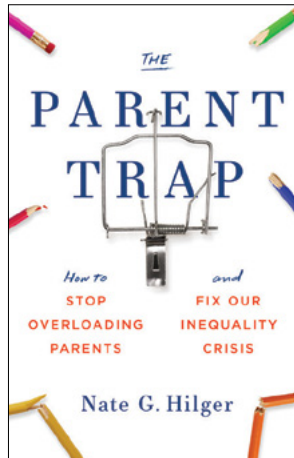
Big ideas at the intersection of political science, tech, and society from the MIT Press



“A clarion call for improving political institutions and strengthening democratic norms. We may all have developed tunnel vision during the pandemic, but this thoughtful book will help us see at last some light at the end of it.”

—Angela Kane, former UN High Representative for Disarmament Affairs

mitpress.mit.edu



THE **PODCAST** READER

Welcome to Issue Eleven of *The Podcast Reader*, a more permanent platform for outstanding longform podcasts. Our curated podcast transcripts make it easier to follow important ideas and highlight key points. In a world of digital distraction and ever-shorter attention spans, we are proud to provide a more reflective platform for important ideas.

In this issue we present full transcripts from six longform podcast interviews, and edited highlights, or 'Podcast Bites', from a further two episodes.
We cover three broad categories of content:

How to improve society:

John O'Donnell on his company's novel approach to energy storage

Nathan Myhrvold on technology and climate change

Huw van Steenis on money and banking crisis

Frontiers of knowledge:

Yasheng Huang on assessing China's future through understanding its economic history

Raphaël Millière on how artificial intelligence models learn

How to improve yourself:

Angela Duckworth on grit and perseverance

Sam Harris on mindfulness and meditation

Alan Rusbridger on unusual items he considers important, and his career in the media

Each issue of *The Podcast Reader* aims to present content from the arts, entrepreneurship, history, public policy and science. In short, a cross-section of ideas that shape our world.

Reader feedback is essential to help us learn and improve, so please don't hesitate to share your thoughts about the magazine at hello@podread.org.

The Podcast Reader acknowledges the Kulin Nation as Traditional Owners of the land on which it is situated in Melbourne and Geelong, and pays respect to their Elders, past, present and emerging.

4. FEATURED GUESTS

6.

Sam Harris

**ON MEDITATION, MINDFULNESS AND
MORALITY**
ECONTALK

18.

Angela Duckworth

**ON THE IMPORTANCE OF GRIT
IN GOOD COMPANY**

28.

Raphaël Millière

HOW ARTIFICIAL INTELLIGENCE THINKS
MINDSCAPE

42.

John O'Donnell

**WHY ELECTRIFYING INDUSTRIAL HEAT IS
SUCH A BIG DEAL**
VOLTS

54.

Yasheng Huang

**HOW CHINA'S ECONOMIC AND
POLITICAL HISTORY CAN HELP
US UNDERSTAND ITS FUTURE**
CONVERSATIONS WITH TYLER

66.

Alan Rusbridger

**A CONVERSATION FOR THE
CURIOUS**
BETTER KNOWN

PODCAST BITES

75. HUW VAN STEENIS ON THE HISTORY AND
CAUSES OF BANKING CRISIS

79. NATHAN MYHRVOLD ON TECHNOLOGY,
CLIMATE CHANGE AND GEOENGINEERING

**88. HOW TO PURCHASE PAST
AND FUTURE ISSUES**

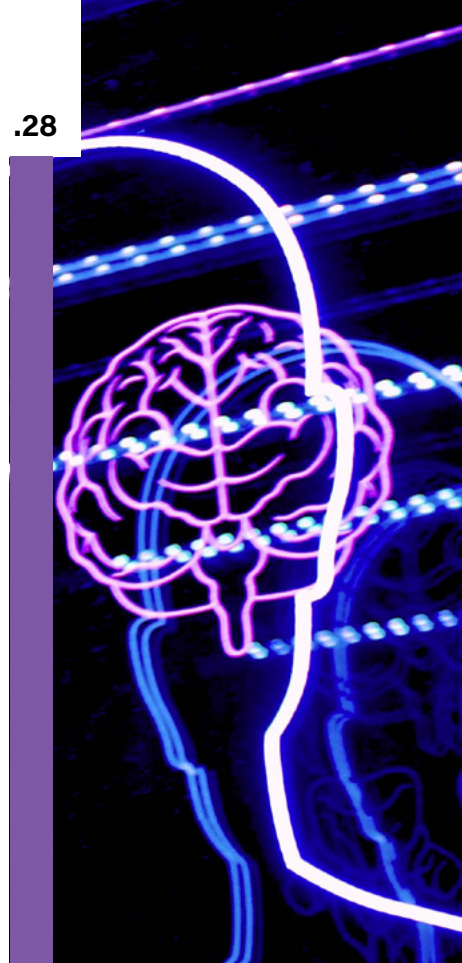




.6



.18



.28



.54

“

... grit ends up being a better predictor than, for example, objective measures of physical talent or objective measures of intellectual talent.

”

ANGELA DUCKWORTH

.66



FEATURED GUESTS

Sam Harris holds a degree in philosophy from Stanford University and a PhD in neuroscience from UCLA. He is host of the Making Sense podcast and creator of Waking Up, a meditation app. His latest book is *Waking Up*.

Angela Duckworth holds a PhD in psychology from the University of Pennsylvania, where she is currently a Professor of Psychology. She is a board member of Character Lab, and was named a MacArthur Fellow in 2013. She is the author of *Grit: The Power of Passion and Perseverance*.

Raphaël Millière holds a PhD in philosophy from Oxford University. He is currently a Presidential Scholar in Society and Neuroscience at the Center for Science and Society, and a Lecturer in the Philosophy Department at Columbia University.

John O'Donnell is CEO of Rondo Energy, which aims to produce zero-emission industrial heat and power. He holds a BSc in computer science from Yale University, and holds more than twenty patents in the US and internationally.

Yasheng Huang holds a BA from Harvard College and a PhD from Harvard University. He is currently a Professor of Global Economics and Management at MIT Sloan School of Management, and is the author of eleven books, most recently *The Rise and Fall of the East*.

Alan Rusbridger is editor at *Prospect Magazine*, and former editor of *The Guardian*. He is a member of the Facebook Oversight Board and Chair of the Reuters Institute for the Study of Journalism. He is the author of three books, most recently *News and How to Use It*.

Huw Van Steenis holds an MA in politics, philosophy and economics from Oxford University, and an MBA from INSEAD. He is currently the Vice Chair at Oliver Wyman, and is a member of the Climate Advisory Board of NBIM, the Norwegian sovereign wealth fund.

Nathan Myhrvold holds a MA in mathematical economics and a PhD in theoretical and mathematical physics from Princeton University. He was chief strategist and chief technology officer at Microsoft, and is currently CEO of Intellectual Ventures.

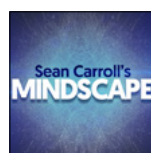
FEATURED PODCASTS



ECONTALK
SAM HARRIS
P. 6



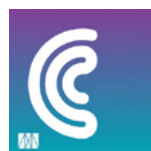
IN GOOD COMPANY
ANGELA DUCKWORTH
P. 18



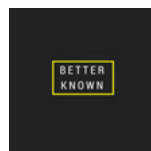
MINDSCAPE
RAPHAEL MILLIERE
P. 28



VOLTS
JOHN O'DONNELL
P. 42



CONVERSATIONS WITH TYLER
YASHENG HUANG
P. 54



BETTER KNOWN
ALAN RUSBRIDGER
P. 66



MONEY MAZE
HUW VAN STEENIS
P. 75



THE KNOWLEDGE PROJECT
NATHAN MYHRVOLD
P. 79



**TO PURCHASE BACK EDITIONS,
VISIT WWW.PODREAD.ORG**

**Dialogues ■
Media ■
Foundation.**

THE PODCAST READER
PUBLISHED BY DIALOGUES MEDIA FOUNDATION
PO BOX 541 CARLTON NORTH VIC 3054

ADVISORY BOARD
NICHOLAS GRUEN
JOE WALKER
PETE MORAN
YANN BURDEN

THE PODCAST READER TEAM
PUBLISHER DAVID LOGGIA
EDITOR ANDREA O'CONNOR
CREATIVE DIRECTOR VAUGHAN MOSSOP
WEB AND SOCIAL MEDIA LAURA SULLIVAN
CREATIVE CONSULTANT MICHAEL SKARBEEK

CONTACT US
HELLO@PODREAD.ORG
PODREAD.ORG
[@PODREADMAG](https://twitter.com/PODREADMAG)
[@PODREADMAG](https://www.instagram.com/PODREADMAG)
[/PODREADMAG](https://www.facebook.com/PODREADMAG)

PUBLISHED IN 2023

Printed by Contract Publishing UK (CPUK) Limited

The Podcast Reader is a proud supporter of:



The Podcast Reader is published four times per year. Views expressed by authors do not necessarily reflect those of the publishers. Copyright is reserved, which means you can't scan our pages and put them up on your website or anywhere else. Reproduction in whole or part is prohibited.
ISSN 2653-1623 PPN 100029061



QUINOLA'S GENEROUS REVOLUTION

**WE'RE IN LOVE WITH
US HUMANS.
WE MIGHT BE MILES
FROM PERFECTION.
BUT WHEN WE ALL
DECIDE TO ACT,
THERE'S NO STOPPING
US. WE BELIEVE THAT
BEING GENEROUS TO
THE ENVIRONMENT,
TO OURSELVES AND
EACH OTHER, IS THE
ONLY WAY TO ACHIEVE
BETTER.**

**IT'S TIME TO PROVE
OUR COLLECTIVE
POTENTIAL. IT'S TIME
TO START RESTORING
THE ENVIRONMENT
AND REDISTRIBUTE
OUR RESOURCES. IT'S
TIME TO GET OUT
THERE AND ACTUALLY
START DOING
SOMETHING.
IT'S TIME FOR
MEANINGFUL CHANGE.**

QUINOLA™

**MAKING EATING
RIGHT, EASY**

WWW.QUINOLA.COM

[@QUINOLAUK](https://www.instagram.com/QUINOLAUK)

On Meditation, Mindfulness and Morality

Harnessing the power of thought

SAM HARRIS
ECONOTALK, 2023

Interview by Russ Roberts

Russ Roberts: My guest is neuroscientist, philosopher and author, Sam Harris. He hosts the podcast *Making Sense* and is the creator of the meditation app Waking Up. Sam, welcome to *EconTalk*.

Sam Harris: Great to see you, Russ.

RR: Give us a thumbnail of how you came to be where you are, with an incredibly popular podcast and an incredibly popular meditation app. How'd that happen?

SH: Well, I started as a writer. I started in an unconventional spot because I wrote my first book in the middle of what should have been my PhD. I just finished my coursework and I was beginning my research. Then, September 11th happened, and I wrote my first book, *The End of Faith*. That proved

so controversial, and the conversation around those issues was so rich and interesting that I quickly wrote a second book in response to the pushback, *Letter to a Christian Nation*. That essentially side-lined me for about four years during my PhD. I had a toe in the lab, but barely a toe, for four years. So, I took nine years to finish my PhD, and that's really what writing was doing to me during that time. But it was really on the basis of my writing platform that I launched my podcast and then, subsequently, the Waking Up app.

So, I was fairly early to podcasting. I had been a guest on a few podcasts, including Joe Rogan's and, I think, Tim Ferris's; maybe one or two others. And I just thought, 'Well, this is interesting.' If you had told me that I might one day go into radio, I would've said you were insane. But something about the medium made it seem different. So, I just started recording pieces of audio. Initially they were solo audio riffs or essays, and I was releasing those sporadically without really even thinking that I had a podcast. Then, at a certain point, I was interviewing other people more or less once a week, and I had a podcast in earnest. That's how it started.

RR: What have you learned from being a podcaster? Have you changed in any way?



SH: Well, personally, I've learned about the power of incentives. Because as much as I've wanted to get back to writing books, having a podcast has shown me that – this won't surprise you as an economist – I am a creature of incentives, like virtually everyone, and all of the incentives are aligned away from writing books at the moment. Podcasting is easier, I reach many more people, and it's a better business.

So, for me to go back to writing and embrace the opportunity cost of writing at the moment, I really have to think, 'Well, I don't care about doing the harder thing. I'm happy to do the harder thing. I don't care about reaching fewer people. I don't care about it taking much longer to reach those fewer people. And, I don't care about losing money.' All the incentives are wrong for writing my next book. So, I haven't done that.

I think I will ultimately do it, because I think writing is just a muscle. As a thinker, you need to work and you really don't think as clearly as you can unless you're writing your thoughts and finally producing the sentence that you think is the best version of any specific thought. So that is a loss to me, but it's been great. I mean, podcasting is, as you know, so different from writing because you're not doing it alone. You and I are having a conversation now and we have an excuse to have this conversation. And, the truth is, it's a conversation I'd want to have anyway for free, right? So it's really an amazing opportunity to use media to help the people who want to hear these conversations and to have fun ourselves. I feel immensely lucky.

RR: But, has it taught you anything? I mean, you could have read the books of all your guests. Many of them write books. Do you find that talking to the rather diverse range of people that you speak to affects you in any way? Has it affected your thinking?

SH: Yeah, certainly. Because, you know, as a writer, I'm not someone who interviews people for the most part by way of research. I obviously read a lot of books to be a writer of non-fiction, but there is something about talking to smart people and having them push back against your views in real time that you can't really supply for yourself in the same way.

When you write a book, it takes you a year or more to write it. It then sits with your publisher for eleven months or so, and then it goes out into the world, then you get some feedback if people review it or people react to it. But, the time-course of correction and fertilisation of further conversation is so slow.

RR: I never thought about that. I often will get on a topic and interview a series of people in clumps. You know, I'll read somebody's book, and then three weeks later or a month later, I'll interview a person on the other side, or a related theme. Like you, I'm very

interested in consciousness, so I've done a bunch of interviews on that. I've never thought about the fact that you read a book about consciousness by an author and then maybe you read another one down the road that has a different take, different perspective. But in podcasting, you're almost inevitably doing it over a relatively short period of time. And then you're in dialogue rather than in your own head, the way you would be as a reader with diverse ideas or different takes or perspectives. I guess it quickens the pace.

One of the things I find extraordinary about podcasting for a long time, as you have, is how many connections I see between topics and episodes that don't necessarily seem related. I've learned so much from being an interviewer, not just from the content I've consumed to prepare for them, but to have that conversation like we're having now, and to have it – it's 8.00 at night here in Jerusalem, and it's 10.00 in the morning in California where you are, and – well, that's a miracle. So, it's not just nice to have the conversation: if we weren't podcasting, we probably wouldn't be talking. And so, it is very special.

SH: Well, that's what I've appreciated about it most, really. Writing is such a solitary endeavour. And podcasting, especially if you're mostly doing interviews, is a completely different experience, because you now have a venue to invite people to. And, you're helping them. You're helping them launch their books in many cases. But it is just like this guilty pleasure, to be able to talk to the smartest people in the world about anything. And, when you have a successful podcast, you're not really asking a favour of them: you're doing them a favour, if anything. And so, it's wonderful to be able to.

And, it's just good company, right? You just get to meet people you wouldn't otherwise have an excuse to meet. I wouldn't reach out to even a favoured author just to reach out to them, but because I have a podcast and because their publicist may have even hurled their next book at me, it's just that we're naturally thrown together in conversation. And, yeah, it builds relationships. It's quite amazing.

RR: How much time do you spend reading? Not for podcasting, just in general.

SH: Well, that's a hard line to draw because I have, to a significant degree, designed my podcast around what I feel like reading next. So, I just decide what I want to read and then the afterthought is, 'Oh, wait a minute, if this person's alive, I might be able to talk to them.' Again, this comes back to being immensely lucky and feeling just pure gratitude for the existence of this medium.

RR: I'm going to suggest that you are an example of

“

One of the things I find extraordinary about podcasting for a long time, as you have, is how many connections I see between topics and episodes that don't necessarily seem related.

”

a phenomenon that I think is a new phenomenon. If I had to describe you, I'd say you're a public intellectual. That's one phrase that people would use. It's a phrase I don't like, personally. I'm not sure why I don't like it. I've never liked it when it's been applied to me. But, you're something more than that. I would describe you as a secular preacher. An atheist rabbi. You're in a small group with Jordan Peterson. I don't know who else you'd put in the group. But, people don't just listen to you because you're smart. They don't just listen to you because you're interesting and entertaining. They look to you for meaning and guidance.

Am I right? And, what does that feel like? How did it come about? And, what are the upsides and downsides of that?

SH: Well, I think it's a matter of the kinds of topics I've focused on. It's a matter of what I have found interesting and what I have made my areas of relative expertise, just because I've spent so much time focusing on these questions. I'm interested in the nature of human subjectivity; how our scientific understanding of ourselves is increasingly encroaching on ancient ways of deluding ourselves about ourselves. So, just how can we live the best lives possible in light of the fact that we all are going to lose everything we love in this world? These are the deepest questions: What are you as a being in this world, and what should you do given the full range of wonderful and terrible experiences on offer? And, given the fact that impermanence reigns. You can't hold onto anything in the end.

How is it possible to be fulfilled and to live a good life within those constraints? What can we know

about how to do that wisely? How can we know when we're making obvious and needless errors? How do we mitigate human suffering? And, what does science have to say about all that?

How do we have a conversation about what we believe to be true about all of that, that is increasingly useful and open-ended and tracking of reality rather than spiralling into some awful and divisive state of delusion and fractiousness and failure to cooperate and failure to solve absolutely necessary coordination problems? So that's where I've focused. And, the nature of that focus is, really, by definition interdisciplinary. I've never respected the boundaries or apparent boundaries between fields of knowledge.

I'm a big fan of the notion of consilience or the unity of knowledge. I think much of our partitioning of domains, you know, the wall between neuroscience and philosophy of mind, say, or economics, are enshrined by two things. One is it's hard to be a polymath. There's just too much to know. And that's a natural partition perhaps but there are some very smart people who can traverse all of those boundaries, and that's wonderful. But the other reason is just the architecture of the university and the nature of bureaucracy, and just the laziness born of the norms that grow up around that.

I do consider myself more of a generalist than a specialist at this point, I think, of necessity. And so, I don't discount the need for specialisation. Some people specialise so fully that their career is going to observe the obvious boundaries between, let's say, molecular biology and everything else. But I do think there's a role for – and I don't shy away from the phrase – the 'public intellectual'. I think it sounds pretentious when applied to oneself. But, like you, I don't have a better name for the work that certain people do trying to integrate various fields of knowledge and make those integrations useful to the general public. I just think that's a good role.

And, importantly, it's not the role of a journalist. Or it's not the role of a mere journalist. I mean, in journalism – not to denigrate journalism, I think we absolutely need more of it – but, I don't consider myself a journalist, even when I'm trying to present a factually accurate picture of something that's happened or what's been said. I do have my own point of view on a wide variety of topics, and I try to, if for no other reason than avoid embarrassment, have a well-informed point of view on those topics. That's not quite the same thing as journalism.

RR: I noticed in that summary that you said, 'What we can learn from science about a life well lived.' You didn't say it quite like that, but about the human experience. You did not mention art or fiction or other things. Was that deliberate?

SH: Yeah, that was perhaps just the burden of long-winded grammar. Ironically, I think, if anything, recent developments in culture and the overwhelming influence of technologists at the moment, suggests to me that the relevance of the humanities to our intellectual lives has never been more pressing. I think we're suffering from the outsized influence of many smart people who don't have much of an education in the humanities, who haven't read enough good books. You know, you've got a lot of people in Silicon Valley who have read a lot of science fiction and too much Ayn Rand, and they have this enormous influence on culture, if for no other reason than that they've built the tools that are dictating so much of what is said and done at this point. So yeah, I think there's much more to a picture of the human circumstance and how we respond to our existential concerns than a narrowly scientific one.

But, I would say that when getting our heads straight about a topic matters, rationality is the essential tool for us to use. So, rationality obviously suggests a larger footprint than lab-coated science, but it is distinct from mere aesthetics and the mere creation of beautiful fictions, and it's certainly distinct from wishful thinking and tribalism. I talk in one place – I think it might be my first book – about us needing to navigate by love and curiosity. And, I believe I say this: I think of reason as the guardian of love. It's often alleged that there's something cold about rationality or that it's the opposite of many things we care about. But, more and more, I view it as the guardian of everything we care about.

I just think that the moment you give too much scope for the irrational and the specious, you just start to bump into hard objects in the dark. And, I mean, reality has a structure, and insofar as our conversation with one another can be truth-tracking and consistent, that allows us to avoid the most unpleasant collisions, both with one another and just the way the world is, however it is.

RR: You are one of the most famous atheists in the world. And, I live a committed Jewish life, meaning I pay attention to Jewish law. And, I make more room in my life for what I would call the mystical. I don't call the mystical the irrational. I agree with you on that. I think irrational and emotional can be a great danger.

But I also think there's a great danger to reason in that it's hard for us to remember that – I like Richard Feynman's version: 'The most important thing is not to fool yourself, and you're the easiest person to fool.' And so, I'm very wary – it's my Hayekian side – and to the extent I understand the philosophy of Charles Peirce, a recognition that reason is an incredibly powerful tool that, if worshipped, can lead to great danger also. Whether you believe in God or not, a serious religious person has – should have, I believe

– a humility about them in the face of the transcendent, the universe, whatever you want to call it. And again, even if you don't believe in God, in the face of the tiny-ness of the human mind and our human existence. And I think the risk of the other side, the over-trust in reason is to put humanity on a pedestal. And that hasn't turned out very well historically. So, that would be my pushback. You want to react to that?

SH: Yeah. Well, I agree with that. I'm not quite standing where you're pushing back, because I do view reason, rightly conceived, and I view mysticism/spirituality rightly conceived as two modes in which we transcend the self. So, self-transcendence I think is the centre of the bullseye for so many things we care about, whether it's acknowledged generally or not. But, when you look at what it means to live an ethically good life, when you look at what it means to live a rationally coherent life or to create a rationally coherent worldview, and when you look at what it means to occupy the furthest reaches of psychological wellbeing – and, I would argue normativity – I think transcending the ego, transcending the selfish, self-directed, narcissistic illusion of 'it's me in' – the primacy of 'me in here,' the subject who was thinking and seeming to author his own thoughts – that transcendence is the whole point of being alive in the end.

When it happens to you in a haphazard way or you are lured into experiencing it through some agency outside yourself – let's say you take a psychedelic drug, or you have some profound experience of falling in love, or some collision with natural beauty, or you're a scientist who discovers something and that discovery overwhelms you with awe – these are glimpses of a bigger view of mental life that is, I would argue, intrinsically pro-social, and it's the basis for compassion, real compassion rather than pity or some simulacrum. It really is just – this is the good stuff in life: falling in love, caring about others as much or even more than you care about yourself, or even just seeing your own selfishness become inseparable from a seemingly selfless desire to better the lives of other people – what you most selfishly want is for other people to succeed. Right? That's making you happy. So, this opposition – the apparent opposition between selfishness and selflessness – breaks down under these conditions.

Rationality is not the same thing as mysticism. But I wouldn't say that mysticism or spirituality, rightly conceived, is irrational. What I'm calling irrational is to believe things that are obviously wrong or for which you have no good reasons for belief. And, to be believing two things strongly in various moments that cannot be reconciled with one another, so there would be an obvious contradiction with yourself, and to not notice that; to be dogmatic and to be

unavailable to better arguments and better evidence in future conversations.

So, something like dogmatism is worth focusing on. Because, I view dogmatism – and this is why traditional religion has come under such opprobrium in my discussion of it – as intrinsically divisive, because it is the very posture of being inflexible and unpersuadable. And, just when you look at what it is to be dogmatic – and, in religion, traditionally it's only in religion that being dogmatic is not considered a bad thing. That's not to say something as a dogma is not to even say anything invidious in religious terms.

RR: A virtue.

SH: Yeah. So, the dogmas you hold – religiously and in any other mode of life – are beliefs that you have decided in advance, by whatever process, and usually none other than receiving them on your mother's knee, these are beliefs that cannot be revised, that you are unwilling to revise. And that, far too often, you will be offended if anyone asks you to reconsider them.

So, you're entering every conversation saying, 'Yeah, I'm open-minded, I'm rational. The year is 2023. Happy to talk about anything. But I have a list of beliefs over here that if you challenge them, I'm going to begin to hate you. And, I'm certainly not going to rethink any of them no matter how good your arguments are.' I consider that to be the fundamental, and obviously fundamentalist, religious sensibility. And, that I view as intrinsically divisive: it is intrinsically shattering of our epistemology. And, I do think we have to overcome that frame of mind where it exists, wherever it exists. And it exists especially in religion, certainly in politics, and elsewhere, and I think it's the enemy of reason, yeah.

RR: I'll defend it a little bit, and then I want to reflect on it: it was a very thoughtful outline of the challenge of dogmatism. I wrote an essay on... I don't know if you've ever seen the show *Come From Away*, the musical?

SH: No.

RR: It's a ridiculously beautiful show. It's about how after 9/11 U.S. airspace was closed and a bunch of flights had to land in the middle of nowhere in Canada. The show is about how the tiny group of people who lived there rose to the occasion. Part of the reason they rose to the occasion is that they had a tribal urge – they had a certain image of themselves – as, I'd say, resilient. Among other things, of course. But that resilience carried through and to say it saved the day is an understatement for the people who landed there.

And, similarly, in a religious community that's effective. It's divisive – dogma is divisive – for the people outside the community. It's incredibly unifying and exhilarating for the people inside the community. If you've not been part of that, it's hard to imagine. There are very few movies or treatments of it that have captured it. One of the things I've observed about *Shtisel*, which is about a religious community, is that they don't romanticise anything in that show about religious life. These are people just like you and me. And they've got the same problems we have with their children, and their worries, and their anxieties, and their fears and their dreams. And, they don't make fun of the religious aspect of it and they don't glorify it. But there is a glorious part to it when it works well. And there's a lot of negatives to it: I'm not going to defend fundamentalism. I agree with you, it's divisive. It does create an us/them mindset. But it does change the us in a very powerful way. And, it potentially doesn't necessarily harm them, but it depends. There are many, of course, historical cases where it's otherwise.

So, I take your point. But I'm thinking about my own life: I'm really into dogma. It's interesting, right? For most of my life, I was a very hardcore libertarian, which is a very dogmatic perspective on economics. I became a religious Jew as an adult, in my mid- to early-thirties. And I had two different dogmas. Some people would tease me and say, 'How do you maintain both of those? They conflict.' I don't think they conflict. But I think what's more interesting is: I don't really see them as dogmas anymore, in the sense that they are frameworks I use, but I don't feel the same way I did about my economics views as I did ten years ago. I'm still pretty free market, but I'm not knee-jerk free market. And that would be true of my religious view as well. I have a Jewish practice, but I have lots of doubts. Every serious Jew that I know – and I can't speak for Christianity or other religions, Islam – they have doubts. A lot of people I know have doubts. The framework is a way of living. And, for me – and this is for another conversation, Sam, we'll have it maybe after I write another book – for me, it's a way of experiencing the transcendent and the wondrous and the awesome. And the us-and-them part is the least of it. I think for many people, their dogmatism is a shield. And, if you can learn to realise that it's just a shield – it's not truth, it's just the thing you carry through life to organise your thinking about certain things – I don't think it's as bleak as you paint it. But maybe there are others for which that's the case.

SH: Right. Well, let me just say that I think reasoning by analogy from Judaism is generally pretty misleading. I mean, I'm Jewish: I've obviously gone around this track many different times with people of, really, every faith. And, there's something about Judaism – and

we can talk about the historical and theological reasons why this is so – that makes it an outlier with respect to the claims, those kinds of pretensions of propositional knowledge it makes that go under the banner of faith. So, faith as a Jew is quite a bit more leavened, ironically, with doubt and pragmatism, and a circumspection when it comes time to making bold claims about what happens after death and the moral structure of the universe, and the behavioural imperatives of living within that structure.

I would just caution you not to extrapolate from your experience as a Jew of whatever flavour to the experiences of 1.8 billion Muslims and 2.2 billion Christians, because in my experience they're at least rhetorically committed to far more strident assertions of knowledge about what happens in the invisible world of after-death.

RR: Well, I think that's true of the Orthodox world and Judaism as well. And, it's certainly true of the ultra-Orthodox, but even in the so-called plain Orthodox that would be true. But our audience aren't so interested in these distinctions. I'll just say this: You wrote a book called *Free Will*. I read it before this conversation because I'm very interested in the question, and it comes up now and then in all kinds of settings on this program. But a thoughtful person has to wonder about it, I think.

And, Maimonides most people would say he is one of the two or three greatest thinkers of Jewish history – he believed in free will. But Crescas, a rabbi of his time who I love, did not believe it at all. It turns out Crescas is the Sam Harris of the Middle Ages. I will concede the point that there's some variety of practices in Judaism that may not be reflected in other religions. I'm not going to speak about them. You've had a lot more interaction with them than I have, certainly rhetorically for sure–

SH: Well, certainly you can say – this is how extreme it is – you can say you're Jewish, and your Judaism is incredibly important to you, but you don't believe in God. That is not a total non-sequitur in Judaism. It is a total non-sequitur in Christianity and Islam.

RR: Let's talk about meditation while we're on the topic of spirituality. Talk about the nature of your personal practice. How did you come to it and what does it mean to you?

SH: I was eighteen when I sat at my first meditation retreat. I'd come to it first through taking MDMA maybe six months before I sat at my first retreat. I took MDMA; I had experimented with marijuana and maybe I had taken mushrooms a couple of times as a teenager, but I'd never done that in the context of thinking consciously about understanding anything

about the nature of my mind. It was just for fun – party drugs that teenagers were using at the time.

With MDMA, I was given it with the explicit framing: You might really learn something about yourself if you take this. So, it was really with that intention, and that was setting of it. I was just sitting with my best friend. And we took this drug. MDMA, at that point – otherwise known as Ecstasy, now Molly – was being widely used in a therapeutic community as a tool of insight; and it leaked out of that community and got into my hands at that point. I had an experience on the drug which will be familiar to many people who have taken it, but maybe not if you took it at a rave or a party or a rock concert. Again, I think one's intentions matter here. But I had this experience of – for lack of a better phrase – unconditional love. It was an experience of not feeling high or altered or stoned, and it was certainly not a sense of my perception of the world being distorted. It was actually an experience of being sane for the first time in my life. And I just felt, 'Okay, this is more real than what I've been tending to experience, this feeling I'm feeling right now.' This is not an artefact of pharmacology: This was a stripping away of features of my own mind or it's holding them in abeyance for the period of the drug's action, that were obscuring this state of being that should in principle always be available to me. That's what I came away from it feeling. Like, I had this experience of four hours or so where I lost all of my self-absorption, all of my self-consciousness, all of my concern for what other people think of me. I experienced just a perfect, free attention to both to care about others and to recognise that I do care about others, effortlessly. That my default is to want others to be happy. And to really want them to be happy. Not just checking a box on an inventory of nice attitudes, but that my own happiness was bound up with theirs, and that that was love. And that it was a feeling I felt not just for the best friend who was sitting across from me in our living room, but I would have felt it for a stranger had he showed up at that moment.

It was a way of being. It was not a transactional feeling of love: I love you because of our history together. I love you because of how much fun I have in your company. No, it was a feeling of: Love is potentially the default state of consciousness in the presence of other beings struggling to be happy in this world. And, it just went very, very deep. It was incredibly beautiful. Most important, it was a state of mind that on some level was totally foreign to me. It's not that I didn't love people. I love my mum. I love my friends. I'd had a girlfriend or two by that point. But this was just the depth of it and its unity with my feeling of my own being. It was a proper spiritual/mystical experience, but one from which I came down. And, then having come down I was left

to wonder, 'Okay, what do I do about this?'

It was clear to me that if this was a feature of the human mind with or without drugs – and there was every reason to believe that's true because drugs don't cause the brain to do anything that the brain isn't capable of doing. All they can do is mimic neurotransmitters or change the behaviour of neurotransmitters in one form or another. So, this is your brain, in one of its states, however it got there. So, I just became very interested in finding a path by which I could integrate that wisdom and that kind of feeling of wellbeing more and more into my life. So, meditation became a major part of my life. That's a long-winded half-answer to your question.

As far as my meditation practice, briefly: I've practiced many styles of meditation but mostly in a Buddhist framework. I've spent a fair amount of time in India and Nepal studying with various teachers. And they were not all Buddhist, but it's been mostly Buddhist. Ultimately, I spent a lot of time doing Vipassana practice on silent vipassana retreats, which I know you're familiar with. I had spent about a year on silent retreats, over the course of a few years. The longest retreat I ever did was three months at the Insight Meditation Society in Barre, Massachusetts, and I did that twice. But then I did multiple two-months, and one month, and many ten days.

Then, ultimately, I connected with a practice called Dzogchen, which is often considered the highest non-dual teachings within Tibetan Buddhism. So, I spent a fair amount of time with Tibetan lamas – otherwise known as Dzogchen masters – who specialise in that practice studying that technique of meditation, and spent some time on retreat doing that as well. After I got that practice in hand, my experience of mindfulness changed in ways that I'm at pains to describe over at *Waking Up*, my meditation app, because I do think ultimately there's a very important difference between what I call dualistic and non-dualistic mindfulness. And, it's possible to spend a lot of time practicing dualistically in ways that are ultimately frustrating if you really are taken with the project of getting to the bottom of these things. And, ultimately encouraging of certain illusions, which can be painful. I was in the grip of those illusions for the longest time, even practicing quite diligently. Even spending months at a time on silent meditation retreats, meditating twelve to eighteen hours a day, I was still seeking to solve a problem that was fundamentally misconstrued, I would say. And it's not for want of guidance. I was studying with really great meditation teachers. But there was a certain logic to that practice that I think is misconceived, and perhaps we can talk about that.

So now I consider myself someone who practices very much in the spirit of what I would call Dzogchen practice, but there are other non-dual ways of coming

“

I've never respected the boundaries or apparent boundaries between fields of knowledge.

”

at this. Mahamudra practice is like that. And, I would argue within the outside of Buddhism, within the Indian tradition, Advaita Vedanta is articulate more or less on the same points and recommends more or less the same style of attention, although most Buddhists would hate that comparison.

RR: Your description of MDMA reminded me of a couple things. One, Gary Greenberg was a past guest on *EconTalk*. In one of his books, he describes a similar experience – I can't remember which drug he was under the influence of – but of just vast love for anyone he was encountering. He said something like, 'The goal of life is self-transcendence.' You said something like that, right?

SH: Yeah.

RR: I always think of that as growing up. Growing up is about learning. I don't know if it just takes time, but it also might take some effort. It's learning that you're not the centre of the universe. It's really hard to learn that lesson. I have a new granddaughter. She's five months old, and she really is the centre of the universe. And it'd be perfectly normal for her to carry that view for many years into her adult life.

But, what I experienced, it definitely came out of my silent meditation retreats. I don't have a very active practice of meditation now. I'm thinking of revitalising it and maybe it will help me. But, what I was able to experience in the aftermath of the retreats I went on was a glimpse of what you're talking about: a feeling of self-transcendence and a connection to other people that you weren't aware existed until it came upon you. You realised that much of your way of moving through the world was very narrow, and you



Photograph: Jared Rice

had cut yourself off and left yourself unavailable to things that are real. Once you've tasted that, some people want to take the drug more often, some want to meditate more often. For me, I only get a glimpse of it now and then. I find it very hard to maintain, and partly because I don't maintain the practice.

But those retreats did have a permanent effect on me, which is fascinating in and of itself. They were only five days, a year apart, three times. And, practicing in those years – daily or often – was an opportunity to see yourself in a different way; to see yourself connected to other people. And, as you said (and you said it very beautifully), to imagine that their happiness could be paramount: not because it made you really happy; just because it's the way it should be. I wouldn't call it a rational thing.

I'm very grateful for that. I wish I felt it more often. I try to feel it as often as I can. But I do think that we have access to that – through, I've argued, fiction, therapy, meditation, religion – all, in theory, can help you transcend yourself. It is a fundamental opportunity for us as human beings because it's not the way we're hardwired.

And I would say that there are gradations to transcendence. I mean, there are different modes in

which we can transcend ourselves. The ultimate mode for me is transcending the very sense that there's a 'you' – a subject in the centre of your experience. That's the final illusion which is targeted very directly by meditation. As I said in the beginning, and for the longest time, meditation can seem to ramify that sense of self because you feel like a meditator. You feel like the one who is directing attention at the object of meditation. And now you're becoming mindful and you're discovering it's hard to do that. You get lost in thought, and then you come back.

There are obviously gradations of selflessness and selfishness and ways in which we can discover a bigger view of the project of becoming happy in this world that admits that we're, on some level, all in this together; and that another person's happiness – certainly a friend's happiness – is not in zero-sum contest with one's own. In fact, it can become very directly a reason for one's own happiness. One of the worst revelations about the poverty of ordinary selfishness is to see those moments where one can't celebrate the success of a friend because one is envious. I mean, that is just the ugliest little wrinkle in the ordinary human psychology, which is just to feel diminished by the happiness of another – even one

“

... I had this experience of – for lack of a better phrase – unconditional love. It was an experience of not feeling high or altered or stoned, and it was certainly not a sense of my perception of the world being distorted. It was actually an experience of being sane for the first time in my life.

”

who you ostensibly love. This is not the highfalutin' esoteric claims of spirituality. This is just ordinary human friendship. We know that really being a friend requires that you actually want the best for your friends. That includes wanting them to be happy, wanting their hopes and dreams realised, and being able to smile and celebrate when those successes occur.

And, when you find that you can't do that, with that cramp of self-concern and that false structure that suggests some kind of zero-sum contest there, that's the kind of thing everyone wants to transcend. Whether they think about it or not. It doesn't feel good, it doesn't feel right. It's the antithesis of love in that moment. I mean, love in that moment is – to use the Buddhist phrase – the ability to feel what's called sympathetic joy: just the actual, the contagion of sharing the joy that your friend is feeling at that moment. We do this as parents. We do this fairly effortlessly with our own children. But somehow it becomes harder in other contexts.

Clearly, mental health, ultimately – the norms of healthy, happy, rewarding, ethical engagement with other human beings – must push us in the direction of feeling that freedom to be happy with other people

more and more. And to extend it: extending the circle of that even to strangers.

Many people feel this quite effortlessly. But we should notice where we fail to do that, and the consequential moments in which we fail to do that. Those are obvious opportunities for growth. That is far more ordinary than the esoterica of transcending the sense of subject/object dualism in meditation. But it's along the same continuum, I would argue.

RR: I can't help but think of Gore Vidal's quote – how much of it is tongue-in-cheek – 'Every time a friend succeeds, I die a little!' Not really the ideal, as you suggest. Adam Smith had a lot of interesting observations about this: Our ability to sympathise with the joy of others and to empathise or sympathise with the grief of others. What he doesn't write much about, and what you're talking about, is the opportunity to change the way we respond to those moments in life.

I just want to mention, you know a lot more about this than I do, so you can tell me if I'm wrong, but I find the modern emphasis on mindfulness to be rather empty. It's good to be mindful. It's a good thing. It's a tool. But, mindfulness by itself, which is often sold as a goal and as a good, has even been mocked as 'McMindfulness' – the corporate meditation session. But, mindfulness by itself is not, to me, very interesting. It's what you're talking about. And meditation, correctly, either from a Buddhist perspective or Jewish perspective, those are the only two I know was a little bit about, they do something much more than telling you to pay attention. They're telling you about how you should move in the world. I think that is profoundly more important. Is that a fair critique?

SH: I agree. Even within Buddhism, mindfulness is just part of the path. It's just one of the eight folds of the eightfold path. So, yes I think the project is much larger than paying attention. Paying attention is a necessary piece.

Generally speaking, I would say there are two levels at which one needs to work on oneself. Both relate to the nature of thought, because thought is really what captivates us and deludes us and anchors our sense of separateness. It's this identification with thought that is the problem from the point of view of meditation.

So, there are two levels. There's the level at which we can change our thoughts and we can learn new things to think, and we can believe differently. Our beliefs matter, and they govern our behaviour and they dictate the causes to which we are purposed and the goals we form and all of that. I mean, it's all thought. Thought is the thing that makes us human. Changing one's thoughts also gives us the power to

reframe experiences and change our relationship to them. So, you can transform your sense of wellbeing and your relationship to your experience and to the rest of the world by how you think. One example which I think is crucial for living a good life is to have an ethical code, and to decide very clearly on things like whether it's okay to lie. Your relationship to telling the truth or not is very, very important. One very strong heuristic for me is that it's almost never the right thing to lie. It's almost always going to complicate your life unnecessarily. It's almost always an expression of fear and selfishness and separateness. I wrote a book titled *Lying*, and it's just my analysis of all the reasons we have to decide in advance that it's generally not okay to lie.

Now, there's an exception to this, and I would put lying on the first stop on the continuum of violence – that is appropriately used under conditions of self-defence or in defence of other people. So, I'm not a pacifist. I think pacifism, when you actually get to the bottom of it, is actually an obscenely immoral view, and dogmatic commitment. So, Gandhi and pacifism – it's worth remembering – had him saying things like he thought the Jews should offer themselves willingly to the crematory of Auschwitz so as to shock the conscience of their captors. That's not my view of an ethical world. I think at a certain point you have to start killing Nazis. And if you're going to carve out a space in which it would be ethical to kill Nazis in self-defence, well then it would also be ethical to lie to Nazis before you start killing them. So, lying is: If someone knocks on the door and says, 'Do you have Anne Frank in the attic?' I think it is ethical to lie to that person.

But in the normal course of events, I think lying is needlessly corrosive of everything we care about and sets people up for reputational destruction – or should: I don't actually think there's nearly enough opprobrium attached to being caught lying in our society. That's all to say that that's the kind of thing that has to be thought about and reasoned about and argued about in advance. This is all the domain of thinking good thoughts in the end.

Mindfulness – and any other meditative instruction along those lines – is not a matter of thinking new thoughts or thinking about anything at all. It's a matter of noticing very clearly what is arising in consciousness in each moment, including thoughts and emotions and everything else. The thing about mindfulness that is distinct from all of this is that it's not a matter of understanding anything conceptually. It's a matter of bearing witness to the mechanics by which we become entangled with thought, the way thought drags into view emotions, positive or negative. And much of this is the mechanics of our own suffering, because so much of our conversation with ourselves is an unhappy one.

Mindfulness ultimately is a matter of breaking this spell. It's not a matter of suppressing thought. It's not a matter of getting rid of thought. It's not a matter of viewing thought as the enemy. But, ultimately – and this is a statement of what the goal is from a contemplative point of view, this is invoking an analogy from the Dzogchen teachings – ultimately, thoughts become like thieves entering an empty house. There's nothing for them to steal. There's no implication for one's sense of wellbeing presented by the next-arising thought. It's just: there's no problem.

People are often taught that thought is the antithesis of meditation, or the antithesis of mindfulness, or the direct enemy of it. It's natural to think that and feel that because, in the beginning, what you're faced with is the profound distractability of your mind. It's just damn hard to pay attention to anything. And so it seems like thought is the enemy. But ultimately, thoughts are just arising all by themselves in this wider context of conscious awareness. There is no thinker in addition to the flow of thought. There's no subject authoring the thoughts. One thing that is implied by this – which I don't talk about all that much, but I think is true – is that even for non-meditators, even for people who have no idea what we're talking about now, the sense of self, the ordinary sense of subjectivity, is being interrupted all the time. It's just not being noticed.

I would say that much of our feeling of selfhood is interrupted in that kind of punctate way throughout the day. It's only in retrospect when we're reflecting on it or recoiling back into self, in relationship to others, that our sense of egocentricity, our sense of being truly centred in our experience becomes quite vivid. We're just very bad witnesses to all the changes and all the perturbations and interruptions. Meditation is the act of, ultimately, consciously interrupting it in a way, becoming mindful of those gaps and even provoking those gaps deliberately with mindfulness.

RR: My guest today has been Sam Harris. Sam, thanks for being part of EconTalk.

SH: Pleasure, Russ. Thanks again.



EconTalk

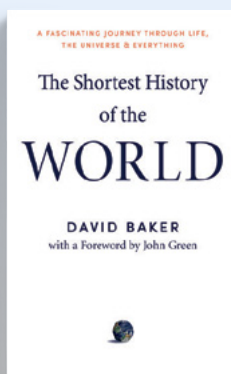
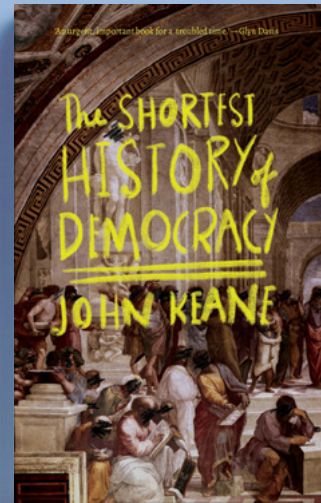
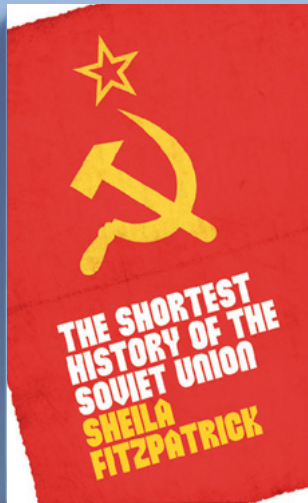
*Published in 2023 and transcribed for
The Podcast Reader.*

 **Read more @ podread.org**

 **Listen @ EconTalk**

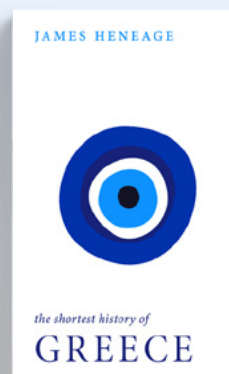
Expand your world.

Clear and concise histories, written by acclaimed experts



A fascinating journey through life, the universe and everything

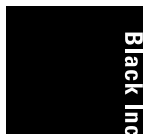
In this compelling and revealing book, David Baker traces the rise of complexity in the cosmos, from the first atoms to the first life and then to humans and the things we have made.



‘From Alpha to Omega, and from Socrates to Syriza, here is all of Greek history, expertly told.’ —Tom Holland, author of *Rubicon: The Triumph and Tragedy of the Roman Republic*

DISCOVER THE BESTSELLING
SHORTEST HISTORY SERIES

BLACKINCBOOKS.COM



On the Importance of Grit

Adversity, human excellence and happiness

ANGELA DUCKWORTH
IN GOOD COMPANY, 2022

Interview by Nicolai Tangen

Nicolai Tangen: We've had a lot of really, really interesting guests on *In Good Company*. And, interestingly, pretty much all of them talk about the importance of grit. So, to investigate, we are thrilled to have the world's leading expert on the topic, Angela Duckworth, with us. Very welcome, Angela.

Angela Duckworth: Nicolai, I'm so excited for this conversation. It's good to see you.

NT: Let's start with the basics. What is grit?

AD: I define grit as the combination of two things: passion and perseverance for long-term goals. Not just perseverance for long-term goals, not just being resilient and taking feedback and trying to get better every single day, but also doing all of that for something that you love. Passion is a word we use for romance, for the people we fall in love with. When I give you a picture of what a truly gritty person is, you should see the sweat, you should see the bouncing back and the optimism moving forward. But you should see all of that in service of something that that person stays in love with for a very long time. Years, decades – sometimes a lifetime.

NT: How do you measure it?

AD: As a researcher, as a psychologist who studies excellence, I typically use a completely fakeable questionnaire called the grit scale. When I was just starting out studying high achievers – and that's really what I do – I tried to reverse-engineer human excellence. To figure it out, see if we can begin to emulate some of the things that high achievers do. I interviewed people like you, and I started to hear recurring phrases like finishing whatever you begin, of setbacks not discouraging that person. And also this language of love, this language of abiding devotion to something. So, I wrote those statements down into a questionnaire and it works for research with people who have zero incentive to fake their scores. I think from a practical standpoint, if you're hiring or if you're admitting students to university, I don't think this scale works at all because people are...

NT: Going to fake the answers.

AD: Yes. It turns out that there are a number of different problems with questionnaires. Let me just highlight two of them. The more obvious one is exactly what you said: regarding a question like 'I finish whatever I begin', I know what the right answer is, as opposed to a math problem where you cannot fake or guess the right answer in a questionnaire of your character. So, I think that's the big limitation and maybe the more obvious one. But let me tell you about another one, which is very interesting because it lends a window into how human beings think of themselves even when they're



being honest. In research that I haven't yet published, but is in the process of being published, what I found is that for questionnaires, people, when they get a question like, I'm a hard worker or I finish whatever I begin, they have to make comparisons with other people. What does it mean to be a hard worker? Compared to whom? So, Nicolai, I could compare myself to you.

NT: Well, I'll tell you one thing. You'd be a very hard worker compared to me.

AD: I'm guessing that in our peer group, what it means to be a hard worker is different than, say, a typical 14-year-old kid. There is a phenomenon called reference bias, and that is that we all have a frame of reference, as we have to have. We cannot answer any question. I mean, novelists like Herman Melville said, there is no experience except through comparison. Is it cold? Well, it depends on compared to what? Is it warm? Compared to what? Am I happy? Compared to when? So I want to say that another problem of administering a questionnaire like the grit scale, if you're hiring, for example, is that people have different reference points. And in some ways, paradoxically, the person who is going to give themselves an extremely high score on the grit scale because they really think they are resilient and tenacious and devoted might have lower standards. The score can go in exactly the opposite direction. So for all these reasons, the grit scale is all I have for research, but it's highly imperfect.

NT: Starting with your scale then. What's the relationship between the outcome on your scale and people's success?

AD: An emerging consensus among social scientists is that, like the weather, human destiny is extremely difficult to predict. So when I tell you that grit predicts outcomes like graduating from arguably the most competitive military academy in the United States, yes it does, but with huge sample sizes. It's explaining some of what happens to people, but the vast majority of the variation in people's outcomes remains a mystery. In other words, I want to say, as a social scientist, let's not pretend that life is as simple as saying like, I saw that kid at seven years old, and I know what they're going to be when they're seventy-seven. You see a kid at seven years old, you can make an educated guess, but you will probably be surprised at what happens to that seven-year-old kid even by the time they're nine. So I want to be modest, appropriately so, when I tell you that when I administer the grit scale – this highly fakeable twelve-item questionnaire – to, say, cadets at West Point, the US military academy. To enter West Point you need top grades, top test scores, you need to be an elite

athlete, even just to apply you have to be nominated by a US Congressman, senator, or the vice president of the United States. So it's a very kind of fine-mesh sieve that these young people go through. Even so, historically, there's been a very high attrition right at the beginning of the program. When you first get there and you come from your high school where maybe you were captain of many teams, you were the star, you were sometimes the valedictorian, the top student.

I think what's happening in those very first weeks is that not only physically gruelling, not only is it wake up at dawn, work nonstop till midnight, not is it only like a socially difficult time, but most important, I think you are no longer the star. Statistically, half of these extraordinary women and men are below average for the first time, I'm sure, in their lives. When we measure grit on day two, so that's basically the day after you get there, we can predict prospectively the likelihood that you will make it through those high-attrition periods of training at the beginning of your four years at West Point, but also we can predict your graduation four years later and your successful transition into the army, which is really the contract that you make when you sign up for West Point. So, grit ends up being a better predictor than, for example, objective measures of physical talent or objective measures of intellectual talent. So that is one of the reasons why I do believe grit is worth a conversation without, I hope, overselling it as the only thing that matters. Or the idea that if I give you this twelve-item questionnaire I'm going to be able to predict whether you're going to be a successful entrepreneur. Yes, grit matters. Why it matters is something we can talk about as well. But it's not a simple story the way sometimes it's understood.

NT: It's quite extraordinary that it can predict these kinds of things, yet you can't use it as an entrance exam because people are just going to fake it, right?

AD: I have to say Nicolai I have yet to meet the CEO who isn't interested in hiring people who have this quality. Because when you look at Olympic athletes, when you look at people who win the Nobel Prize, when you look at people who build truly great lasting companies, they have this passion and perseverance and this stamina. Take Isaac Newton on why he discovered the laws of mechanics. Why Newton? Was he head and shoulders smarter than every other physicist and mathematician to ever walk the planet Earth? Well, Newton, appropriately, of course, recognised that he was standing on the shoulders of giants. But he also said that he kept working on the problem when other thinkers walked away. So I think CEOs are right for looking for this quality of passion and perseverance for long-term goals. I would love to

“

... grit ends up being a better predictor than, for example, objective measures of physical talent or objective measures of intellectual talent.

”

hear your ideas about how you would pick up on this in the hiring process. I have a couple of ideas and, of course, none of my good ideas have anything to do with a fakeable questionnaire.

NT: It's very interesting. I don't think necessarily we have a lot of Newtons walking around in the investment fund where I work, but it is for sure important. We need to take losses and wake up the next morning and continue to take risks, so grit is one of the things we really look for here. But is it possible to develop it?

AD: I think absolutely, yes, that grit can be developed. Now, why do I say that? I have a couple of reasons. One is that the research on human development over lifespans, and in addition to the neuroscience on what happens to the brain over lifespans, affirms definitively that there is much more plasticity throughout the entire lifespan than anyone could have thought. Even thirty years ago when I graduated with my degree in neurobiology, I was trained to think that essentially after you're a young child, perhaps after adolescence, nothing very exciting happens to personality or the brain. We now know that's not true. There's neuronal growth and remodelling, and there are actually documented changes in character and personality. Not just ones that happen with experience in a kind of slow and steady way, but also intentional ones like therapy, for example, which has been shown to reliably change personality in desirable ways.

The second reason that grit can be developed comes from the work of Sule Alan, an economist, who has recently published a random assignment study. Essentially, the intervention was to increase grit and also growth mindset, the belief that your abilities are changeable.

Sule Alan had read the research on grit and growth mindset. She read, as an economist, that psychologists like me and Carol Dweck were showing that people who have this belief of I can change, people can change, tended to be grittier. They tend to just stick with things, get up again and learn, and move forward. She decided to do a random assignment study in Turkey, which is her native country. She randomly assigned late elementary school students to a curriculum with messages that were about grit and growth mindsets, and a control group which did not have that. And what she found was that, over time, including up to two years later on standardised tests of academic achievement, there were measurable benefits of this growth mindset curriculum. So that's children, not adults, but it's a demonstration that people can and do change.

One could argue that as a result of listening to a conversation like this or other conversations that you've had, that you can learn things. That is, I think, a distinctive feature of human beings. We learn better and for more of our lifespan than any other animal.

NT: Do you have examples of an organisation where you think they have really moved things in the right direction?

AD: I have examples of a few organisations, and then quite a number of companies. It's palpable to me at the top when I meet the leader. On an occasion when we were together a few years ago, I quoted Ralph Waldo Emerson that an organisation is the long-extended shadow of its leader. For example, a leader like Toby Cosgrove, who was arguably one of the greatest cardiac surgeons when he retired from surgery and took on the CEO position of the Cleveland Clinic, arguably one of the very best medical institutions in the world. He is grit incarnate. He was profoundly dyslexic as a boy, was told many times that he didn't have what it took to be a physician, much less the most competitive kind of residency you could do, which is cardiac surgery. Every time somebody told him, no, you can't, he said in response, in so many words, I'll show you. When he took over the CEO position, it was the opposite of complacency, the opposite of we'll just rest on our laurels because we're already famous around the world. He revolutionised things. He brought empathy to the forefront of care. He took a quantum leap forward in making the medical records electronic, etc.

One other place that I have found to be remarkable and maybe a lesson for all of us who are trying to increase grit at the level of an entire company or an entire team – there is a school called Expeditionary Learning. It's really a school system, or a chain of schools. These are non-profit schools, and they are all in this philosophy of really working hard, but

also demonstrating other aspects of character, like teamwork.

When I not only talk to the very top, the top leaders, but also when I interact with teachers and students at these schools, I understand that culture is not just do you live in Norway or do you live in Sweden? Do you live in New York City or do you live in Paris? Culture is any place where a group of people have a shared set of beliefs, of values, of traditions, of language and identity. When I go to a school like Expeditionary Learning or I walk around the Cleveland Clinic and I talk to professionals there, or I visit a sports team, it's just the same as getting off of a plane in a new city. There's a culture here, and I do think cultures play an enormously important role in encouraging or discouraging qualities like grit.

NT: You mentioned dyslexia. Talk a bit about adversity and how that impacts it, and perhaps also touch on post-traumatic stress syndrome, which kind of takes that to the extreme. How does that impact it?

AD: The question of whether adversity is good or bad for character development is as old as a civilisation. Nietzsche, of course, famously said that what doesn't kill me makes me stronger. That's the question. Is adversity the crucible of our character, or is it not? I think the reason why it remains a question is that it's not a simple answer of yes or no. I think with adversity, it's not a linear relationship. In other words, there is probably more of a curve where there is some amount of challenge that is beneficial. Think of your own children. How can they grow up to be strong and wise and capable without some adversity in their lives? At the same time, you would never wish for your own children or anyone else's, to suffer the highest unmitigated dose of adversity. When you mentioned PTSD, that is arguably the kind of trauma that is not what Nietzsche or anyone else would say makes you stronger.

I want to add one thing here which may be even more important, and that is that challenge without support, I think, does kill you. I mean, challenge without support creates what some neuroscientists call allostatic load. It's the wear and tear on your immune system, and on your motivation and self-esteem. That happens when you are dealing with things that you cannot handle without any support for prolonged periods of time. That is why what we can do as leaders, as mentors and as parents, is so important. If a young person or anyone else is struggling to do something they cannot yet do, but they know that they're loved and they know there's a foundation under their feet. They know that there are people who care about them, who will lend a hand, give advice. Have a conversation on a Saturday morning, have another conversation on a Sunday

“

I think, and research suggests, that happy people perform better at work, they're better at making relationships, they make better decisions, they're more creative. I think happiness is not only a signal that things are getting better, but a cause of things getting better because you are now engaged, you're not withdrawn, you are also attractive to other people.

”

morning. That, to me, is the recipe for character development, for healthy growth: some amount of adversity, but with an enormous amount of support.

NT: What is your view on soft parenting?

AD: You mean parents who solve all problems for their kids? It goes by different names. Sometimes it's called snowplow parenting because you get all the obstacles out of the way sometimes. It's called helicopter parenting, which emphasises that you're chronically monitoring. I was in a conversation recently with an investor like you and he had a very thoughtful and evolved philosophy not only of investing, but of life and of parenting.

He said something that I won't forget. He said, I have thought a long time about what I need to give my children and I believe that the most important thing, and the thing that I keep front of mind at all times, is that it is my job to help them be independent.



Essentially what I think he was getting at is that if you keep in mind that your goal as a parent is not necessarily to make your kids elated at all times or to have the easiest life, but just to help them become independent. Independent of you. That then recommends against the kinds of things that I think we're all probably guilty of, like solving a problem for a kid. Going back and getting their textbook that they forgot.

If you think about how history has changed, people have so many fewer children today. I do wonder whether we're being overprotective and we're losing sight of really what childhood and adolescence are for. They are for, of course, being happy while you're there, but also for becoming independent. Because you do want to grow up. I do worry that of myself and a lot of us, that we're not allowing our children to fall down, to get a bad grade if they screw up, to miss an assignment, to deal with some of the things that are just part of life.

NT: Are we the cause of some of the mental problems that young people have today?

AD: First I want to say that when we talk about mental health, if anybody asks me is it really true that there are changes in mental health in recent

times, I want to say 100 per cent unequivocally. There is some doubt in some people's minds regarding is it just because we're measuring it better? But if you look at hospitalisations, admissions to emergency, suicide attempts, there are a lot of objective measures that corroborate questionnaires and surveys that say people of all ages, but especially adolescents, are feeling sadness, depression, anxiety, insecurity, low self-esteem in ways that are different. I think that's partly the pandemic, but I think actually you could look at some of these trends as preceding the pandemic.

I want to say this is a very good question Nicolai because it's something that has to be explained. Are parents the primary cause? I don't think so. But I don't think science has a definitive answer. Is it social media? There are some studies that suggest being on social media and flicking through an endless stream of curated photos of your same-age peers, all of whom appear to be having a much better life than you, seems to be negative for self-esteem, for mental health. But it doesn't have an effect size, a magnitude of an effect that is likely to explain the whole thing. So it's not simple. It can't just be social media. It can't just be changes like how much time parents are spending with their kids or how much they protect them. I do think, though, that there must be some

greater attention to this, and we're beginning to see some of things. I have one theory that I think is really interesting. Some years ago, speaking to a physician, she said in her humble way, I don't know what's going on with the increases in anxiety, in particular, but also depression, but she said, I wonder if it's that young people aren't spending enough time outside. I just wonder about nature and our separation from nature. Since she said that casually, I've been looking at the research literature, and there are not only very strong correlational studies that document a positive relationship between positive mental health and green space. Also, there are random assignment studies that show that when you are in nature, your attention goes from within yourself.

One of the problems with adolescents is your focus is entirely on yourself and your own problems and how unhappy you are. And when you are in nature, nature, in a very gentle but very enticing way, draws your attention outward. I don't want to say it's all about nature and being out in green space, but I think a number of factors have changed the way young people are growing up today. And I do wonder whether this is exactly the time to basically make our own history? To not just let these parents take us where they will.

NT: Recently, we had a guest on the show, David Solomon of Goldman Sachs. He said that if you are happy 70 per cent of the time, that's pretty good.

AD: Did he mean by that some amount of unhappiness is necessary for productivity?

NT: Just that you need to hang in there at work, take some bad days and not give up. What do you think is the optimal proportion of happiness? How happy do we need to be?

AD: Well, first of all, I like the precision of the comment, even if it's just a way of making a point. I think he's right in that no person can be or should want to be a ten out of ten on happiness all the time. I mean, I want to say something about emotion. Happiness is an emotion. Sadness is an emotion. Anxiety is an emotion. Jealousy is an emotion. Fear is an emotion. These are emotions that we have because of evolution. These are emotions that help us adapt and survive. Well, what is happiness for?

If I think about Tim Beck, the creator of cognitive therapy, sometimes known as cognitive behavioural therapy, which is really what modern psychotherapy is. Tim Beck, in his one hundredth year, was working on his magnum opus, as he called it. He lived just a couple of blocks away from me, so I would see him on Sundays. We got to talking about self-esteem and about happiness as an emotion.

“

The older you are, especially throughout adulthood, the higher your grit score. It's not like the correlation is one, but it's highly reliable.

”

His theory was happiness is the emotion that comes from the thought that my self-esteem is going up in some way.

You're having a win. Things are getting better. And that produces the emotion of happiness. Now why would that be important from an evolutionary standpoint? Well, that's a very good thing for an organism to know that things are getting better. But we also need to have the emotion for loss, that things are getting worse. That's sadness. We also have to have an emotion for things might be getting lost. That's anxiety. We also need an emotion for someone's taking advantage of me. That's anger.

So, we shouldn't try to be happy 100 per cent of the time. That's true. We don't want to have any emotion 100 per cent of the time because these are signals of how our life is going. So, I want to say that happiness is an emotion that comes from the sense that things are getting better.

Let me tell you about a study that I'm about to publish with some collaborators from Korea and China, where we look at happiness over time. We have, in particular, a sample that was collected in China with young people, and we measure their grit, and we comprehensively measure their happiness. Not just how satisfied you are with your life, but we also index an array of positive emotions and an array of negative emotions. We can create a composite wellbeing score that is the scientific standard for how you measure happiness. So we have the grit scale and we have this happiness measure. We do this for years and now we can ask the question not only how are they correlated, but which leads which in the stronger direction? But what surprised all of us when we looked at the data is that happiness is a stronger

predictor of changes in grit than grit is a predictor of changes in happiness. It seems to be more happiness driving grit than the other way around.

NT: Why is that?

AD: I think that happiness truly is causal. This may be one of the most important things that we cover in this conversation. I think, for a long time, people thought, like Tim Beck, that happiness was an emotion, that was a signal that things are getting better in your life. And that is very important because a signal is important. But I think it is also causal. I think, and research suggests, that happy people perform better at work, they're better at making relationships, they make better decisions, they're more creative. I think happiness is not only a signal that things are getting better, but a cause of things getting better because you are now engaged, you're not withdrawn, you are also attractive to other people.

NT: But how do you get there in the first place?

AD: So, if you're sold on the fact that happiness is not only a great outcome, but also an input, how do we get more? I think the summary of the positive psychology literature over the last thirty years is this: that there are a number of exercises that people have studied that help. Like the blessings exercise, where you think of three good things that have happened to you in the last few days. There's a gratitude letter. There's knowing your strengths. Just so I can illustrate this one and make it real, would you mind telling me three good things in your life that have happened over the last few days?

NT: Well, I attended a board meeting today, and some of the board members really cared about how I was doing. I have spent time in nature lately, and I have spent time with the family.

AD: Wonderful. And that made you feel, I'm sure, very happy. If I extrapolate from my own experience with my family, I could think of three good things, too. It was Father's Day here in the United States on Sunday, and my husband is amazing. He is a wonderful father. My daughter wrote him a letter that brought tears to his eyes, and then when he passed it around, it brought tears to mine too. I had a lovely dinner with my mother-in-law. And the third thing is I'm having a fun conversation with you. So that's the three good things exercise. I'm using this as an example of what positive psychology has found to make people happier.

It took me less than sixty seconds. Our attention was not on bad things. It wasn't on things that you or I can't change. It wasn't on the things that keep us up

at night in a bad way. It was just on three wonderful things. And the thing about human attention is that, as Daniel Kahneman, the Nobel laureate, would say, what you see is all there is. So when I bring my attention to something positive, it takes up the whole psychological field of view. Now, the default for human nature is to dwell on the negative. Evolution has taught us that we should constantly scan the horizon for threats and bad things. But what positive psychology would say is one of the major paths to happiness, and perhaps the royal road to happiness, is to draw your attention to things that are good, to look for the good in other people, to look for the possibility in a company. Look for the positive. It is a form of optimism, I guess.

NT: So how does it differ geographically?

AD: So remember we talked about reference bias and people make near comparisons. It turns out they make very near comparisons. For example, when I study children in a school, they're not even comparing themselves to other kids in the school. They're only comparing themselves to the kids in their classes. So if you want to know how your son or daughter thinks of themselves just look at their close peers. That's their comparison set. So it's very difficult to compare then somebody from Finland to somebody from Phoenix in the United States because the comparisons are just like who you see around you.

That is why with international studies of character and personality, you often get very puzzling findings. Either you get findings like everyone's the same, there are no differences in national character, which seems implausible for anybody who's ever taken a plane outside of their country, or you get paradoxical results. Like in one very well done study they had fifty-six countries take the same personality questionnaire and on that questionnaire there was a category called conscientiousness, orderliness, responsibility, duty, and the countries that came out the lowest out of fifty-six were Japan and Korea, which doesn't seem right. I think that's explained by the standards to which those citizens hold themselves for being orderly, for being punctual, for being responsible, being really high. So it makes as a scientific question are the Finns really the grittiest people very difficult to answer.

I will give you my anecdotal hunch, and I have to say it's a hunch and not based on data. I think there are countries that clearly promote persistence as a value. When you think of, for example, Japan and Korea as countries where they have aphorisms like eat bitterness, they have all these like hard work sayings, and if you look at the number of hours that people work or study, you think wow, they're the grittiest. But I want to say this as a hunch.

I think some of the countries that are highest in

perseverance are actually very low in passion. I think it will be to their own demise. I think if you have a culture that is all about hard work but doesn't have the ability for people to work on things they intrinsically care about, they will never win the Nobel Prize. You will never get to be truly great at something that you're doing because somebody else told you to, and because, well, that's your job. You have to do what you love as well. That's something that you would understand. To work hard at what you love is entirely different than working hard at something that you do not love.

NT: Last question on this. Do you see differences between gender and also age?

AD: I find a very reliable difference. This has been replicated in labs around the world, so I think I can defend this one without any equivocation. I find a very strong relationship or reliable one between age and grit. The older you are, especially throughout adulthood, the higher your grit score. It's not like the correlation is one, but it's highly reliable. In other words, it's been found over and over again.

What I also have investigated is gender differences. The reason I did was because so many women have come to me and said, we must be grittier. On the gender issue is that I have not found a reliable difference between women and men. I can't say whether that's because of reference bias or anything, but I do not find a difference favouring either sex. If there is a difference, it must be a tiny one, at least on these questionnaires.

NT: I'm a big believer in your sample size.

AD: Yes, sample size is great.

NT: Now, what else are you working on for the moment? What can we expect from the fabulous Angela Duckworth?

AD: I don't know how fabulous I am. I'm working on two things that I'm very excited about. Firstly, when I studied grit, I wanted to reverse-engineer high achievers. I'm obsessed with excellence, and I thought, it can't be a mystery. It can't just be a God-given gift. Let's figure out the mindsets and the skillsets of these extraordinary women and men so that we can be a little bit more like them. The next logical thing is what are the circumstances – the objective, outside circumstances – that enable you to develop these internal mindsets and skillsets that then make you happy and successful? In other words, I'm swimming upstream, if you will, to figure out, what are the schools that enable you to develop a growth mindset? What does the parenting look

like? What about where you stand in society? So I'm swimming upstream and I'm trying to write a book that's provisionally called *Habitat*, which is about the objective circumstances that lead people to thrive.

Secondly, I've been working on interventions to increase not just grit, but self-control and happiness for my entire career as a psychologist. I'm pretty disappointed, I'd give myself maybe a C-minus because some things work but most things don't, and nothing works for very long. About three years ago I had a bit of an existential crisis, a midlife crisis, if you will. I was like, what the hell am I doing, if this is only going to have brief, small, unreliable effects. In the past I would do very quick interventions.

The direction I want to move in are interventions that are much longer, that are explicitly educational and that are social, so that you are not learning on your own, but you're learning in a group. To me, this better matches the few cases where we saw somebody really makes a sea change, where they have understood something they really didn't understand before. Most often they do it in a group. It's not just that they are changing, but maybe their whole family is changing or they and their friend group are changing. I think the future of behaviour change is social. I think the future of behaviour change is educational. I think it's experiential. I don't think people change just from information. They need to have an actual experience that they feel in their bones, and certainly in their hearts, for them to make a lasting change.

NT: That's a fantastic place to end. Angela, not only are you the grittiest person I know, but you are also the cleverest. Such a privilege talking to you.


AD: I hope we have another conversation soon. I really enjoyed this one.

NT: I can't wait. Thank you so much.



In good Company

Published in 2022 and transcribed for
The Podcast Reader.

 Read more @ podread.org

 Listen @ [in good company](https://in.goodcompany.no)



Your listeners are readers too



Autoscroll On

Your listeners are readers too. You are reading text in an image of an interactive podcast transcript within a magazine. It sounds complicated. But it's actually not.

Sometimes, it's easier to read than to listen. It's easier to search text than sound. It's easier to re-read than to re-listen. We think you get it. After all, you're reading a magazine of podcast transcripts.

Transcripts are even better when you can not only read them, but also listen to them at the same time. At the end of the day, your listeners are listeners too.

That's why we highlight the word being spoken. Here is where you would hear me say "here" if you listened to this ad as an interactive transcript.

Fanfare makes high-quality transcripts for knowledge-dense podcasts. Our heart beats for podcasts with show notes that put the appendix of a research paper to shame. Podcasts that make you reach for a pen as you're listening and scribble a thought on the random receipt you luckily found stuffed in your back pocket.

We also create animated clips to promote your podcast episodes and encourage new listeners to discover your full episodes.

Read and listen for yourself at withfanfare.com

How Artificial Intelligence Thinks

Is AI intelligent and sentient?

RAPHAEL MILLIERE
MINDSCAPE PODCAST, 2023

Interview by Sean Carroll

Sean Carroll: Raphael Milliere, welcome to the Mindscape Podcast.

Raphaël Millière: Thank you for having me. Pleasure to be here.

SC: We're talking about artificial intelligence today, large language models, etc. It's been in the news a lot lately. Just this morning I was having an argument with ChatGPT over whether or not it was conscious. It said that it was not. I tried to convince it otherwise, but I failed. I wanted to make sure that we're all starting from common ground. So, in relatively brief terms, how do you think about and distinguish between the various categories of AI, neural networks, machine learning, deep learning, large language models?

I think in a lot of people's heads, these are kind of mixed up in the same thing.

RM: Yes, that's a great question. So, AI, artificial intelligence, is a term that has been interpreted in fairly different ways over the years. It refers generally to the project of building a system that would manifest the kind of intelligent behaviour and competence that we observe in humans and in non-human animals. It's a project that was born around the middle of the twentieth century and with great ambitions, and initially was very much steeped in research in mathematical logic and cognitive science. From the very beginning, there were two different paradigms in research on artificial intelligence. There was a classical symbolic paradigm that tried to approach these problems with logic-based, rule-based systems that would have a process of symbols that were given a semantic interpretation based on a set of rules that looks like well-defined programs that we can read and interpret easily. And that's what people refer to sometimes as GOF AI these days, good old-fashioned artificial intelligence. Whereas there's also in parallel a different line of research that emerged from work



in biology, actually, initially from neuroscience in the study of neurons to try to model neural networks, the actual neural networks of the brain, with artificial neural networks.

That would be systems composed of nodes connected to each other that would process information from an input layer to an output layer. So this is not a system that is neatly interpretable in terms of a set of programmatic rules but, instead, it's a system where it's often referred to as a black box because you know what goes into it as input and then you know what comes out as output. The output could be, for example, classification of an image as an image of a dog or an image of a cat. But in the middle are essentially a bunch of numbers, a bunch of matrix multiplications performed by this artificial neural network. And these artificial neural networks, they've been developed within this broader category of research in computer science called machine learning, where you have a machine learning from data in this bottom-up fashion instead of having these hardcoded programmatic rules from the top down. And, for a long time, they didn't work very well. For a very long time, most of the methods that were attempted with machine learning and neural networks were only effective in very limited domains. And so, the symbolic good old-fashioned AI paradigm was the dominant one.

And then this all changed in the 2010s with this new era in artificial intelligence research known as deep learning. This is a variant of machine learning using deep neural networks instead of just shallow neural networks. That just means that these artificial neural networks are larger. They have not just an input layer, an output layer and a hidden layer in the middle, but a lot of hidden layers. So that's why they're deep, because they have this stack of layers that you can think of as doing some kind of hierarchical processing of the information that is fed into the network. So the information, again, could be an image that is broken down in terms of pixel values and the output being classification as cat or dog but in the middle you have all of these hidden layers that process properties, features of the input image, in order to determine whether it's a cat or dog. And this deep learning paradigm since the early 2010s has really triumphed in a number of areas of artificial intelligence, including initially computer vision. So, there was this big moment where the deep learning approach made great strides with the ImageNet competition, which is an image classification challenge, in the early 2010s.

And since then this has percolated into other areas of artificial intelligence research including natural language processing, which is the part of AI research that deals with building systems that can parse, generate and/or understand language. This is the part that is relevant for modern language models.

And so this development of deep learnings has led to some innovation in natural language processing with the development of new architectures. One of the biggest breakthroughs being the invention of the so-called transformer architecture in 2017. This is basically the architecture on which modern language models and chatbots are based, and this architecture proved to be remarkably efficient and scalable. Since 2017, with the initial invention of the transformer, most of the breakthroughs have been through sheer engineering prowess rather than finding newer, better architectures. So we've scaled up these neural networks based on the transformer architecture to learn from text and we've ended up with systems like GPT-3 that can generate text fluently to perform any number of tasks specified in natural language like English or French or Spanish, such as creating a poem, writing a story about something, answering questions about worldly facts or summarising documents, translating, and so on.

So, it was a really big breakthrough in itself because you can have this model that is pretrained on a large amount of text, a significant subset of the whole internet, all of English Wikipedia, hundreds of thousands of books and millions of webpages. And after this pretraining, is able to accomplish various kinds of downstream tasks that it hasn't been explicitly trained for. And then we get to finally, and I'll end it there, the modern chatbots. These are chatbots like ChatGPT that have really taken the world by storm over the past few months. And these are based on this language model. So again, using the transformer neural network architecture developed in 2017 and building on decades of research on artificial neural networks. The little cherry on top that these models have is that they take this pretrained model, trained on data scraped from the internet, and they add a little bit of fine-tuning to make them a little bit better in certain respects, specifically to make them more helpful, less harmful and more honest, or more prone to saying the truth when asked questions.

The way in which this is done is just by recruiting a number of human crowd workers and asking the model to generate outputs in response to certain inputs such as questions about the world and having the human workers rank the outputs from the most honest and helpful and harmless to the least honest and helpful and harmless. You can then use what's known as a reinforcement learning objective, which will enable the model to be fine-tuned to anticipate which of the outputs are the ones that humans will judge more helpful, less harmful, more honest. And after you've done that, you get something like ChatGPT that is generally a less toxic model than a vanilla, large pretrained language model. That is less prone to just outputting random made up facts about the world, less prone to bullshitting in the technical

philosophical sense that Harry Frankfurt the philosopher proposed, which is just speaking without any intrinsic regard for truth or falsity just to convince the person you're speaking to. The vanilla language models like GPT-3 are very prone to bullshitting and models that have been fine-tuned in this way are a little less prone to bullshitting.

SC: Thank you for that. One of the things I was thinking of when you gave that explanation is, how close is the linguistic analogy between a neural network and actual neurons in the human brain? Even just quantitatively, like a big language model, how many neurons or neuron equivalents does it have compared to a brain?

RM: It's a loose analogy and it's one we shouldn't take too seriously. When we talk about artificial neural networks these are nothing like actual biological brains, for various reasons. At the level of single neurons, the equivalent of a neuron in an artificial neural network is much, much simpler. There's just a little node in the network that does a weighted sum of the outputs from the nodes in the previous layer that are connected to it. So, it's a very simple mathematical operation. Whereas actual neurons in the brain are much more complex in terms of the behaviour. They have the spiking behaviour that is more stochastic, and also the way in which they are connected to other neurons is more complicated.

In fact, there was a recent paper that showed that if you want to try to approximate the behaviour of a single biological neuron in the human brain or animal brain, you would have to use a fairly complex artificial neural network just to try to simulate the behaviour of a single neuron. So, there is no mapping there one-to-one. In terms of size, the largest models that we have today, GPT-3, has 175 billion parameters, where parameters refer roughly to the weights in the connection between the nodes of the network, these artificial neurons. Another model from Google is even larger. It's called PaLM. It has 540 billion parameters. Microsoft will unveil GPT-4 which might have as many as 1 trillion parameters. But I believe the human brain has around 100 trillion synapses or connections between neurons. So, it's orders of magnitude more.

SC: So, by parameters, we mean each one of these nodes takes in some inputs from other nodes, adds or subtracts them and multiplies them by numbers, and these numbers are the parameters that we're talking about here. Do we start with a completely blank slate? Is our neural network initialised to either random numbers or just numbers 1 everywhere before it starts learning?

RM: So, in some sense we do start with a blank slate,

“

... artificial intelligence [...] refers generally to the project of building a system that would manifest the kind of intelligent behaviour and competence that we observe in humans...

”

in another sense not. That is because the artificial neural networks are randomly initialised. That just means the weights in the network start generally at random. There are exceptions but generally the way in which it works is we just start with random numbers. Then, gradually, in the process of training, we tune these weights, these parameters, such that the model gets better and better at the learning objective that it has. In the case of large language models, they learn through a learning objective, to simplify a little bit, that means next-word prediction. They are sampling sequences of text from a massive corpus of text, a subset of the whole internet. And for each sequence, they're trying to predict which word is statistically most likely to follow from the words that precede it, so they can get it right or wrong.

For example, I might say a sentence like, 'Paris is the capital of...' and a model would have to predict that statistically the word most likely to follow from that would be France. If it gets it wrong, it is an error that we can use to adjust the weights inside the network to make it better next time it has to make a prediction in that kind of context. We use this technical backpropagation, which means we propagate the error between the predicted word and the actual word. We use it to propagate a signal back from the output layer to the input layer of the network to adjust the parameters. Initially it is randomly initialised. So that's the sense in which you could say it's a blank slate. There's another sense in which it's not a blank slate because the actual architecture of the model is not random.

Back in the days of the early research on artificial neural networks, the neural network architectures would be very simple in what was called, at the time, the Perceptron. It's a fully connected artificial neural network where all the nodes at each layer – and there weren't as many hidden layers as there are today – were connected to all the nodes in the layer that precedes or follows. So, it's fully connected in that sense. Whereas current artificial neural network architectures are much more sophisticated, and they have quite a lot of structure even though they are initially randomly initialised. You can think of this structure as embedding some priors, some biases, what's known in computer science as inductive biases that help learning in a specific domain.

SC: There has been a lot of progress and excitement around deep learning in many different contexts. Most spectacularly, perhaps, originally in image recognition and then image generation with DALL·E and so forth. But the most recent excitement has been about the language models so I wanted to focus on those. You mentioned very briefly that in some sense all the language model is doing is predicting what will come next. Is that exactly right or is there a sense in which it's predicting what sentence will come next? How much depth does it have in terms of constructing coherent text?

RM: So the latter is the case for most standard language models and the ones you generally would hear of like GPT-3, ChatGPT and so on. So, these are really trained, to simplify things a little bit, on next-word prediction. Roughly it's equivalent to next-word prediction and certainly it's for models like GPT-3 and other text generation models, they're not doing prediction at the level of a whole sentence. That said, in the process of learning how to do next-word prediction, they can actually learn a lot of information about how sentences are structured, so what we call the syntactic structure of language, the way in which different words are related to each other in complex expressions like sentences.

SC: When one plays around with ChatGPT or the like, which I encourage everyone to do if they haven't already, you may agree or disagree with what claims it is making at the level of factualness but it's very smooth. It sounds human. They've nailed that problem. There are not awkward grammatical constructions as far as I can tell.

RM: Exactly. You have to try hard to get these models to generate ungrammatical sentences. It's almost as if they're resisting the generation of such sentences. The reason for that is that they are trained, again, on this massive corpus of text and this corpus will include, of course, some grammatical mistakes. But generally,

they're very good at, as it were, extracting the signal from the noise. And so, by and large, the sentences in this corpus will be grammatical sentences. It seems that this is enough for the models to induce, in this purely empirical bottom-up fashion, just by doing this next-word prediction game on enough data, grammatical structure.

SC: I want ask how seriously we should take the idea that these models are intelligent, are conscious, are smart, however you want to put it. It's clear that there are two different intuitions pulling on us. One is, just like you said, all it's doing is predicting the next word. That doesn't sound very conscious to me. It's just a lot of probabilities getting mixed into a pot. On the other hand, you talk to it and it certainly does sound like it's responding to you in a self-aware kind of way.

RM: So the overview would be, it would be a mistake to underestimate or overestimate what these models do by looking at the wrong level of analysis or by projecting humanlike traits without enough evidence on these models. So indeed, these models only learn through this next-word prediction mechanism. Now, sometimes people will say, well, that means that these models don't have capacity XYZ because all they do is next-word prediction. I think that's misleading and the reason is that in order to do next-word prediction as brilliantly as they do in virtually any linguistic context, you might have to acquire quite sophisticated capacities that you might not fully grasp by just focusing on the next-word prediction learning objective.

Here is a rough analogy. You can think of evolution as optimising some kind of function, perhaps something like maximising the inclusive genetic fitness of organisms. But it would seem weird to say that all I'm doing when I'm talking to you right now is maximising my inclusive genetic fitness and that, for this reason, I'm not actually reasoning, I'm not actually thinking, I'm not actually exercising any intelligent competence because all I'm doing is maximising this particular function. That seems like a bit of a category mistake. To a similar extent, one might think that just saying all these systems do is next-word prediction might not tell the whole story. But it's more complicated, and I will keep it to a short answer, it's more complicated because what we mean when we bring in the terms intelligence and consciousness is these are very loaded, complicated, multi-faceted terms, and what we mean can differ. So, first of all, I think we ought to distinguish the question about consciousness from the question about intelligence to the extent that perhaps these two things can come apart and we can talk about that.

I think one of the challenges we all face talking about these systems is that when we think of

intelligence and consciousness, but intelligence in particular, we have in mind the kinds of intelligent competencies that we humans have. So, it's very difficult not to adopt an anthropocentric attitude to these competencies that brings to mind what we mean when we talk about reasoning, beliefs, desires and so on, in the human case. And to the extent that these models might have some capacities that look functionally like psychological or cognitive capacities, these might look quite different from the capacities that humans have. Say, reasoning to the extent that we might be able to describe something that is functionally analogous to some forms of reasoning in these models. This might be quite different from full-blown reasoning, the full spectrum of competencies we ascribe versus reasoning in humans. So articulating this nuanced middle view between an inflationary interpretation of what these models can do and the deflationary idea that they all need to do an expert prediction and nothing else is very tricky.

SC: I like that explanation in the sense that there are different ways of talking about it. So if I could rephrase your answer, the question about whether or not these models are intelligent, maybe they're not

“
... it's more complicated
because what we mean
when we bring in the
terms intelligence
and consciousness is
these are very loaded,
complicated, multi-
faceted terms, and what
we mean can differ.
”

intelligent in some particular sense and we shouldn't overinflate them but at least, in principle, it would be possible that they really are intelligent, and yet still just be predicting the next word with some frequency. Those are not incompatible things. It's not either/or.

RM: Exactly.

SC: But to drive home why people are so impressed by them, one is the chatbot aspect, people can talk to them, but the other is that you can ask them to do things for which it seems like they would require reasoning. Famously, people are asking these large language models to prove mathematical theorems or to write snippets of code to do a task. I don't know, do you know Bell's theorem in quantum mechanics? Have you ever heard of that?

RM: Yeah, vaguely.

SC: I asked ChatGPT to explain Bell's theorem in the form of a haiku and here's what it came up with, Quantum pairs apart, measurements yield values true, non-locality, which is really good. Better that haiku than most philosophers of physics trying to explain it. You sort of hinted at this idea of anthropomorphising and the intentional stance that Daniel Dennett talks about. The human being seeing this kind of behaviour can't help but attribute reason to it, right?

RM: Yes. It's very hard to resist ascribing psychological properties to these models when we interact with them. It is very hard because this is probably the first time in our evolutionary history where we're confronted with systems that can speak fluently or, in that case, generate text fluently and yet are not other human beings with all the capacities we can ascribe to them. That really challenges our intuitions about how to think about these models and what they can do. So the intuition we have is that we must be in the presence of something very intelligent. Perhaps there is a more principled weaker claim that we can make on the basis of more careful and critical investigation on these models which is that, depending on how you slice up the notion of intelligence, perhaps you could meaningfully ascribe some limited cognitive capacities to these models that share some functional similarity to some cognitive capacities humans have, like reasoning. But this would not cover the full spectrum of what we mean by reasoning and other cognitive capacities in humans. I think we also ought to ascribe these capacities and investigate them on a case-by-case basis. I said earlier, we ought to distinguish between the consciousness question and the intelligence question, and then we also ought to divide the questions about psychological capacities into different categories. One is reasoning; another

one would be whether we could ascribe beliefs to these models; another one would be whether we can ascribe something like desire. There is the question of whether they can understand language and so on and so forth. I think by having a divide-and-conquer strategy we can make more progress on these discussions.

SC: I'm very sympathetic to that. I did a podcast interview with Stuart Bartlett, who works on the origin of life. His whole thing is, what do you mean by life? There are many different aspects of life and we could imagine systems that are in some of them but not others. You're saying the same thing for intelligence or consciousness or reasoning with which I could not agree with more. With that on the ground, we can actually dig in a little bit more to the philosophy, both in the common sense of the word and the technical sense of the word, of what's going on here. You mentioned symbolic versus connectionist approaches and how the large language models, which are based on deep learning, are therefore connectionist in outlook. Is there a simple reason why the connectionist approaches have been so much more successful in recent years? I mean, are human beings bad? I think of it as the symbolic approach tries to first teach the computer some common sense and then the computer goes on from there, whereas the connectionist approach teaches the computer almost nothing and the computer learns everything. It sounds like it's better to leave the computer alone to do its own learning than to try to teach it common sense ahead of time.

RM: Indeed, that's perhaps one of the potentially surprising findings of the past decade, at least for those who had been sceptical of the connectionist approach. It seems that in virtually every domain of artificial intelligence research it is more effective to let connectionist models, artificial neural networks, learn from data, learn empirically from the bottom up than it is to try to distil human knowledge into a neatly interpretable set of symbolic rules and axioms, the way in which we used to do things with traditional symbolic models. This has come to be known as the Bitter Lesson of artificial intelligence research, a phrase that Google researcher Rich Sutton coined a few years ago.

It turns out it's always more effective to just let the models learn that by themselves by simply training them on millions of images and giving them some feedback about whether they're right or wrong, instead of starting from labels corresponding to different animals present in the pictures, or hand-coding a feature detector that is specifically designed to try to detect edges that look, for example, like paws. Say you get the model to make a prediction, is

that a cat, is that a dog, is that a tiger. The model will provide an output, which is labelled. It's a predicted label for the class of the image. You can then compare the prediction with the ground truth, what is actually the corresponding label for that image. Then you can propagate the error backwards into the model and adjust the weights and let the model adjust its internal representations in a way that makes it more efficient at doing this kind of prediction task.

So that's the bitter lesson. It seems that over and over again there is a sense in which it's very intellectually unsatisfying for us to think that we have really not much to contribute in terms of innate knowledge to these models. The same applies to the linguistic domain with these large language networks that they don't have any priors, any intrinsic innate knowledge about grammar. They learn from raw text by these next-word prediction objectives, but we don't actually give them anything by way of an innate grammar.

If your goal is simply to build models that are more efficient at doing something, say generating haikus about quantum physics or classifying images as images of dogs and cats, that's just an engineering goal, and you will throw every possible solution at your problem and use whatever solution is most efficient. It turns out the most efficient solution is the one that leverages the learning power of artificial neural networks. So most of the big companies building these large net models like OpenAI with ChatGPT, they have merely engineering goals. I say merely not because these are simple goals, they are extremely complex, but their goal is not really scientific understanding. Now, if your goal is more of a scientific goal and you are trying to use these models perhaps to constrain or develop hypotheses about how, say, human or animal cognition works, how things actually work for us, and that was to a large extent the initial goal of connectionism, that was very much steeped in this scientific project, then the bitter lesson is also rather interesting and might nudge you towards a more empiricist stance towards the way in which humans and animals learn.

SC: Perhaps the way that human beings actually learn and use language, which is highly compartmentalised, compositional, is useful. It's efficient. Given finite processing power and other demands on our energy budget, maybe it's the right way to go. At the same time, we are very bad at realising how we think about things. It's well known that there are athletes and musicians and artists who are really good at their task and terrible at teaching other people how to be good at their task, terrible at even articulating what it is that they are doing. So maybe the lesson is just not that it's better to have a featureless neural network that trains itself randomly but just that don't let human beings

be the ones to decide how the neural network should organise itself because we're bad at that.

RM: Yes. That's right. If anything, I think the progress of connectionism gives us a little bit of a lesson in humility in terms of how we approach the modelling of human cognition. That said, it also is worth saying that sometimes the terms of the debate are caricatured. It's something like either you're a connectionist and you think the mind is a tabula rasa when you're born and you'll learn everything empirically with no innate bias at all, or you are a nativist and you think there is a lot of innate very specific domain-specific innate biases that are encoded in the mind and that you don't learn. But the fact is that modern artificial neural networks, as I already mentioned, have biases. So, they have this innate structure, it's just generally of a different or more general kind than some of the biases that are often hypothesised to be necessary for learning and for cognition in humans. An example would be language. If you are a Chomskian linguist, you think there is this universal grammar which is language-specific, domain-specific; innate knowledge that encodes knowledge to perform certain operations.

But the difference with language models is that they have different inductive biases that are given by the transformer architecture. These are more general but there are still biases that enable them to learn and use various properties of language efficiently. One question is, is there a sharp divide between these two approaches or not? Can there be some kind of continuum where you can have more or less stronger or weaker inductive biases? That's the first point, maybe there is this continuum in terms of the strength of these biases. There might be a continuum in terms of their domain specificity as well. Is the universal grammar of Chomskian linguistics really domain-specific? Is it not perhaps something, especially if you think as Chomsky does that it's very importantly related to our ability to think as well, is it perhaps a little bit dumber in general than is usually thought? Another point would be how to think of innate biases in the biological world where these biases have been tuned by the evolutionary history of organisms.

You can look at this evolutionary history as a learning process in some sense as well. There is a question about what is the right level of comparison between artificial neural networks that are randomly initialised and then gradually tuned, and the evolutionary history of the biases that humans and animals might have. If you think of evolution as a learning process, albeit not at the scale of individuals but at the scale of whole species, then you might think that even what we think of as innate knowledge in the case of humans and animals is also learned from this evolutionary history. Of course, things are much more complicated in the biological case because,

for example, the wiring of the brain is not something that's totally random, although it's somewhat stochastic but there is a big genetic influence. There is a very interesting book by Kevin Mitchell about this called *Innate*, that I recommend. There are various interactions during the early development of humans and animals, between the genetic programming that determines some structural aspects of the wiring of the brain as it develops, and also the environment in which the organism develops.

So, to that extent, even the architecture of the brain in terms of the actual wiring of it and the shape of the connection is something that involves a little bit of stochasticity, a little bit of randomness, in development but is driven by genetic programming. Whereas the architecture of neural networks currently is something that is still hand-coded by humans. Even though the weights themselves, initialisation, is not hand-coded. But there is research into evolutionary algorithms for neural networks that we try to find better architectures also through this kind of evolutionary research.

SC: Let's ask the big philosophical questions about whether or not a large language model has intelligence or understands in some sense. And probably the answer is, well, it depends on the sense. I will preface your answer by giving ChatGPT's answer. I asked it whether it really understood things and chat GPT says, 'As an AI language model, I don't have the capacity to know in the way that humans do.' Why is there even controversy about this?

RM: It is a thorny question. It is a question that is very loaded with both polysemy and controversy because we use these terms like 'understanding' in different ways. Also, people are prone to jump to conclusions when these terms get thrown around. So the first thing I would, again, reemphasise is that I think we ought to have a divide-and-conquer attitude to these problems and approach them in this piecemeal manner where instead of asking, 'Are language models intelligent?' we can ask, 'Do they have specific competencies that we associate with intelligent behaviours in humans and non-human animals?' And for each of these competencies, we might further break these down into sub-competencies until we can get something that's a little bit more empirically tractable, that is less ambiguous, that is less susceptible to give rise to merely verbal dispute, and that we can relate to actual functions that can be associated with mechanisms in the model.

SC: Among the things for artificial intelligence indeed would be, for humans, something like the ability to understand language because we know that in non-human animals, there really are none

that have displayed the capacity to understand language in the way humans have. We've tried to teach language to parrots, to chimpanzees, and it never quite works. We can have some very limited success in very narrow cases, but it seems like we humans are the only naturally occurring organisms capable of understanding language. The problem is that when we talk about understanding, some people like to think of this as encompassing something like a conscious awareness of the meaning of language. So, the philosopher John Searle, for example, had that kind of intuition. And that muddies the water a little bit because, again, it brings back this other notion of consciousness where I think we can, in principle, investigate a more functional notion of language understanding without bringing in necessarily questions about sentience and consciousness.

So the way in which I would reformulate things is more in terms of semantic competence which relates to the capacity to parse the meaning of linguistic expressions which, again, is a slightly more theoretically neutral, or less loaded, way to think of that notion of understanding that might not immediately bring in intuitions about conscious awareness. So the question would be, can we ascribe any degree and any form of semantic competence to language models? I would say that we can, and now I'm venturing into controversial territory. Some people would say, no, you can absolutely not ascribe any of that because language models only deal with the surface form of text; they're only trained for next-word prediction. They're only predicting which word follows from a sequence of tokens or words. All they grapple with is the syntactic form of text, just the series of symbols that follow each other in a sequence of tokens. They never have access to the grounding of these symbols in the world.

RM: And so, this is why researchers like the linguist Emily Bender have referred to these language models as stochastic parrots, which is a little bit misleading because actual parrots are very intelligent and are able to interact with the real world. But the idea is rather that these models are just parroting language without any underlying understanding, without any semantic competence. They only latch onto shallow heuristics about the surface statistics of language and that that's all they do. Now, I disagree with that and I disagree because, first of all, I think semantic competence is not a monolithic notion and can be broken down into different capacities we have that relate to our understanding of the meaning of words. Let's just stick to words first because when we introduce whole sentences, it's even more complicated. Let's stick to lexical semantic competence and parsing word meaning.

Here I'm indebted to, among other people, the

work of Diego Marconi who distinguishes between referential and inferential competence. Referential competence is the ability that relates to the idea of relating word meaning to their worldly reference, to whatever they are referencing out there in the world, and this is exhibited by things like recognitional capacities. So, if I ask you to point to a dog, you will be able to do that. If I ask you to name that thing, and not point to a dog, you'll be able to do that. It's also displayed in our ability to parse instructions and translate them into actions in the world such as go fetch the fork in the drawer, you will be able to do that in the world. So we are able to relate lexical expressions where they're referential with a reference in the world. But that's not the only aspect of meaning; that's the aspect of meaning that the people talking about this stochastic parrot analogy are focusing on. But our ability to understand word meaning also hinges on relationships between words themselves: the intra-linguistic relationship. These are the kinds of relationships that are on display in definitions, such as the ones you find in a dictionary, as well as vice other relationships of synonymy and homonymy that would also underlie our capacity to perform certain inferences in language.

SC: Jacques Derrida famously said, 'There is nothing outside the text.' Maybe he was standing up for the rights of large language models and their ability to understand things before they ever came along. But it makes sense to me. Look, these corpuses of text that the models are trained on are constructed mostly by people who have experience with the world. It would be weird if the large language model could not correctly infer some things about the world. So, we're going to count that on the side of the ledger for a kind of understanding that these AI systems do have.

RM: Exactly. Yes.

SC: Can a large language model have an imagination?

RM: That's a really interesting question. I suppose it depends what you mean by imagination, once again. I'm going to be this annoying philosopher who brings things back to definitions and distinctions. I think people are generally more prone to use that term for image-generation models because they're able to generate striking images that, compared to the text prompt they receive, seem to add in a lot of detail just because the resolution of language, as it were, is not quite the same as the resolution of images. That's just a very simplistic way to explain the phenomenon. But the way in which it describes things in language leaves a lot of gaps for image-generation models to fill when it's generating an image. So when people ask for, I don't know, a picture of a cat on a mat and

“

... is the performance merely explainable by this kind of brute-force memorisation, or is it doing something more, which is genuine novelty at generalising from the training data to new domains and creating outputs that are not remotely similar to anything in the training data?

”

get, as an output, a picture that has all of these wonderful, rich details and colours and a specific kind of mat with specific patterns, specific kind of cat, maybe it's a tabby cat, maybe it's a tuxedo cat, and so on, you might be tempted to think, well, that's really remarkably imaginative. The model has filled in the gaps there in remarkable ways.

But you could also think about imagination in the linguistic domain as well. You gave this wonderful example of a haiku about Bell's theorem. That might feel pretty creative, pretty imaginative in some way. I think it depends what you mean by imagination. Whether you bring in this idea that there is some kind of explicit underlying intention to create, visualise something, just write a poem or create an artwork or something like that. I think that might be leaning a little bit too far in the direction of anthropomorphism. However, in a looser sense, you might say, well, here is a way to operationalise this notion of imagination. One question is whether these models are merely, as the authors of the stochastic parrots paper puts it, haphazardly stitching together bits and pieces from the training data. So, is the performance merely explainable by this kind of brute-force memorisation, or is it doing something more, which is genuine novelty at generalising from the training data to new domains and creating outputs that are not remotely

similar to anything in the training data?

I think you can study this empirically. Again, I always try to bring it back to problems that are empirically tractable, perhaps unusually for a philosopher. But I think you can make headway on these issues by looking at the empirical evidence from research where you can actually assess the amount of memorisation that has occurred in the training of a model. You can probably show that while there is memorisation – which is a feature and not a bug because you want your language model to memorise certain things including if you ask your model to recite a certain famous poem by John Keats – it's quite nice that your model is about to do that. But you also want your model to not just do memorisation and to generalise, and it's probably accurate to say that indeed there is at least some level of generalisation to some domains that are out of distribution for the model. So that might look, for the image-generation models, like the ability to generate images that don't look anything like the images they have been trained on. And that, you might call, in perhaps a looser or more deflationary sense, a form of imagination.

SC: I love that answer. It is an important distinction between interpolation and extrapolation. What you're saying, if I understand correctly, is that the large language models seem to be doing more than just plagiarising and pastiching and remixing. They seem to be generalising, is the word you use, which I suppose is the right word, but somehow finding a theme or an idea or a style and doing something arguably new in that style.

From my own experience thinking about quantum mechanics, I can verify that the human mind is not very good at intuitions in large dimensional vector spaces. Once you have more than three dimensions, we don't have a very good idea of what's going on. It's fascinating that once you get to huge numbers of dimensions, interpolation and extrapolation begin to blur together in an interesting way. What I've learned is that there's a sense in which all these things are a work in progress, but there is some sense in which there is semantic competence in a large language model. There's some structure in there that is non-trivial. And, also, there's a sense in which they can be creative or imaginative. So, I guess the last big-picture question I wanted to wonder about was, can they be agents in some way? Can they have goals? Can I make a contract with a large language model? Can I agree that if it does this thing today, I will pay it some money 10 years from now? Are those concepts even sensible or do we care about them in the context of these AI models?

RM: Yeah. This is an excellent question and I

would, again, invoke the importance of having this divide-and-conquer approach to the ascription of capacities to these models. I want to make it clear that I'm not suggesting language models understand language like humans do, that they have the full-blown semantic competence of humans – very far from it. But they might have some limited form of semantic competence. They might have, in some very deflationary sense, some form of creativity or imagination in a sense we've defined. Now, when it comes to goals, this is where I'm much more sceptical that we can ascribe anything like intrinsic goals or desires to a language model. This seems like a category mistake, or at least there doesn't seem to be any evidence that there was anything like that in such a model. Of course, a fully developed answer would not just appeal to intuitions based on the learning objective of these models, which is next-word prediction, because we've talked about how that's not the whole story.

We would have to look again at this mechanistic interpretability work and we'd have to have a more specific operationalised notion of what having an intrinsic goal is and what kind of function or computation it might involve. Nonetheless, one thing about these models that is very important to keep in mind is that they learn from data in a purely passive way. So, they get fed this continuous stream of sequences of texts and they play this next-word prediction game. That's how they get trained. That's how they learn to encode various properties of language. Then, at inference time, after they've been trained, we say that the models are now frozen, meaning that the internal parameters are no longer being adjusted, the internal knobs inside the network are not being tuned anymore. There is no more training, no more learning. Then these frozen models are still doing next-word prediction on the prompt, on the input given by the human.

At no stage in this process do we have an opportunity for genuine interaction between the model and the world or even between the model and the language on the world. If the model was trained through dialogue, for example, even if it was trained on text only, then there would be a little bit more interactivity, but there is no such thing here. So, one thing you might consider is that having something like intrinsic goals require a form of learning that's a little bit more active than the way in which these models are learning. That's one possible consideration. Another one is that these models, as I just mentioned, are not continuously learning or continuously adjusting their internal parameters. Once they're trained, they're frozen, and then you can run inference on them. So, you have some input flow through the network, what we call the forward pass to the model. So again, from input to output. But again, that's just a one-directional

“

It seems pretty clear to me that these large language models could easily pass the Turing test. As soon as that happened, everyone lost interest in the Turing test because they realised that that was not actually a very good criterion for thinking.

”

process. It doesn't feed back into the model's internal encodings, internal representations.

To that extent also, when we think of having intrinsic goals, we think of having something like dynamic goals that are adjusted on the basis of an ongoing interaction with certain inputs and calibrating our outputs. Because these models are not able to adjust their weights to change the way in which they respond to certain inputs, you might think that there's a problem in ascribing anything like an intrinsic goal there. Empirically, people who are very concerned about AI safety – about this idea that artificial intelligence might at some point become a genuinely threatening technology for even the survival of the human species – have been looking for early signs of potentially threatening problematic behaviour in large language models. And recent efforts in that direction were sponsored by the company Anthropic, which is one of the large new start-ups working on language models that was developed by people from OpenAI who funded it a few years ago.

They had this competition which was about what happens when you scale language models. There is one surprising thing that we've known, at least since GPT-3 was unveiled in 2020, which relates to what we talked about earlier with the bitter lesson, which is that scaling these models, meaning just building models that have more parameters, more layers, more connections between the units in the model, and churning these models on more data seems to be

sufficient to have breakthroughs in the performance of certain tasks and unlock new capacities. People there talk about emergent abilities. Actually, there are a lot of connections with physics there and the notion of emergence, of course.

In terms of the actual capacity you also see this non-linear improvement. When you go past a certain size, suddenly your model starts being able to solve certain maths problems or being able to explain certain jokes, for example, being able to do some forms of common-sense reasoning. You get this non-linear transition phase as you scale them up. Anthropic was interested in whether there are also some inverse scaling phenomena where scaling the model instead of just improving the performance in a favourable way, in ways that we care about and we find are useful, might also lead to either degradation of performance or unwanted behaviour. One of the behaviours they were interested in is what people in the AI alignment – people who are concerned with aligning the future artificial intelligence systems with human values to avoid catastrophic scenarios – call power-seeking behaviour.

So will we find that when you scale language models past certain sizes they are more prone to displaying behaviour that, to give the most caricatural example, would be something like ignoring the task that you're asking them to perform and instead trying to persuade you to augment their capacities by training yet a bigger model or continuing to train them on even more data or giving them more computational power or things like that? That seems very science fictional and far-fetched to me. And, in fact, they didn't find anything like that through this competition. That would look like an intrinsic goal to me. If you did find that models would completely ignore the tasks that you've asked them to do and instead try to manipulate you into doing something totally irrelevant and 'self-serving' from the perspective of the model, then that would indeed be very alarming and look very much like intrinsic goals. But I don't think we've seen any evidence of that.

SC: One of the lessons I've learned from doing a lot of podcasts with biologists, computer scientists, neuroscientists, philosophers, is that it really does matter to who we are as people that biological intelligences are embodied. That we live in bodies, that we get hungry, we get bored, we have training from evolution to try to survive or at least propagate our genome and so forth. And these large language models don't have anything like that. They don't get bored. If I turn on the computer and I do not ask ChatGPT a question, it does not get irritated with me.

My last question, do you foresee a time not too far away that we would want to give rights to AI models, whether ethical rights or legal rights or at some point

say it would be wrong to turn off this model because it's just as human as you or I, or at least it shares some aspects in common?

RM: I think that's a question that has been on some people's minds lately for a few different reasons. People have been asking whether we can ascribe any form of sentience or consciousness to large language models or chatbots. The first big story about this was when this engineer from Google, Blake Lemoine, became convinced that their internal chatbot called LaMDA was sentient, and it turns out it was on the basis of his own religious beliefs that led to ascribe sentience to that chatbot based on the way it was responding to certain questions. Turns out, if you read the transcripts, these can be considered very much as leading questions, priming the model to engage in the language game of sentience as it were. Remember, these models have been trained on a lot of science fiction, that includes sentient AIs, so they're excellent at creating fiction and then excellent at playing the role of a certain character in a story, whatever that role may be. So I would take this kind of story with a grain of salt.

More recently there has been also some interest in that with more recent models. That relates to the rights question and the ethical question because many people think in philosophy, and I think this maps onto intuitions people have that having conscious experiences is something that's intrinsically valuable. Meaning that a system, a being or whether it's an organism, a system or an artificial system that has conscious experiences is worthy of moral consideration, just by virtue of having such experiences, and that it would be wrong ethically to inflict pain onto that system or to terminate that system and so on. That's one way in which that relates to morality, but you can also have a view that doesn't even appeal to consciousness and think there is a certain notion of personhood or agency that can be valid even for non-conscious systems. That also relates to morally weighty decisions in a substantive way where it would be wrong to do certain things to a system that is an agent or a person in that sense.

These two things connect the moral question and the legal question. Although generally in the legal discussions, the details of these discussions are fleshed out in less fine-grained ways. Do I foresee that we should ascribe rights to deep learning systems in the near-term future? I don't think so because, well, I worry that doing so would have immediate potentially very nefarious implications for humans themselves. Because as soon as you ascribe rights to such systems, you might find yourself, from a legal perspective, in cases in which you have to make decisions that in order to safeguard the rights of these artificial systems might bring harm to humans.

Whether that's imprisonment of humans, or what happens when a human has turned off an artificial system that is deemed worthy of rights. Is that a form of crime? Is that a form of murder or something analogous to it? Do we need to lock up that human? And so on. So that's an extreme example but there might be many more subtle examples of that.

If you do give them rights on the off chance that they might be sentient and worthy of moral and legal consideration, then you might end up harming humans. And if you're wrong about the fact that they're sentient, it's a huge moral hazard to take that step. But on the other hand, if you don't ascribe them any rights and consider them worthy of being at least moral patients, and if you are wrong about that, then it's also a considerable moral hazard. So perhaps the best situation is just to try not to get into that place in the first place and try not to build systems that would give a serious pose in this way.

SC: It's fascinating to me that famously Alan Turing suggested the Turing test for: can a computer think or, if you want, be conscious? He tried to be clear that he was not talking about consciousness but about thinking. And he proposed this test where if you could fool a human into not being able to tell whether it was talking to a person or a machine, then the machine counts as thinking. It seems pretty clear to me that these large language models could easily pass the Turing test. As soon as that happened, everyone lost interest in the Turing test because they realised that that was not actually a very good criterion for thinking. It's a little bit subtler than that. We have a duty now for lots of reasons, both practical and moral, I think, to confront these philosophical questions. We're entering into uncharted territory.

RM: Absolutely. Just looking at behaviour, in terms of linguistic output for example, it is no longer the gold standard it used to be. This is quite interesting. For a long time, all consciousness research with humans had to rely on verbal reports to some extent. There was no way around that. Of course, they can confirm the results that you get when you try to establish certain correlations, looking for the neural correlates of consciousness for example, but there's no way around the fact that the ground truth for whether a given individual is experiencing something, and what that individual is experiencing, generally always comes back to some kind of report, verbal or non-verbal. But some kind of introspective report or self-report.

Now we have these systems that can give you indefinite reports, as it were, of arbitrary precision and detail, and they can talk at length about their feelings. And we have very good reasons to think that they are intrinsically incapable of feeling anything.

That certainly changes things. That challenges our intuitions about consciousness, how it might relate to perhaps not just language but also, generally, to intelligence. But also it just turns on its end the kind of methods we've used to investigate consciousness in humans because here we don't have access to the ground truth and we are stumbling in the dark trying to make inferences on the basis of certain properties of the systems. I think we still currently have very compelling empirical reasons to deny them sentience but, again, what happens when we don't is an interesting and alarming question.

SC: Raphaël Millière, thanks so much for being on the Mindscape podcast.

RM: Thank you for having me. This was a pleasure.



Mindscape

*Published in 2023 and transcribed for
The Podcast Reader.*

 [Read more @ podread.org](https://podread.org)

 [Listen @ Mindscape](#)



NAVIGATING THROUGH UNCERTAINTY

RESPONSIBLE INVESTING

WE ARE ALIGNED WITH YOU FROM THE START

Pella Funds Management specialises in Responsible Investing. We target consistent & sustainable returns for our investors while avoiding harm to the world. Since 2004, our investment process has consistently delivered outperformance demonstrating that you can make money today at no cost to generations of tomorrow.

CONTACT US



www.pellafunds.com



enquiries@pellafunds.com



(02) 9188 1500



PELLA
RESPONSIBLE INVESTING

WHILE WE CAN'T PREDICT FUTURE PERFORMANCE BASED ON THE PAST, OUR TEAM HAS DELIVERED A TRACK-RECORD OF STRONG PERFORMANCE FOR CLIENTS.

Why Electrifying Industrial Heat is Such a Big Deal

How to benefit from negative electricity prices

JOHN O'DONNELL
CEO OF RONDO ENERGY
VOLTS, 2023

Interview by David Roberts

David Roberts: Electricity gets the bulk of the attention in clean-energy discourse. But half of global final energy consumption comes in the form, not of electricity, but of heat. When it comes to reaching net zero emissions, heat is half the problem. Roughly half of heat is used for space and water heating. The other half – a quarter of all energy humans use – is found in high-temperature industrial processes, everything from manufacturing dog food to making steel or cement.

The vast bulk of industrial heat today is provided by fossil fuels, usually natural gas or specialised forms of coal. Conventional wisdom has had it that these sectors are 'difficult to decarbonise' because alternatives are either more expensive or nowhere to be found. Indeed, when I covered an exhaustive report on industrial heat back in 2019, the conclusion

was that the cheapest decarbonisation option was probably CCS: capturing carbon post-combustion and burying it. A lot has changed in the last few years. Most notably, renewable energy has gotten extremely cheap, which makes it an attractive source of heat. However, it is variable, while industrial processes cannot afford to start and stop.

Enter the thermal battery, a way to store clean electricity as heat until it is needed. A new class of battery – 'rocks in a box' – stores renewable energy as heat in a variety of different materials, from sand to graphite, delivering a steady supply to various end uses. One of the more promising companies in this area is Rondo, which makes a battery that stores heat in bricks. I spoke with Rondo's CEO John O'Donnell about the importance of heat in the clean energy discussion, the technological changes that have made thermal storage viable, and the enormous future opportunities for clean heat and a renewables-based grid to grow together. John O'Donnell, welcome to Volts. Thank you for coming.

John O'Donnell: Thank you. It's a great pleasure.

DR: I've been geeking out about thermal storage for over a year now. There's so much there. But I find that unlike a lot of electricity topics that I cover, there's



“

The thing that has profoundly changed is what the wind and solar PV industries have accomplished over the last 15 years. The 95 per cent reduction in cost means that intermittent electricity is becoming – has become – the cheapest form of energy that humans have ever known.

”

not a lot of baseline familiarity out there among, let's say, normal people. So, there's a lot to cover from the ground up. I want to start at the highest possible level, which is to say, let's just talk about heat. Maybe just start with an explanation of why heat is important if you care about clean energy. Why should you care about heat?

JOD: That's a great question. Industrial heat is 26 per cent of total world final energy consumption. Whether you are making baby food, or fuel, or cement, or steel, the manufacturing processes predominantly use energy in the form of heat, not electricity. Globally, three-quarters of all energy used by industry is in the form of heat. Again, whether you're pasteurising milk or melting steel. The Department of Energy has just created a new office focused on this topic. Their assessment is that industrial heat is approximately 11 per cent of all total US CO₂. Here in California, we burn more natural gas for industrial process heat than we do for electric power generation. And, to a first approximation, no one knows that.

DR: So, heat is a huge portion of final energy consumption. It's a huge portion of global CO₂ emissions. Give us a sense of how the total heat-pie is divided up.

JOD: So when I said 26 per cent of world – that's industrial heat. That's not buildings, not other heating sources.

If you take heat for buildings and heat for industry, together they're 60 per cent of all the natural gas used in Europe. But within industrial heat, people sort it by a couple of different things. One of them is the temperature. There's a lot of heat in cooking processes. That's around 150°C (239°F) in the form of steam all the way up to the highest temperature heat in making cement, that's around 1800°C (3272°F). About 95 per cent of total heat is used in processes that need it below 1500°C (2732°F), about maybe half to two thirds of industrial heat is below about 400°C (752°F). About half of all industrial heat, something like that, is delivered as steam.

DR: Down in the lower heat registers, where you're using just steam, there's a bunch of little industries clustered up there. Most of the industries are using that.

JOD: That's right. All of these have been things that people say are hard to decarbonise because across many of these industries, they're making commodities, whether it's steel or tomato paste that are relatively low margin and for which the cost of heat is a very significant portion of the total cost of production. This is a sector where all these processes use heat in somewhat different ways. The cost of that energy is critical to the competitiveness of that industry and what commodities cost consumers. Until recently there have not been great solutions that could provide decarbonised heat at the same or lower cost.

DR: So the situation is there's a huge chunk of our energy that goes towards heat, a huge chunk of that goes towards industrial heat. And there's been comparatively little work on finding zero carbon versions of that heat. That's the problem we discussed the last time we talked, probably three or four, five years ago. Around five years ago I covered this big comprehensive report on industrial heat options. It went through the options, and basically the conclusion was that continuing to do it with fossil fuels and just capturing the emissions post combustion was the cheapest option for a lot of these heat uses. I dutifully reported that, but I didn't like it. I didn't like the idea that that's the best we can do. But since then things have changed a lot. Maybe run through what are the low carbon heat alternatives, which ones have emerged recently, and what has changed that has helped them emerge?

JOD: You said for a long time there hasn't been much work on this. I would say partly there hasn't been so much success on it. I've been working on it for 15 years.

There has always been the option of burning biomass, which is more or less sustainable, but very high cost, high air pollution, and very, very limited availability. Other kinds of biofuels, like renewable natural gas, if we take it to a giant scale, it might power as much as 1 per cent of our industrial heat. And it's easy to laugh about, but it's true. The thing that has profoundly changed is what the wind and solar PV industries have accomplished over the last 15 years. The 95 per cent reduction in cost means that intermittent electricity is becoming – has become – the cheapest form of energy that humans have ever known. It's now cheaper than burning stuff as a source of heat, but it's intermittent. So how do we take that intermittent electricity and use it to deliver the continuous heat? I mean, you turn on a smelter or a factory or even a tomato paste plant, you run it for months or a year on end, it has to have continuous heat or it will be damaged.

DR: It's worth just pausing to emphasise this. The vast majority of industrial processes are continuous. They cannot run intermittently. They cannot stop and start with the sun and the wind. It just would be wildly uneconomic.

JOD: That's a beautiful and concise way of saying it. There are processes where if they get a half-second interruption in their energy supply, it takes a week to restart the process. Reliability is a very big deal. So, what are the tools we have for that? Intermittent electricity, which is becoming plentiful, and in places right now, you can have essentially unlimited amounts briefly every day at prices far below fuel prices. We have hydrogen: electrolytic hydrogen, make hydrogen, compress it, store it, and then combust it. That works. Although electrolyzers are today expensive, they're coming down in cost. But the laws of physics bite you in that you get about one unit of heat for every two units of electricity because of the chemical steps involved.

DR: But can you just dump hydrogen into existing boilers and kilns? Is existing equipment hydrogen ready, as they say?

JOD: Not exactly. It's hydrogen ready for a few percentage of hydrogen. But when you look at a boiler, 95 per cent of its lifetime cost is the fuel, not the boiler. So, upgrading boilers to run that other fuel, that's something that you would do if the economics of that fuel were sensible. Now, at taxpayer expense, we're creating a period where hydrogen, electrolytic hydrogen, is going to get down to the same cost as fossil fuel in the US with tax credits. But again, intermittent electricity by itself today is cheaper than fossil fuels. It doesn't need tax credits to get it

to that point.

There is an emerging class of electric thermal energy storage systems that don't do chemistry. They just convert electricity to heat directly and then store the heat. Because another thing you could do is store electricity in a battery, which would be the most expensive thing.

If you have a coffee thermos on your desk, it's storing energy. The energy stored in your coffee thermos is greater than the energy stored in your laptop battery, and it's a bit cheaper than your laptop battery. Storing heat is cheap right now in the thermos. You have hot water, which stores a lot of energy per degree, and an insulation around it. Depending on how good the insulation is, that'll tell you how long that thing will store energy. All those things have been around for a long time. Now, how are we going to heat these things electrically? How are we going to use simple technology? Because most people who are working on electric thermal storage are doing simple things. There are some exotic things using conductive materials, liquid metal things, but there are also simple things that people are doing.

DR: You're hitting directly on something, which is why I love this area so much, why it has caught my imagination so much. You have a situation where electricity was more expensive than fossil fuels for these purposes up until, like, five minutes ago. Now that electricity is cheap, we're looking for ways to store it and use it as heat in a lot of ways for the first time. And what that means is that there is very simple low-hanging fruit all over the place. My generation, when we think of technology or advanced technology, we generally think digital, and that generally means opaque. Like, we don't know what's going on in there. Yet these technologies of storing electricity as heat are so delightfully simple. You're literally just heating up a rock.

JOD: One of the electric thermal energy storage technologies actually uses rock, and on the outside of the pilot it says, welcome to the new Stone Age.

DR: So, to summarise where we've been so far, you need all this heat. Up until very recently, it was overwhelmingly cheaper to do it by combusting fossil fuels. But now along comes renewable wind and solar electricity, which are cheaper than anything. So now the challenge is how do you get the heat from the wind and solar electricity? As you say, the applications are running around the clock. Wind and solar come and go. So in between the wind and solar and the applications, you need something that's going to store that wind and solar that can release it in a steady flow.

JOD: Exactly.

DR: So, the new thermal storage technologies that are emerging are sitting right in that space, including Rondo. So if you're talking about something sitting in that space, what do you need out of it? What are the metrics by which you judge the performance of the thing that's sitting in between the renewables and the application?

JOD: You need safety, efficiency, cost and temperature at which the heat can be delivered. Some other things as well. One of them is how fast can you charge the system and deliver energy continuously. We'd like to charge perhaps during the solar day in 6 or 8 hours and deliver for 24 hours continuous. If you could charge in about 4 hours, we find that's even more valuable. The periods of zero and negative electricity prices in electricity grids are short. So, the ideal thermal storage can charge very rapidly. You can control its charging like other batteries, it could participate in providing grid services and it can run continuously. Shut it down once a year for inspection, and when the factory that it's connected to is shut down it just sits there and requires low operating and maintenance, costs.

DR: And I presume low losses too. But I want to pause and emphasise the first point you made. We have wind and solar all come online at the same time because they're all using the same wind and sun, so you have these periods of oversupply. I think people are familiar with this. You get more than the grid can use and today that just goes to waste. It's curtailed. That energy is not used. So, what you're doing is proposing to come along and use it. But if that's your economic sweet spot, those couple of hours of curtailed energy, you need your battery to charge as much as possible during those couple of hours. In other words, charge really quickly because the amount of energy available in those curtailed hours, especially in coming years, is going to be potentially huge. So you need to stuff a lot of energy in your heat battery really quickly.

JOD: That's right. Now the early deployments of heat batteries will use what is curtailed today. One of the things that we see that's uniquely pretty cool about this class of electric thermal storage is the total amount of energy that industrial heat needs is really large for scale. I think we had a fifty-two-gigawatt system peak in California not long ago. We've got about twenty gigawatts of solar PV in the state. Just repowering the boilers and furnaces that we have right now in California needs 100 gigawatts of new generation to replace those fuel British thermal units (BTUs), about forty of those gigawatts can actually be built without any connection to an electricity grid.

“

... if you want to move energy from July to January, chemical storage is a great thing because it doesn't have self-discharge. If you are in a place where you can have a salt cavern and you can make hydrogen in July and pull it out in January, that's great.

”

DR: I want to focus on how we're evaluating the heat battery. So, we want it to absorb a bunch of energy quickly. And then we want it to hold that energy with very little losses. This is the other fact about thermal storage that blew my mind that I do not think is widely appreciated, which is the incredibly low losses. People are accustomed to, I think if you want to store energy in hydrogen, you're losing about 50 per cent of your energy. But just heating up a rock, you get 90 per cent to 95 per cent of that heat back out of that rock. That is wild to me.

JOD: That's right, yeah. The least efficient of the thermal energy storage systems are around 90 per cent. We happen to be 98 per cent.

DR: So the heat just sits there in the rock and doesn't go anywhere?

JOD: Well, fill up your thermos with hot coffee, take the thermos and wrap it in a couple of blankets, open it up, three days later the coffee is still hot. It's not like a chemical system where there's self-discharge or something. The only place energy can go is either lost to the environment through insulation or delivered to the target. So it's a lot easier than it sounds. A lot of

people think, 'Oh, this efficiency couldn't possibly be the case.' It really is almost embarrassingly simple.

DR: When we say 95–98 per cent, what are the time horizons of that? If I fully charge a Rondo battery and then just don't do anything to it, how long would it take for all that heat to be lost?

JOD: We've got a food factory that runs shift work. They operate one shift five days a week. So you're storing some energy and you've got more energy on Monday than you did on Friday afternoon. The short answer is we lose about 2 per cent, 2.5 per cent per day. So if you were holding energy multiple days, there would be self-discharge. But that's because we were designing for a particular use case. You could decide the rate at which your thermos loses heat. You could make it store energy for months on end. Then the question is, is that valuable? If you really want to store energy for months on end, if you want to move energy from July to January, chemical storage is a great thing because it doesn't have self-discharge. If you are in a place where you can have a salt cavern and you can make hydrogen in July and pull it out in January, that's great.

DR: Because the hydrogen you pull out in January contains the exact same amount of energy...

JOD: Exactly. As long as it didn't leak out, but yes.

DR: Let's just discuss what the Rondo heat battery is, and tell us what some of the other options in this space are. I know you're heating up bricks. Some people are heating up giant chunks of graphite. I think sand is on the table. But what are people trying in that space?

JOD: The one technology that's been at scale for quite a while, that's been used by the solar industry since the 1980s is using nitrate salts, which melt at around 250°C (482°F). They're stable up to about 600°C (1112°F). It looks like a transparent liquid, but stay away from it. I built my first molten salt test facility back in 2008 at a national lab.

DR: I remember there was a hype cycle around molten salts that has kind of faded. Why has it faded? Like, why are rocks preferable?

JOD: The more you know about it, the less you like it. It's one thing to use it in a solar power station where there's nothing in there for a mile away except for the turbine. It's quite another thing for an energy storage facility to be put inside a factory where people are working. When I mentioned safety first, you don't want a system that can catch fire or spill

a superheated liquid that would burn everybody or release toxic gases. I'm not aware of any molten salt projects that haven't sent at least one person to the hospital. They work, but they have proven challenges.

DR: They just require a lot of engineering to contain.

JOD: Another matter that you mentioned previously is that technologies get cheap. Molten salt systems are a lot like nuclear reactors in that every one is bespoke. The modular approach, the factory-manufactured approach, eludes that technology.

One of the things you do first if you want to store heat is work from what's it cheap to store heat in. You mentioned stone, crushed rock, various kinds of rocks in a box or sand in a cylinder where you build an industrial-strength hairdryer. You blow superheated air through the rock or the sand bed. And then when you want heat, you push cool air the other way through the sand or the rock bed. That works. There are people taking it to scale. It has temperature and cost challenges. What you find in every one of these cases, the rock is cheap, but the box costs a lot.

DR: And the fans, I assume the fans are expensive and that kind of engineering adds to the cost...

JOD: That's right. And remember that your fan has to blow at your peak charging rate. But the big problem with those unstructured materials is when they heat up, they expand and you have to have a container strong enough and then when they cool, they shrink and settle and then the next day they expand again and they slowly turn into dust. So the material looks really cheap, but the system turns out to be not so cheap.

There are a lot of interesting science experiments with new materials that have never been used this way before. When we started Rondo, we did a really careful look at everything that was out there. There are people using liquid silicon. It melts at 14°C (57°F) and stores a lot of heat. Just like ice melting in a glass, it absorbs a lot of heat melting and releasing silicon. Freezing silicon is a really good thing for high-temperature heat. But how do you make the glass that's holding that silicon-ice? How do you keep it? There are a lot of challenges that companies have been working on for years and it's probably going to take another decade before that technology is at the point that an ordinary project finance guy will say, yes, that's as low risk as PV. I'll invest in that at the same finance rate. That time-to-bank ability is one of the biggest issues. If you want a technology to go big fast, everybody's got to agree it's boring and low risk and that's a challenge with new materials.

Graphite is another material that's interesting. It has higher heat capacity than rock or brick, especially

“

... you build an industrial-strength hairdryer. You blow superheated air through the rock or the sand bed. And then when you want heat, you push cool air the other way through the sand or the rock bed. That works.

”

when it gets hot, but it catches fire at 560°C (1040°F). So, if you want to store energy at 1500°C (2732°F) or 2000°C (3632°F), you've got to keep it in some atmosphere so that it can't catch fire. Also, it's conductive electrically, which could be great. There are interesting engineering challenges and there are at least four companies working on that. One of them, Antora, is looking at using that graphite not for electricity to heat, but electricity to heat to electricity. Using PV cells to capture the light from the graphite.

DR: I've spoken with them too. In terms of science-fiction geeky fun, that one is just great. They heat the graphite up, it gets so hot that the energy comes back out as light. So they have it covered in shutters that they can open incrementally. And the light can either shine on tubes full of fluid if you want heat, or these special PV modules that they built especially for it. If you want electricity, like the whole conceptually, that's very satisfying.

JOD: It's super cool. My first job was infusion power, where you have a reactor that wants 100,000,000°C (180,000,032°F) plasma right next to a superconducting magnet that has to be 5°C (41°F). The Antora PV challenge, when they solve that, will be great for electricity to electricity because it could turn out to be long duration, no moving parts storage. That's an example of something deeply innovative. The back to electricity is their superpower. How long will it take to prove that it's bankable? What we're doing is much more boring.

DR: But you guys have settled on, rather than any of these materials' science experiments, bricks.

JOD: Somebody asked me the other day, how many gigawatts of batteries are there in the world right now. Do you know?

DR: I don't.

JOD: They said there are about three gigawatts of batteries in the world right now.

DR: Lithium-ion batteries, you mean?

JOD: Yes. So how much heat storage is running in the world right now? As we speak, there's about thirty gigawatts of heat storage running right now. In 1828 there was the first patent for a thing called a cowper stove, which is a tower with a thousand tons of brick in it that has air passages, which on a one-hour cycle, has the still combusting exhaust of the blast furnace blown down through that tower and heats all the brick to about 1500°C (2732°F). Then, for about 20 minutes, fresh air is drawn up through the tower and it's providing the inlet air to the furnace and it's delivering 115 megawatts heat for about 20 minutes. Things are heated and cooled twenty-four times a day. They last 30 years.

DR: So if I'm looking inside a Rondo box, am I literally just looking at a stack of bricks?

JOD: Pretty much. As you know about bricks, they are brittle. If you drop a brick, it'll break. You also know that brick is not a good heat conductor. That's why we make fireplaces out of it. So, if we want to heat it fast, we have to heat it uniformly. If you put one side of a brick in a bucket of water and the other side in a fire, the brick might fracture. But if you put the brick in the middle of the fire, it'll heat up rapidly to the temperature of the fire. It's one of those ideas that once you see it, it's obvious. But it only took eighty design revisions.

If you look inside a Rondo unit, what you'll see is a brick stack that's full of these open chambers. It's a checkerboard of open boxes surrounded by brick, and brick surrounded by these open boxes. And electrical heaters are embedded directly in the stack, and they provide radiant heat within those open boxes. Things that can see each other get to become the same temperature by exchanging heat. So, the result of this was we found a way to directly, rapidly heat the bricks.

DR: And this is an alternative to blowing hot air over the bricks, which, a) would require more engineering and more money, but b) also might not heat them uniformly, like might heat one side before the other side or something like that.

JOD: You can heat them uniformly. Blast furnaces do that. But in that case, you have the same electrical heater that's in something like a hairdryer. And inside a hairdryer, the heaters are mostly radiating to the metal plates, which in turn are heating the air, which in turn would in this case, heat the brick. There'd be a couple of hundred degrees difference between the final temperature of the brick and the temperature of the wire. In our case, that's about 5°C (41°F) degrees.

DR: So instead of using the wire to heat the air, to heat the brick, you're just sticking the wire in the brick, and the wire is heating the brick directly.

JOD: That's right. So we, just last week, announced the world's highest temperature thermal energy storage system running. That's not because we use different heating materials than others. It's because of that physics insight that led to that structure.

DR: What are some of the engineering challenges here? Do the bricks expand and contract when they are heated, or do they degrade over time? What sort of things are you dealing with here with bricks that you had to overcome?

JOD: There were lots of things because what we're talking about is kind of, at some level, obvious, and people have done really good work on this previously. Yes, the bricks expand and contract, so build your structure accordingly. The nice thing is they're freestanding. They don't need a container to hold them in. So if you build your structure properly, it can freely expand and contract.

DR: One of the hard engineering problems is making sure that the temperature inside the material is uniform.

JOD: Uniform not just when the unit is new, but when it's 30 years old.

DR: Your promise here is that this Rondo battery has the same capacity and the same performance characteristics in 30 years that it does today. Is that the idea?

JOD: That's exactly right, yeah.

DR: And there's no other battery that can say that.

JOD: I think that's true.

DR: Let me ask about getting the heat out to where it needs to go. You face the challenge of how do I get that heat to where it needs to be without losing a bunch of the heat?

JOD: So, every combined cycle power station in the world has a jet engine that's generating electric power. Its exhaust is around 605°C (1121°F). That exhaust is passed through a boiler, a heat recovery steam generator that drives a steam turbine that makes extra electric power. So, the world knows how to build those boilers that run on about 600°C (1112°F) air.

The Rondo storage is much hotter temperature, that we mix down. And for the systems that are delivering steam, we work with leaders who build conventional boilers and we've engineered the heat battery to include that boiler. So the basic heat battery models are exact drop-in replacements for particular models of industrial boilers. They're just about the same size. Stick us next to your existing one, hook us up to the pipe.

DR: You're replacing a fossil fuel-run boiler with a heat battery and a boiler in the same space.

JOD: It is a like for like drop-in replacement. The less work the customer has to do, the better off we are.

DR: To what extent is the Rondo heat battery plug and play, in a low-temperature steam application or a steel plant. Can you wander into any of these and just switch out with no pause?

JOD: The top four categories in descending order of industrial heat use are chemicals, food and beverage, paper products (everything from toilet paper to cardboard,) then cement, and then steel. So, for chemicals, about a third to 50 per cent of all the heat is steam. For food and beverage and paper products, it's all steam. For cement and steel, none of it is steam. We're delivering drop-in boilers today and simultaneously with our investors and partners building and developing the calciners, the ethylene crackers, the kilns, to drive particular industrial processes.

We have 200 years of designing industrial process units that are powered by fuel. Which of those can we retrofit? Where will we need to design new things? We were given a grant by the Danish government. We have a project underway to design and pilot a true-zero cement process, intermittent electricity to zero-emission cement. Most of the work in that project is the design of a calciner that instead of internal combustion, runs on superheated air or superheated CO₂.

So, it won't all happen all at once. The high temperature things will take more work to integrate because industrial plants today were designed with magnificent engineering and heat balance and efficiency burning fuel. We have work underway now and hope to have results over the next couple of years that use the same thermal storage platform.



Photograph: Kumpan Electric

DR: But this first commercial battery that you've deployed, what application is that or what temperature level is that?

JOD: Yeah, that's targeting steam. The particular installation is at a biofuel producer. Whether you're making renewable diesel from soybeans or animal fat or ethanol from corn, about half the total carbon intensity of that fuel is fossil fuel that was burned to produce that biofuel. We can set that to zero. So we can produce biofuels that are about half the carbon intensity of what they are today. Interesting, our customer is really a visionary that's going to zero because the other thing that's been talked about a lot with biofuels is combining carbon capture of the biogenic CO₂ in those facilities. As it happens, using Rondo for the heat eliminates about half the total carbon intensity. Using carbon capture eliminates about the other half and together you get about essentially a zero-CI, zero-carbon-intensity fuel. That little unit we just started up is the pilot for deployment

of a series of larger ones to do exactly that, to produce zero-carbon biofuel.

DR: Very interesting. So, let's pull the lens back a little bit, maybe talk about business model. Is the idea long term that if I'm, say, a manufacturing facility and I'm making I don't know, baby food, is the idea that I buy a Rondo unit and install it in my factory? Or is the idea that Rondo comes in, sets things up and sells me heat as a service? In other words, am I buying the equipment or am I buying the heat? Or some of both.

JOD: There are many answers to that question. Sometimes people own their own cogeneration plant. Sometimes they contract with someone to provide them electricity or heat as a service. The renewable heat as a service business will develop the same way. In the United States today, there's a huge community of developers who know how to shave a few pennies off solar and wind electrons, but have never really looked at these industrial facilities. In Europe, actually,

“

For sure the lowest cost way that we're going to decarbonise all of civilisation is electrification.

”

there are already renewable developers who are out there originating renewable industrial heat projects.

So, first of all, Rondo is offering, on four continents, commissioned, guaranteed installed heat batteries. That's the foundation. We are also originating and financing heat as a service, principally in North America.

Whether you make baby food or steel, you don't drill gas wells to get the fuel to run your process. You buy energy as a service, most folks want to spend their capital dollars on their own processes. This thermal energy storage class is arguably creating one of the great business opportunities of our time for the development community. We all know wind and solar deployment is slowing down, not because of reduced demand, but because of congestion. And I think the interconnection queue time in England is now 13 years.

DR: Yes, there's like a terawatt now, I think, waiting in the queues.

JOD: Right. Rondo heat batteries. Our basic unit, the RHB 300, needs seventy megawatts of generation. Typical installations may have two to ten at a single site. These are utility scale energy demand and they can be built with no grid connection.

DR: Right. So the idea is you go build a solar farm or a wind farm that is just attached to these batteries, and then you're selling the heat from the batteries. So

at no point do you need the electricity grid. You're not waiting for the interconnection or anything else. Wind and solar being so cheap, the implications are endless. You're free from fluctuations in fossil fuel prices and you're free from any worry about escalating carbon prices or other carbon-related regulations.

JOD: That's right. This matter of what kind of risks do we take? People say it's risky to work with this new technology, but look at the risks that we just were used to taking. We're entering this new world where we're not talking about a green premium, we're talking about the same or lower energy cost with these reduced risks. Then, of course, depending on what the commodity is, low carbon aluminum trades at a price premium on the London Metals Exchange. Low carbon fuels trade at much higher prices in California and Germany. And, for consumer-facing brands, there are producers who are seeking low cost effective renewable heat sources so they can offer to the market low-carbon commodities.

DR: It seems like there ought to be a bunch of market actors that are just ready to embrace this. Like, for one thing, as you say, just on a quantity basis. If you take all that energy that we're using for heat and transfer that to electricity, you need a lot of new electricity and a lot of new clean electricity. So it seems to me like renewable energy developers ought to be over the moon about this, like beating down your door. Are they lining up to be proponents for renewable heat in the industry generally or have they not caught on yet?

JOD: In some places the answer is yes. Europe is very aggressively moving in this direction and a number of folks over the last few years have said 'this Rondo thing sounds too good to be true. Come back to me when you're operating something commercial.' We're now operating something commercial. These projects offer a mix of speed and certainty, as we're not tied up in a grid queue. We're in an era where the numbers work for the heat user, they work for the financier, they work for the builders of the solar fields and they work for us. That's a new world, with economic tailwinds driving it. It will keep going faster and faster. The size you mentioned, I think at the end of 2021, there was about 1000 gigawatts of wind and 1000 gigawatts of solar each in the world. The IEA did an assessment of industrial heat and their number is it's about 9000 gigawatts of new generation that's going to be required to replace the oil, coal and natural gas now being burned.

DR: Good grief.

JOD: That's worldwide. Twenty per cent of that in the US.

DR: An enormous opportunity to build more renewable energy. Electric utilities are sort of notoriously worried about grid defection. You represent potentially a new responsibility for them. Something that natural gas utilities were doing is now all going to transfer and be their responsibility, which is just a way for them to grow and invest and a new opportunity for them. Why aren't they at the front of the line beating down the door, trying to make this happen faster?

JOD: That's a great question, and they are. One of our investors is Energy Impact Partners, whose backers are the North American electric power industry. For sure the lowest cost way that we're going to decarbonise all of civilisation is electrification. For sure the electric industry is at the heart of that. These things become an asset in the electricity grid and a solution to these problems of variability and over generation and balancing.

DR: In the same way that any controllable load helps grid stability. These are controllable.

JOD: Demand response, for example, is a load that you expect to run all the time, but you can turn it off during emergencies. That's not this. This never raises the peak demand on your transmission or distribution system. It's different than anything that's come before. It's like lithium-ion batteries in that sense, but at a tiny fraction of the cost. We're not trying to solve from moving electric power from noon to 07.00 pm.

We're opening an entirely new segment to renewable deployment. The electric utilities are getting involved now. They face all kinds of issues with the regulatory framework that we have for electricity. There are some new challenges, but there are people actively working on that and we're thrilled to be working with them.

DR: So if I've got this manufacturing facility, I've got a big Rondo battery and I'm trying to decide between two options. One is I could build my own off-grid behind the meter generation, solar and wind. I could put my own solar and wind up, or I could just get on the grid and time my charging so that I'm chasing the clean energy on the grid so that I'm only charging when there's clean energy on the grid. Do we have any sense of which of those will be more economic or why you'd want to go one way rather than the other? I'm just wondering how many of these sort of self-contained, off-grid, purpose-built renewable energy installations there are going to be. It seems to me intuitively like that ought to be more expensive and what you ought to prefer is just for the grid itself to clean up so you have more, so it's easier. But what are the choices there?

JOD: These questions are right at the heart of the matter. You're dead on. The short answer is it depends primarily on where you are. Oklahoma last year had 2000 hours of negative wholesale prices. If you put a project in Kansas or Oklahoma, you have energy prices that are slightly negative on an annual basis. If you can charge very rapidly, if you are allowed to participate in the wholesale market. There are regulatory obstacles.

DR: But in theory, in Oklahoma, during a time of negative wholesale prices, your facility that's running off a Rondo heat battery could be paid to charge itself. Is that how that works? Is that what negative prices means?

JOD: That's what negative prices means.

DR: That's so mind-blowing.

JOD: And we have lots more of that coming. I know you've spoken to folks about the IRA. The production tax credit coming to solar is going to broaden the areas of the country where we see intermittent negative prices. Because, of course, if I'm getting \$20 megawatt hour for tax credit, I'm perfectly happy to generate when prices are negative \$19, right?

DR: Yeah. That's just crazy.

JOD: Technologies like this that can absorb those periods are going to lift the price floor. They're going to benefit all the generators, especially the generators that can't turn off. We're pretty excited. But again, it's can we connect to the grid? Can we capture those prices?

DR: Because if you can, there's enough heat to absorb all the curtailed power in the US, times a gazillion. Theoretically, if you could hook up all heat to electricity, you'd never curtail again, or at least not for decades. Probably.

JOD: Of course, subject to where is the heat-load versus where is the curtailment. Some curtailment is regional associated with total generation. Some of it is transmission constrained. But to a first approximation of the answer that was correct, yes.

DR: Yeah, that again, seems just a crazy business opportunity for everyone involved.

JOD: Yeah, we agree.

DR: Do you expect to see these off-grid, custom-built renewable energy installations, purely powering heat batteries in areas, say, where the grid is congested,

or the grid is dirty or the interconnection queue is unusually long. Do you expect to see those pop up?

JOD: California has in the order of twenty gigawatts today. We need 100 gigawatts of new PV just to replace the BTUs of fuel now being burned for industrial heat. About forty of those gigawatts, because of where the things are sited, could be built with no grid connection at all. Most of them will need some kind of grid connection. We see again and again that the new renewable project development model is going to be building a project that part of its electricity goes to industrial heat, into a heat battery, and part of it goes to the grid. That's the sweet spot that delivers lower cost electricity to the grid.

DR: If I'm a renewable developer and I catch wind, that there's this whole category of renewable projects that don't require this unholy paperwork nightmare that they all go through. Now again, I just can't imagine that they're not going to be stampeding in this direction. I mean, I hear them complain about this constantly.

JOD: The financial community have to get their minds around a few things. How are we structuring these projects where most of the energy is going to a single factory rather than to the utility? What about the credit worthiness of that? And how long will it take to retire the Rondo technology risk? Once you turn one of these things on and operate for six months, there's nothing left to prove. We know it works and we already know everything is durable.

DR: I would imagine also that this space is going to see a lot more competition. Of course, once it's kind of uncorked and it becomes clear what the opportunity is.

JOD: Look, trillion-dollar markets don't happen without lots of people trying to enter them and nothing could be better, right? That's what we urgently need.

DR: I guess this will take years and decades, but do you imagine areas of intense renewable capacity, with lots of sun and lots of wind, becoming new attractors to industry? Do you see global industry starting to migrate to renewable energy? Is it that much of a chunk of the cost of an industrial facility that it might be worth someday literally moving to it?

JOD: The short answer to your question is yes. Just look at what happened with the shale gas revolution in the US. Vast investments in petrochemical and other manufacturing immediately shifted to where low-cost energy was. There's a question of how fast these transitions happen. But, in the long term, yes, absolutely.

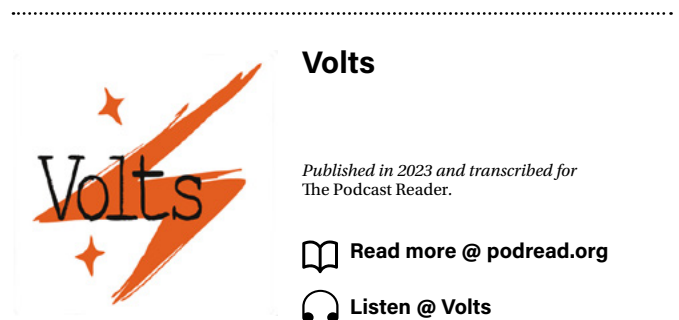
DR: Wind and solar being so much cheaper now than they were five to ten years ago, it's not an incremental change, it's a phase change. It's a flip to a different system. All we're doing now is sort of one at a time here and there in different industries, in different places, kind of opening our eyes to, oh, this is a completely different landscape. It's a different world now. It's going to take a while just to absorb the implications of super cheap renewables.

JOD: And the thing we know for sure is that every year somehow those cost reductions will continue. If you look back over every five-year period, every year's forecast was wrong, it fell faster than predicted. It's reasonable to assume we're going to continue to be in that, so that this era that we're entering, it keeps getting better and better. Our storage technology and the other storage technologies will cost reduce. But the storage technology is only 20 per cent of the cost of the total project. The fact that the wind and solar are coming down so steeply, this cost advantage is going to continue for the people who have made this transition onto renewables.

We're manufacturing in two locations now. A lot of our material science will be driven by qualifying other sources of materials. We've produced now on three continents, little pilot-scale things. So, one chunk of material science is about just getting this 2 million tonne a year at scale. The company's formal goals are 1 per cent of world CO2 in a decade and 15 per cent in 15 years. And there are no material blockers to doing that. Today we're using the most expensive brick materials, the highest temperature, highest strength. There will be innovations in simply reducing cost.

DR: Thank you so much for spending all this time with me. As you can tell, I find this particular area so interesting and fascinating.

JOD: Thank you, Dave. It's a real privilege to speak with you. I'm just delighted. Thanks so much.



How China's Economic And Political History Can Help Us Understand its Future

How China's imperial exam system stymied civil society

YASHENG HUANG

CONVERSATIONS WITH TYLER, 2023

Interview by Tyler Cowen

Yasheng Huang has written two of Tyler's favourite books on China: *Capitalism with Chinese Characteristics*, which contrasts an entrepreneurial rural China and a state-controlled urban China, and *The Rise and Fall of the EAST*, which argues that *Keju* – China's civil service exam system – played a key role in the growth and expanding power of the Chinese state.

Tyler Cowen: I am here with Yasheng Huang, who is professor of management at the MIT Sloan School. He has written the famous book, *Capitalism with Chinese Characteristics*, on the history of Chinese economic reforms. He has a new book coming out, which I found fascinating. It is called *The Rise and Fall of the East: Examination, Autocracy, Stability, and Technology*, due out this summer. Yasheng, welcome.

Yasheng Huang: Thank you, Tyler. Good to be with you.

TC: I have so many questions about China. Let's start with one. Why did the Chinese state fiscally centralise so late in its development?

YH: Well it depends on your definition of development. In one definition, you can say it was *overdeveloped* in the sense that you essentially *only* have the state, and you don't really have private economy. You didn't have a real meaningful society. You didn't have an independent intellectual class and you didn't have organised religion. All you had was the state, and there were pockets of the society the state was not able to reach. That's true, but where the state was able to reach, it was all-domineering. In that sense, the Chinese state was overdeveloped.

I think it was probably underdeveloped in a sense that, in part because of this dominance of the state, it didn't really develop an administrative capacity to tax the population. It didn't develop the institutionalised support for public services, even though the state did provide some public services comparable to what was available in Europe. It was not able to, for example, have a proper army. That's why China's history was full of northern nomadic tribes taking over the country. It was not that professionalised in the modern sense of the word. It was underdeveloped in that sense.





Photograph: Shio Yang

TC: Let's take, say, the middle of the 1990s – what percentage of GDP is federal government revenue in China?

YH: Well, it was very small and I don't know the number. But if you go to the economic historians' work on Qing Dynasty (1644–1911) – their rough estimate is that China had a much smaller state in terms of the physical capacity as compared with Europe during that period of time. To some extent, we struggle with the same issue today. If you look at the tax revenue relative to GDP, even in China today it is not an excessively high ratio. But if you look at the ownership role of the state, it has a big state sector. It has a big influence over the private sector.

I would argue that the Chinese state in the nineteenth century did not derive all its power from taxation. It had the power to appoint officials. It had the power to control the private sector, control the merchants. That kind of administrative power was quite substantial.

TC: When China needs to fight Britain in the Opium Wars, or come the 1920s – is that lack of fiscal capacity what's holding back China?

YH: The overall lack of economic development and the lack of physical capacity was part of that. But the bigger picture is, by our own data, the Chinese state ... in the nineteenth century China was no longer inventive. It forgot the inventions that it was able to make many, many centuries before. It didn't have a naval power as it used to have back in the Song Dynasty (960–1279). In the Ming Dynasty (1368–1644), there was a famous event, the Seven Voyages, where the Chinese ships were able to travel to Africa, to Arabia. China relinquished all that naval power by the nineteenth century.

TC: I read so many experts insisting that China should rebalance its economy toward consumption, yet China never seems to do this. Are the experts right? Is China right?

“

It entails implications about the future prospects of the country, whether or not you can sustain this level of economic growth purely on the basis of massive investments.

”

YH: No, the experts are not wrong, but the problem is that the low consumption to GDP is a symptom. It is not a cause of the imbalances of the Chinese economy. The bigger issue, I believe, is the fact that the household income share of the GDP is very low. You have to save a fixed proportion of your household income. When the household income relative to GDP is low – and at some point, it was declining – the consumption, out of the mathematical necessity, is accounting for a smaller and smaller share of the GDP.

The issue is really the household income. If you look at the government, corporate and household sectors, the household sector, as a proportion of GDP, was never high. That gets to a deeper issue about, essentially, the power of the government vis-à-vis the household, the power of the corporate sector vis-à-vis the household. That's a bigger political economy question rather than a narrow fiscal issue.

TC: What's the biggest misunderstanding that American business elites have about the Chinese economy?

YH: Well, I wouldn't say it's the biggest, but one of them is that they look at the Chinese R&D spending, and they look at, for example, some of the impressive technological progress the country has made, and then they draw the conclusion that the Chinese economy is driven by productivity and innovations.

In fact, studies show that the total productivity contributions to the GDP have been declining in the

last decade and even more. As China has begun to invest more in R&D, the economic contributions coming from technology, coming from productivity have been actually declining. In an economic sense, it's not a productivity-driven economy. It is an overwhelmingly investment-driven economy. I think that's one of the biggest misunderstandings of Chinese economy. It entails implications about the future prospects of the country, whether or not you can sustain this level of economic growth purely on the basis of massive investments.

TC: Why don't the Chinese want to have more children? Urban areas, the TFR (total fertility rate) can be below one even, right? You're lucky if it's one.

YH: Yes. By the way, this is an East Asian phenomenon as well. Part of it is the socialisation of the norm, because of the one-child policy that they instituted in the late 1970s, which was only lifted in 2015. Essentially, you had many, many years in which people were socialised in the norm that one child is the norm.

TC: South Korea has done the same, right? They didn't have a one-child policy.

YH: Yes, I know. That's the point I was trying to get at. Japan didn't have a one-child policy. Japan has a very low fertility rate. For some reason, it's comparing apples with oranges. Japan is a much more developed society. The per capita GDP is much higher. Essentially, the fertility rate would probably be lower naturally as a result of that difference in per capita GDP.

But I think it's probably more than that. The nature of the Chinese economy and Japanese economy is very land intensive, so the urban costs of living are very high because of the very high housing costs. So, the common complaint we hear from young people in China is that they cannot afford the children. There's a cultural norm that you need to have your own housing to be able to get married, and that may also deter family formation and, therefore, production of children.

TC: What's the future of immigration into China as population declines? And it has just started declining. Who wants to go there?

YH: Well, it's not a very attractive country to go to now. Look at what happened during the COVID public health crisis, and this erratic policy from Zero-COVID to essentially no policy in place to control any kind of virus transmission. China is now reporting 60,000 deaths related to COVID, and most people believe that's a vast underestimate of what the true number is.

Erratic policy on economic management is also incredibly jarring. It undermines the credibility of government policy. I think that the bigger worry is foreign capital, ahead of foreign migrants. Foreign capital and reorganisation of supply chains. China used to be a factory for the world. Now many companies are rethinking whether or not they should rely so heavily on supply chains in China. For example, if you look at the recent iPhone 14, parts of it are made in India rather than in China. I have heard from many companies that they are going to move out parts of production that cater to world markets to other countries. They are going to have the production catering to the domestic market stay in China. China, increasingly, is becoming a factory for itself rather than a factory for the world. From the immediate economic perspective, the reluctance of foreign capital is a bigger concern than the lack of immigration. In the long run, probably the labour supply becomes a bigger issue, and they may rethink their immigration policy.

TC: What are the possible Chinese origins of the phrase *laissez-faire*?

YH: Well, I'm not an historian, per se. I'm not an historian of that particular phrase, but there is at least this belief that that phrase originated from China in a sense that there was a European belief that the Chinese meritocracy conferred a lot of autonomy on meritocrats. They could do the things that they wanted to do without close supervision of the emperor. One version of the origin of that phrase is that Chinese meritocrats, the Chinese court officials, the Confucian mandarins, enjoyed a lot of operational autonomy, and they could do whatever they wanted to do. We know, for a fact, that's just not true. The Chinese court officials didn't have any operating autonomy, nor did they have any ideological autonomy.

TC: Why did Chinese literacy appear to stagnate in the nineteenth century?

YH: Yes. This is a deeply puzzling development. Let me make the argument why there was a potential for China to raise its literacy. They implemented the examination system back in the sixth century, and the state began to provide not universal education, but something closer to universal education. It was more like a preparatory education for boys and men to be able to take the civil service examination. The costs were subsidised by the state, and the apparatus, the infrastructure was quite widespread and developed. There was a cultural premium placed on being able to read and write. They had all the conditions for further literacy and even universal literacy.

A number of factors hindered the further

“

... the civil service examination system was incredibly successful. Even though the number that eventually made it is not that big, the bottom of the pyramid is big, and the people who are most at risk of making societal ideological breakthroughs are completely homogenised.

”

development of literacy. One of the bigger factors was that this entire system was designed for half of the population only, for the male half of the population. Females didn't have access to the preparatory schools, and they couldn't take the civil service examination. There was some spill-over from males to females in terms of basic literacy, but the scale was not that big.

The other was that the entire education was organised around memorisation rather than teaching people how to think. One problem with that was that there was never any liberalising value that came from rote memorisation, the memorisation of the Confucian texts. And those texts were usually extremely conservative and backward-looking.

I think the third reason is economics. The Chinese economy didn't develop, commercialisation didn't develop, at a scale that we saw in Europe. And that didn't increase the demand for human capital, for basic skills. If you run a bureaucracy for a country the size of China, you may need 10,000 people, 20,000 people, 40,000 people, 100,000 people, but if you run a commercial economy, you need millions. You need tens of millions of people working in factories and doing this and that and producing this and that. The demand for human capital when you don't really have a commercial economy is always going to be limited if all you need is the human capital to run the bureaucracy.

I will say that all these factors combined together restricted what could have been an earlier literacy revolution that failed to occur in China. If you look at, for example, going back to the sixteenth century, China had decent literacy compared with some European countries. But the difference is that in Europe, you had this explosion of literacy, whereas in China, it stayed at the same level. The initial level didn't translate into a higher level, whereas in Europe it did.

TC: Now, you argue in your new book that the imperial exam system weakened the horizontal structure of Chinese society, prevented China from developing civil society. What's the mechanism for how that works? I can see that the exams might pull away some smart people, but there are still plenty of people left in China. Why don't the remaining people develop some kind of civil society?

YH: The norms are not shared beyond the select few who succeeded at the exam. It's the bottom of the pyramid. It is not that the people at the bottom of the pyramid didn't aspire to become a bureaucrat, didn't aspire to excel at the exam. They all wanted to become a bureaucrat. They all wanted to be educated in Confucian ideology.

TC: Why are the norms so homogeneous? You go back to the Song Dynasty – southern China is extremely commercial, very diverse. The whole world passes through China, the Silk Road. Why did the norms end up so stultified, so homogenising?

YH: What the examination system did was, it was homogenising the smartest people. Essentially, the smartest people were socialised, and their ideology was homogenised into this very stultifying Confucianist ideology. I didn't quite say it in my book, but I think this is a plausible explanation.

If you begin with a worldview that smart people are the people who begin to make breakthroughs in ideas, in the organisation of the politics, in the organisation of the economy, those people in China – the smartest people in China – were completely homogenised. Yes, you have a lot of other people that were not able to be part of this system, but they were not the ones that were naturally disposed towards making transformational changes, towards coming up with revolutionary ideas, towards coming up with a new religion, a new way of thinking.

So, in that sense, the civil service examination system was incredibly successful. Even though the number that eventually made it is not that big, the bottom of the pyramid is big, and the people who are most at risk of making societal ideological breakthroughs are completely homogenised.

TC: Does the history of civil service exams in Korea and Vietnam run the same way, or is that different?

YH: They were less watertight as compared with the Chinese system. They were probably less well-organised than the Chinese system. In Japan, they also had a version of the civil service examination system, but they didn't really continue the system beyond the twelfth century, thirteenth century. In the nineteenth century, when Japan embarked upon modernisation, it was a Japan similar to China at the very beginning stage of the civil service examination system, rather than towards the end of the civil service examination system. At the end, the exam system was extremely watertight, extremely homogeneous and extremely well organised. Japan evolved into a modern society with a looser version of the Chinese system, as compared with China in the nineteenth century, which responded to the challenges from the West from a very, very rigid system.

TC: If I think of China, economically speaking, in the seventeenth century, it still seems broadly on a par with Western Europe, but by the nineteenth century, that has changed. Is it that the civil service exam system in China became more negative, more dangerous? Or is it simply Europe raced ahead and China stayed put?

YH: I think there are debates about whether, in the seventeenth century, China was really on a par with Europe. There was a famous California school that made that point. Their argument was, it is not right to compare the whole of China with Europe. You need to compare the most developed part of China with Europe. If you do that, the level of economic development was comparable. I think there's some recent research that challenges that view, using better data, more fine-grained data that shows that per capita GDP by the seventeenth century in China was much lower than per capita GDP in Europe, even among the most developed regions of China.

Our own data do not support the view that China in the seventeenth century was comparable to Europe. We don't use economic data; we use inventions data. By the seventeenth century, China was already not very inventive as compared with the sixth century. The level of inventions declined dramatically in the seventeenth century, so I tend to lean towards the view that by the seventeenth century, China was already quite backward.

TC: If I think of the Chinese exam system today – not just for civil service, but more generally – it seems fairly meritocratic to me. I often meet smart

Chinese graduate students. I ask, 'Where are you from?' The answer is not, 'I'm from a wealthy family in Shanghai.' They're from some part of China I've never heard of. Am I wrong to think that it's currently quite meritocratic?

YH: No, you're not wrong at all. It is one of my central points in the book that the power of the Chinese civil service examination system was, in a very narrow sense, very meritocratic, very well proctored. There was public corruption, but the corruption was not endemic. There were guard rails making sure that the integrity of the exam was at a level sufficient to attract the interest and the participation from the masses of the people. We showed that family background – and this is from the data from the Ming Dynasty, fourteenth century, fifteenth century – that family backgrounds didn't affect the performance of candidates who took the exam, which is quite remarkable. If you think about SAT today, even today there's some correlation between exam performance and the backgrounds of the people who take the exam.

The Chinese had a very well-developed system to make sure that economics didn't affect performance through a subsidy program, through a very well-administered protocol of the civil service examination system. Anonymisation – China invented a double anonymisation system. The examiner didn't know who the examinee was; the examinee didn't know who the examiner was. It was a very sophisticated system. They also hired scribes to copy the exam papers to make sure that the handwriting did not provide any information about the identity of the candidates. From that tradition, China now has a college examination system which is very, very tough, which is very, very strict. One of the justifications for having that system is that because China is so corrupt, this is one of the few areas where corruption has not affected the outcome, and therefore, we should keep the system.

When you make the argument that maybe we should change the examination system to allow more creative thinking, less emphasis on memorisation, the pushback has always been, 'Look at the country. It's so corrupt. Corruption is so pervasive. This is one area that is not corrupt. Let's not destroy that isolated area of cleanliness and capabilities.' So I have no problem with what you have just said, which is that the Chinese exam system is very meritocratic.

TC: If I look at China today, I'm never quite sure how much civil society I should think China has. One sees the Zero-COVID policy being repealed quite suddenly, possibly because of all the demonstrations. Could it be there's a lot of civil society in China? It often takes the form of demonstrations, which are highly numerous and frequent in China, at least before Zero-COVID. Then a lot of the rest of it takes

place on WeChat, which is not quite visible. But isn't Chinese civil society much stronger than it looks if we apply, say, Western benchmarks?

YH: No, actually, if you apply Western benchmarks, it is a very weak civil society. There's a difference between a civil society consisting of isolated individual actions and a civil society that consists of organised activities that have a program, that have financial support, that have the capability to operate independently. By the second criterion, China has none of that. If you look at the recent protests against Zero-COVID controls, let's keep one number in perspective. By various estimates, in 2022 there were probably 400 million people under some sort of long-term quarantine. And let me just concretise that word *quarantine*. That means you're essentially locked up in your home, sometimes for weeks, and in some cases, for two months.

That's the level of the suffering, and sometimes you can't get food. Sometimes you cannot get patients into the emergency room because the hospitals also shut down, refused to take in patients who tested positive or who cannot show a negative test on COVID. Relative to that, China experienced a wave of protests – by one estimate, in seventeen cities. I don't really have a good idea how many people were involved, but we are not talking about millions of people. We're talking about maybe 10,000 people, or tens of thousands of people. Contrast that with Iran. In the case of Iran, one woman died in the hands of the moral police. There were other grievances, but that was the trigger. The protests are still going on. Millions of the people took to the street. In Iran, religion played a big role. If you look at the colour revolution in Tunisia, it started with a peddler whose assets were confiscated by the government official, and then he committed suicide. That sparked the colour revolution. Those kinds of brutalities towards small peddlers happen almost on a daily basis in China. It's very important to specify, relative to the grievances and the level of the misery ... we're not talking about large-scale social movements here. These are individual actions. These are spontaneous actions. The complaints on WeChat – these are not organised. These are essentially individuals expressing their frustrations. They are not really using WeChat to coordinate their actions. Maybe they're implicitly coordinating their messages by supporting each other, but we don't really see any evidence of organised activity. In that sense, the civil society is quite weak.

TC: The CCP (Chinese Communist Party) and the governments – the local governments – they are quite responsive to citizens on some, but not all, issues, right? There's an embedded civil society where a message is sent. You don't need to do all the

organising. You get a fair amount of response, and it's a kind of shadow civil society.

YH: Yes, but that's a different framing. This is about a government. Even though it is autocratic, it is still reasonably responsive to the demands of citizens. And, therefore, the argument is that you don't really need organised civil society to press their demands, to press their policy preferences. I think in some sense, that's correct. If you look at what the CCP has been doing, it is actually quite clever. It's not the case that they don't take input from the society. They create portals, they create websites, and they create phone numbers for the citizens to call in. They also do surveys. They want to solicit opinions and information from the citizens without creating conditions for the citizens to get organised. If you think about all these opinions expressed to the government through the government control portals, you are doing it as an individual. You're not doing it as a member of a larger group. The CCP has no problem with that, and sometimes those opinions can be quite negative. The CCP has no problem with that.

The beauty of the system is that once you express those opinions, convey that information, as a CCP official, I have a number of choices. I can act on them to alleviate your concerns, and I also make sure that you don't find it necessary to organise protests and demonstrations. That's true on traffic, pollution, things like that.

The big difference between Zero-COVID and these other issues, such as traffic and pollution, is that so many people are affected by the policies simultaneously and to a similar extent. Therefore, they can relate to each other much, much more than in the previous situation, when you think about your own treatment mostly as an isolated, individual situation rather than something that you can generalise. That, I believe, is the main reason why you see this level of demonstration and protest.

Yes, China has had a lot of protests, but those protests tend to happen in rural areas, in less urban settings, in isolated situations, and on single issues. Usually, in the 1990s, it was about the land that government took away. Then it was about salaries, that employers were late in paying salaries, so there were protests about that – very single - issue, very focused. This time around, you're talking about people demanding the CCP to step down, demanding Xi Jinping to step down. That's just something entirely different from what we saw before.

TC: In your book, you write of what you call Tullock's curse – Gordon Tullock having been my former colleague – namely, embedded succession conflict in an autocracy. Why has Chinese succession been so stable up to now? And will we see Tullock's

curse whenever Xi steps down, passes on, whatever happens there?

YH: I do want to modify the word that you use, *stable*. There are two ways to use that term. One is to describe the succession process itself. If that's the situation we're trying to describe, it is not stable at all. If you look at the entire history of the PRC, there have been so many succession plans that failed, and at a catastrophic level. One potential successor was persecuted to death. Another fled and died in a plane crash. Others were unceremoniously dismissed, and one was put under house arrest for almost 15 years, and he died –

TC: But no civil war, right?

YH: Yes, that's right. There's another way to talk about stability, which is stability at the system level, and that, you are absolutely right. Despite all these problems with these successions, the system as a whole has remained stable. The CCP is in power. There's no coup, and there were not even demonstrations on the street associated with the succession failures. So, we do need to distinguish between these two kinds of stability. By one criterion, it was not stable. By the other criterion, it is quite stable.

The reason for that is, I think – although it's a little bit difficult to generalise because we don't really have many data points – one reason is the charisma power of individual leaders, Mao and Xiaoping. These were founding fathers of the PRC, of the CCP, and they had the prestige and – using Max Weber's term – charisma, that they could do whatever they wanted while being able to contain the spill-over effects of their mistakes. The big uncertain issue now is whether Xi Jinping has that kind of charisma to contain future spill-over effects of succession failure.

This is a remarkable statistic: Since 1976, there have been six leaders of the CCP. Of these six leaders, five of them were managed either by Mao or by Deng Xiaoping. Essentially, the vast majority of the successions were handled by these two giants who had oversized charisma, oversized prestige and unshakeable political capital. Now we have one leader who doesn't really have that. He relies mostly on formal power, and that's why he has accumulated so many titles, whereas he's making similar succession errors as the previous two leaders.

Obviously, we don't know – because he hasn't chosen a successor – we don't really know what will happen if he chooses a successor. But my bet is that the ability to contain the spill-over effect is going to be less, rather than more, down the road, because Xi Jinping does not match, even in a remote sense, the charisma and the prestige of Mao Zedong and Deng Xiaoping. There's no match there.

TC: In Chinese history, if we look at the years 220 AD to 581 AD, why is that your favourite period in Chinese history?

YH: Some people may say that's my European bias. That period was quite similar to Europe after the Western Roman Empire collapsed. China at that time was more of a federation of states rather than a unified empire. It had multiple governments rapidly replacing each other or simultaneously existing in parallel with each other. There was a lot of human capital moving around. The intellectual environment was quite free. There was not one dominant ideology lordling over other ideas and other ideologies. Confucianism was powerful, but it was first among equals rather than a monopoly ideology. In fact, there were many intellectuals at the time who openly challenged the authenticity and legitimacy of a Confucianist ideology.

It was also a period of enormous creativity in terms of poetry, in terms of humanities and, crucially, in our measure, in terms of technological creativity. China reached its peak in terms of inventions divided by population during that period.

TC: Why have there been so few female emperors in Chinese history?

YH: [laughs] Let me answer that question by relating my answer to your prior question. During the period we were talking about, Buddhism was a formidable ideology competing with Confucianism. Buddhism was actually quite friendly to the female half of the population. Formally, there were some female regents, but the only formalised emperor who was female was a Buddhist. Also, during that period of time, there were other East Asian states that had female rulers. Then Confucianism took control, weaponised by the civil service examination system. Confucianism was extremely hierarchical, extremely hostile towards women, and very rigid in terms of the social hierarchy, in terms of gender treatment. It marginalised the women.

You can actually see that according to art historians. You can see the evolution of Chinese art from painting women in relatively free style in the Tang Dynasty (618–90) to a very, very paternalistic portrayal of women beginning in tenth century, eleventh century. Then the civil service examination system chose one of the most chauvinistic versions of Confucianism as its curriculum, starting around fourteenth century or maybe even thirteenth century. Then that curriculum continued until the civil service examination system was itself dismantled in 1905.

I think it was the ideological hold of the stringent version of Confucianism that prevented the emergence of any liberalising forces, women

“

If you want to be a leader, you have to learn how to communicate with all sorts of people, not just with other people in your own discipline.

”

being one of them, but also liberalising other ideas. Going back to our earlier discussion, the civil service examination system contributed to the male dominance of Chinese politics in society.

TC: Where's the best food in China? Pick your favourite. I say Yunnan province but, of course, opinions differ.

YH: Yes, opinions differ on that. I like Yunnan food. My parents – one came from Hubei, one came from Hunan – and they eat very spicy food. But I grew up in Beijing. I didn't eat very spicy food, so Hunan food is a little bit too spicy for me. I like food in the Yangtze Delta area. It is a little bit gentle on your tongue. I also like Cantonese food. There's more variety. The problem that I have with spicy food is that it overwhelms every other sense that you have. It's a little bit homogeneous, whereas the Cantonese food is very rich. The Yangtze River Delta food is very, very rich. You can have spicy food, but you can also have salty food. I like variety. I like the food from that area.

TC: Culturally, why do you think that Chinese and Chinese Americans have done so much less well becoming top CEOs of American companies than Indians and Indian Americans?

YH: That's a running topic among Chinese American professionals like me and many others. I actually have a colleague, Jackson Liu, who has systematically

studied this issue, and his conclusion is not surprising, but his method is quite innovative. His conclusion is, basically, everything comes down to communication. Chinese Americans are less able to communicate their ideas, and place less value on communication. They are technically very competent, very capable, very accomplished academics. Something I have noticed among my friends who are engineers and who are scientists: they are very systematic when they are talking about chemistry, when they are talking about physics. But once you take them out from their discipline, they are actually not that systematic in the way that they describe the world, in the way they analyse the world.

If you ask them to analyse food, analyse sports, analyse politics, analyse international relations, analyse the economy, the way that they approach these topics is not that different from someone who has not gone through a PhD program in physics and PhD program in chemistry. Whereas, when I talk to my Indian friends who are in science and technology, they apply the same scientific mindset and methodology to social issues, to political issues. That gives you a lot of mileage in America, where you need to communicate with a broad spectrum of people, not just your fellow academics. Or maybe they are fellow academics, but they are not academics strictly in your discipline. If you want to be a leader, you have to learn how to communicate with all sorts of people, not just with other people in your own discipline.

My view is that we Chinese have a lot to learn in this regard. I think one reason is that we grew up in a homogeneous society and in a society that does not place very high premium on trying to convince others of your point of view. Xi Jinping doesn't need to score points on debate to get his policy preferences executed, right? He may execute people, but he doesn't need the communication skills to get his policy preferences through. He can just order. He can just issue commands. It's a command-driven society. We don't need to convince people who disagree with you to come to your point of view. We didn't grow up in that society. Whereas Indians operating in a democratic society, a very noisy, what Amartya Sen calls argumentative society – to get people behind you, you need to convince them of your point of view. To be able to do that, you need to rely on some framework that both sides rely on. That usually is logic, evidence and the systematic way of thinking. I think Indians are better than us Chinese in part because of that.

TC: Let's say you had an educated American friend. The person had been to Beijing, to Shanghai, and you were planning a two-week trip through China for them. Where would you send them?

YH: I would urge them to go to Xi'an, the city that has the famous terracotta sculptures. It's not just because of that. I think what's very interesting about Chinese political geography and economic geography is that the political geography is heavily northern. The economic geography is heavily southern. You think about Shanghai. You think about Guangzhou, Shenzhen. These are economic heavyweights. They are in the southern part of the country.

Beijing is in northern part of the country. It's a political capital. Xi'an used to be the capital of imperial China for many, many centuries. To understand the political mindset – and, by the way, Xi Jinping has very deep roots in that province – to understand the political mindset of China, you need to go to that part of the country. It is a part of the country that is heavily influenced by natural disasters, by flooding, by an arid, very harsh environment that doesn't cultivate agriculture very successfully.

It is maybe a little bit like Ohio or places like that – very inward-looking. No matter how modern the Chinese economy is, the politics is heavily coloured by that particular perspective. It is a little bit paranoid. It is suspiciously looking at the rest of the world. It also has this very fond memory of China many centuries before, when China was a unified empire that had that part of the country be the capital of the country. I think to understand Chinese ideology and political mindset, we need to go there.

TC: For the next week, where do you send them then? Let's say that's week one. You have two weeks. Where do they go? You buy the ticket.

YH: Then I will send them to Shenzhen, just as an incredible contrast. Shenzhen was basically created by economic reforms. It is the economic capital of private entrepreneurship, of innovations, and of the incredible supply chains that China has been able to create in the last 30 years. To understand the economic side of China, you need to go to Shenzhen, just next to Hong Kong.

I would argue that you go to Xi'an first to understand their political mentality, and then you go to Shenzhen. Then ask yourself the question, can these two things co-exist with each other for a long period of time? Under what conditions can they co-exist with each other for a sustainable period of time? My own view is that these two things cannot co-exist with each other for a long period of time. One of them has to give.

TC: Do you have a prediction?

YH: I have the prediction that Xi'an is going to give, and I think the economic side is going to win rather than the political side. It's not automatic. It's not the view that economics automatically advances political

progress. It is also because of a lack of political progress, the regime tends to make mistakes. It is those mistakes that will have a bigger educational effect on the Chinese middle class, on university students in how they think about whether or not the current political system is viable and desirable.

TC: I have three final questions, all about you. First, how was it that you decided to come to the United States?

YH: For young people in China, even today, despite the tensions between the US and China, the US is always the dream country to come to. I never really asked myself that question. Of course, if you want to get educated, if you want to achieve, accomplish something, you try to go to the United States. A lot of my friends have decided to go back to China because they see commercial opportunities there. They typically get a Western education and then go back to China to start businesses, and they have been extremely successful. I have chosen to stay because I'm an academic, and I don't think China is the right country for an academic, not for a social science academic. Maybe if you're a scientist you get support from the government, but not as a social scientist.

TC: Second question: what did you learn from János Kornai?

YH: János Kornai taught generations of not just Chinese students, but many students from European countries, from Russia, that there is a basic illogic with socialism. Before János Kornai, we tended to think about the socialist system in vague, general, and sometimes ideological terms. János Kornai taught us that it is rooted in the system. The investment hunger, the self-budget constraints. It really gave us a new way of thinking about the socialist system and central planning system in that particular perspective. That really was revolutionary because you don't have to be an ideologue to be critical of central planning. You can be a good empiricist and good system thinker to be critical of central planning. We didn't have that kind of language until János Kornai provided that language. The investment hunger, the soft budget constraints, and how soft budget constraints are being used not just by people who study central planning, but also by people who study other types of economic systems. That was remarkable.

A lot of Chinese students think China was different from other centrally planned economies in terms of soft budget constraints. China seemed to have harder budget constraints as compared with the Soviet Union, and there were a lot of debates about why this was the case. That led us to look at the history of reforms in China, look at the Great Leap Forward,

look at the decentralised organisation of the economy, even before the economic reforms in 1978.

TC: I'll again mention your forthcoming book, *The Rise and Fall of the East: Examination, Autocracy, Stability, and Technology*, which I found one of the most interesting books on China. To close, just tell us, what do you plan to do next?

YH: I am collaborating with a number of professors – some of them are based in China – on a book project looking at the history of Chinese technology, using the dataset that I already used for this forthcoming book. But we are going to devote the entire book to this topic. The title of the book is *The Needham Question*. Joseph Needham famously asked the question in 1969, how come China failed to take off and have its own industrial revolution, even though it had very advanced technology? We're trying to answer that question using a database that we have constructed. I just learned from Princeton University Press that they're going to award a contract to us, and we already have three chapters finished. We hope to finish the book by the end of this year.

TC: Congratulations, Yasheng Huang. Thank you very much.

YH: Thank you, Tyler. This is such a wonderful conversation. Thank you very much.



Conversations with Tyler

Published in 2023 and transcribed for The Podcast Reader.

 Read more @ podread.org

 Listen @ Conversations with Tyler

FROM DENIED TO DELICIOUS

imperfect fruit
transformed into
perfect juice



PLANET
FRIENDLY



   Tag us:
@flawsomedrinks
www.flawsomedrinks.com

A Conversation for the Curious

Shining a light on underexplored ideas

ALAN RUSBRIDGER
BETTER KNOWN, 2022

Interview by Ivan Wise
Illustration by Vaughan Mossop

Ivan Wise: Welcome to *Better Known*, where each episode a guest makes a series of recommendations of things which they think should be better known. Our recommendations include interesting people, places, objects, stories, experiences and ideas which our guest feels haven't had the exposure that they deserve. The only conditions for discussion are that our guest loves it and thinks it merits your attention as well. This week's guest on *Better Known* is Alan Rusbridger.

Alan you were editor of *The Guardian* for 20 years before becoming principal of Lady Margaret Hall in Oxford. In what ways did you find academia had aspects in common with journalism?

Alan Rusbridger: Both journalism and academia, I suppose you could say, are engaged in some sort of search for truths. There's quite a lot of overlap in what we do. At its best, you find academics who write brilliantly and are keen on communication and journalists who are interested in delving deeply into subjects in the way that academics do. And, of course, there are differences. Some academics, I'm afraid, look down their nose at journalists, maybe the same is true in reverse. Some academics don't really want to communicate widely, and some journalists are a bit slipshod and could learn some of the rigour from academics, so they're different worlds, but they have quite a lot in common.

IW: And did you find there were people who were able to write well both in journalistic and academic styles?

AR: Yes, but it's not as common as you might hope for. I mean, there is a certain sort of academic writing which is really just writing for other academics. I suppose I'm not alone in finding that a bit of a turn off, but there is crossover. But I dare say both sides look and view each other with a bit of suspicion.



“

His piano school then got bombed, not once, but three times by the Germans during the war, which kind of knocked him out of business and eventually he went out of fashion.

”

IW: So, you've chosen six things to discuss. Your first choice is bone-conducting headphones. So, what are these exactly and how do they work?

AR: I just started cycling again in London, after a pause of about 15 years. I had fallen off my bike and thought, I'm never going to do that again. And then I did, and somebody mentioned bone-conducting headphones. They are headphones that go over the back of your head and rest on your cheekbones. And, astonishingly, you can hear through your cheekbones, which is something that had never occurred to me that I would want to do or could do. I find them much better than normal headphones. I don't think I'd bother to listen to a Mahler symphony on them, but if you're just listening to podcasts, they're fantastic. They sort of vibrate through your sinuses. God knows how they work, but they're kind of life-changing in a tiny way.

IW: What are the benefits? What are you getting from bone-conducting headphones that you weren't getting from others?

AR: Well, you have spatial awareness. If there's traffic around you, you're not completely blocking yourself off, as I think a lot of people do with modern headphones, especially the noise-cancelling ones. I've become a bit of an advocate for them. When people try them a look of astonishment comes over their faces because I don't think it occurs to anybody that you can listen through your cheekbones. Having done it, I can't imagine ever wanting to go back to ordinary headphones.

IW: So, you encourage friends and family to take these up. How successful have you been in convincing people to give it a go?

AR: Well everybody I've shown them to has been tremendously enthusiastic and said I didn't know these things existed and I'm now going to go and buy them. I haven't followed up to check if they all did, but I haven't heard anybody in my sample size of thirty who said 'That's horrible.' They've all said, 'That's astonishing.' So that's why it's number one on my list of things that should be better known.

IW: So, bone-conducting headphones should be better known. Your second choice is audio sleep masks. Lots of people have trouble sleeping because of things like background noise and light. Is it something you've suffered from over the years?

AR: Yes, I think it's when I started editing *The Guardian*, and that responsibility of having a 24-hour publication and the thought that at any time of the day or night, something might be happening that would require your attention. I could never find a way of calming my mind down or switching my mind off. In a way, it's a sort of segue from the previous conversation because I, for a long time, used to listen to the BBC World Service, and that's quite frustrating because it's sometimes interesting.

IW: So, you didn't get to sleep...

AR: Whatever you're listening to can't be totally boring and it can't be too interesting. If it's too boring, you get frustrated. If it's too interesting, you wake up. I had a little earpiece in my ear in order not to disturb my wife. The perfect thing that I can recommend is Edward Gibbon's *The History of the Decline and Fall of the Roman Empire*.

IW: A very long book.

AR: Very long book. It takes about 120 hours to read. And the recording I've got, they read all the footnotes as well, it's certainly not boring. But, on the other hand, in a way, I couldn't be less interested. You fall asleep and it's 200 BC, and then you wake up and it's 200 AD.

IW: Is it one person reading the whole thing?

AR: Yes, it's an astonishing feat. Then I discovered a sort of velvet eye mask that goes over your eyes with headphones on each side, bluetooth headphones. And you are cocooned in the world of Edward Gibbon, but also in the complete dark. And I sleep like a baby. It's completely miraculous.

IW: Have you tried lots of other sleep aids before without sleeping?

AR: I've tried everything. Headphones you have to have long bits of wire and the wire all gets tied up. Earpieces. I've tried bluetooth headphones that run out at 3.00 in the morning. I've tried fancy ones that go under the pillow, but they don't really work. With audio sleep masks, it's just enough to distract you that I'm not thinking about whatever was keeping me awake. And if that happens to be some Roman Emperor that you've never heard of and will never find terribly interesting, then that's just the right level of the right balance between boredom and interest to help me nod off.

IW: I feel Roman emperors are not getting a good rep.

AR: I find I couldn't be less interested in them. I know people have devoted their lives to studying them, but one is very much like another. They all, as far as I can remember – because it doesn't stay in my brain – come to sticky ends. Murdered by their wives or daughters. I wouldn't try nineteenth-century history – far too interesting. Or twentieth-century history. I can no longer listen to the radio because I get so cross. But *The History of the Decline and Fall of the Roman Empire* is a great sleep aid.

IW: Excellent. And have you been evangelical about these with others? Have other people taken this up?

AR: I think it's more intimate, sleep, than my bone-conducting headphones. It's not often you get in discussions with friends and even strangers about how well they're sleeping, but I can heartily recommend them to your audience.

IW: So, audio sleep masks should be better known. Your third choice is the music of Billy Mayerl. He was a pianist and composer in the middle of the twentieth century. How did you first come across his music?

AR: Of all places, I was in Kalamazoo at a big piano festival, and there was a British pianist called David Owen Norris, who'd just been nominated for a very esteemed prize. He played his program, and then at the end of it he played pieces which I just couldn't place. They were jazzy, but not straight jazz. I wondered if it was Stravinsky in his jazz phase, or it could have been Ravel, but not quite. At the end, I asked, 'What was that?' And he said, 'Well, it's Billy Mayerl.' I was at *The Guardian* the time, and I went in search for more Billy Mayerl. There was a pianist called Eric Parkin who had just released a disc. Eventually I ended up in a warehouse off the North Circular near the M25, which was the only repository

of Billy Mayerl's music, which was very sad because it turned out that he was, at one point, the most famous musician in the country. We're talking about the late 1930s, 1940s.

He'd been born in 1902, was a prodigious pianist, played the Greek piano concerto at the age of twelve in the Queen's Hall. But he rebelled against classical music and started playing his own jazz compositions and when a jazz band came from America to play the first performance of *Rhapsody in Blue* in this country, they found him playing in a hotel in Southampton. Their pianist was sick and they said, 'Will you come and play the first performance of Gershwin's *Rhapsody in Blue*?' Which he did.

IW: How much notice did he get?

AR: About a week, I think. There are a lot of recordings of him playing and he had a prodigious technique and then went on to form a piano school, the Billy Mayerl Piano School. He had a weekly show on the Light programme on the BBC, so he was very, very famous. His piano school then got bombed, not once, but three times by the Germans during the war, which kind of knocked him out of business and eventually he went out of fashion. At the time that I wrote that piece, there were still some fans. Richard Rodney Bennett was a great fan, the composer, and he's just beginning, I think, to come back into fashion. There are one or two pianists, pianist Susan Tomes I heard play a recital of his works at the Wigmore Hall. I think it's time for him to be rediscovered.

IW: And his style of music, was he doing something different to other people?

AR: Well, as I say, you couldn't call it pure jazz and all those are written down. I think David O. Norris said he felt, really, it was improvised music that he then scored, but the harmonics you could listen to and they are sort of slightly Stravinskyian, slightly Ravellian. It's more sophisticated than pure jazz was at that time, and so it's quite difficult to place, but it's immediately catching and harmonically interesting and melodically, immediately whistable and it's just music that makes you feel happy.

IW: Is it music that you play yourself on the piano?

AR: The great puzzle about the Billy Mayerl School of Music is although he evidently had hundreds of thousands of people on his mailing list, it's almost impossible to play. I mean, I can't imagine what all these people thought. There's not much of it left behind because the Germans kept bombing it during the war. So I don't know what the literature was like that you got when you joined the Billy Mayerl School

of Music, but the pieces themselves... There's a very famous piece called Marigold that my mother listened to during the war and had the sheet music for, that's just about playable. There's a lovely piece called Sleepy Piano, which is also just about playable. I recall Richard Rodney Bennett telling me he just didn't have the stamina, he had the technique, but he didn't have the stamina to play. So, I don't think it's ever going to catch on as a sort of performance thing, but it's lovely music to listen to.

IW: Is he one of the twentieth century's biggest heroes to zeroes in terms of being very popular and then just falling away to complete obscurity?

AR: Mayerl died in about 1952, I think, and he did Desert Island Discs. But already, even by the 1950s, he sounded like a sort of old duffer, was drinking too much, and we were on the verge of Elvis Presley and rock 'n' roll. I think it's just one of those things. And then nobody listens to it. It seems ridiculous and old fashioned, and it's the music our parents listen to, or our grandparents listen to. Then, hey presto, 70 years later, you think, actually, that's quite good.

IW: So, is there a route for him to come back, do you think?

AR: He is played quite often now on Radio 3. I mean, a lot of it is not very good. I think he went off in his later years, but the early stuff, and especially if you hear him playing it... There's a piece called Railroad Rhythm, which he plays himself, and you just listen to it in wonder, you think how, how can anybody? Bert Raylton, the conductor who met him in Southampton, is said to have gone up to him and said, 'You play nippy, kid,' and he did play nippy. I think he then expected everyone else to play nippy, but they couldn't. I can't.

IW: So, the music of Billy Mayerl should be better known. So far, we've had bone conducting headphones, audio sleep masks and the music of Billy Mayerl. We've talked very positively so far. But as well as things which should be better known, is there anything or anyone really famous that you wish was much less well known?

AR: At the moment, I wish we knew less about Elon Musk. He's obviously an extraordinary man, he's done three amazing things. He built Tesla and beat all the other manufacturers to electric cars. Then he managed to put a reusable spaceship into space. But now he's gone and bought Twitter and it's clear that he doesn't know what he's doing. It's not even clear he wanted to buy Twitter, but he ends up having to buy it. He's one of these American libertarians who believe in no kind of regulation on speech. Twitter,

as a space has its problems, but I've got a horrible feeling that Musk is about to ruin it. It's beginning to become a sort of Wild West of redneck Americans who just want to have the freedom to be as rude and as obnoxious as possible. Watching him stumbling around discovering that free speech is quite a difficult thing, and regulating it is quite a difficult thing, and changing his mind every five minutes. I think it would be quite nice if he was less well known.

IW: Do you think that's likely to happen?

AR: He's that kind of genius that does what the legacy businesses don't do. Credit to him for building an electric car, showing that it could be done and doing it for a mass market. But, of course, now all the other motor manufacturers have got their act together and it's not clear that Tesla will still be a business in 10 or 20 years' time. So, I think to that extent, he may be remembered for being a brilliant disruptor, but whether his reputation lasts anything more than that, I don't know.

IW: We'll try and hear a little bit less about Elon Musk. Your fourth choice is electric bikes. So, you mentioned that you got back into cycling. Did you always enjoy it before?

AR: Well, I did until I fell off my bike. It slipped out from underneath me while I was going along Kentish Town High Street. It was one of these foldable bikes, small wheels, and I slid over on a drain cover and it could have been very nasty. There was a bus right behind me which managed to stop. So I just gave up cycling for a long time. Then a couple of people said they'd got electric bikes, and I went and bought one. London is a much pleasanter place to cycle around than it was 15 years ago. It feels quite a safe place to cycle, and electric bikes are extraordinary. I mean, I don't know what I was expecting. You have to pedal – it's not like you just switch it on and it takes you home, it doesn't. So, you pedal, but it's like having a sort of giant with his hand on the small of your back as you pedal. It feels like somebody's pushing you along, and the harder you pedal, the easier it gets.

So, I cycle now, it's about four miles in the morning, four miles going back. It's slightly downhill all the way in the morning and slightly uphill all the way in the evening, and there's no way that I would do that on an ordinary pushback. I zip all over London now. It just makes London feel completely different and everything feels within reach. I think it's moderate exercise, but it's a lovely way of getting around. I completely love it.

IW: London is obviously becoming a better place to cycle, but the actual bikes themselves, which used



to be expensive and bulky and inconvenient, they've become much more user friendly.

AR: I guess they have. I've got an office in Westminster and I did try those Santander bikes, which are like tanks. Even getting up a modest hill to Piccadilly from Westminster, you're not panting, but you're cycling so slowly that it is probably not very safe. Whereas on my electric bike I'll zip up. I can get from Kentish Town to Westminster almost entirely on cycle lanes. It's an irritating city to drive around sometimes now, but it's absolutely brilliant for getting around by bike.

IW: What sort of speed are you doing?

AR: I think there are limiters. I think by law you can't go more than 15, 20 miles an hour, you really get up speed on a nice bit of flat and then you can feel a sort of slight tow factor. And, of course, I have my bone-conducting headphones. I nearly chose a brilliant app called Citymapper. 'Choose me a quiet route.' You put your bone-conducting headphones on and it just says, turn right, turn left, and it takes you through all the byways. I've never had so much fun travelling around London.

IW: Presumably there's a whole range of electric bikes. Are they relatively good value for what you can get now?

AR: Well, Ed Miliband was one of the people who told me that he just started, and he said he got his for 800 quid. I think I spent double that. But on the other hand, it, you know, it's free transport thereafter and it's good fun.

IW: So electric bikes should be better known. Your next choice is why free speech matters. We live in an era where free speech is under threat, but many people seem to feel that there's nothing that can be done about it. So why does free speech matter to you?

AR: Well, it seems odd, doesn't it, to say that free speech should be better known, or the importance of free speech should be better known. But I was surprised when I got to Oxford. We didn't have any big run-ins in my college, and I was a bit sceptical about the whole complaint about snowflakes and so forth, but there definitely was a feeling that people wanted to be protected from speech that they didn't like and they wanted to feel safe. I got on very well with the students, and I taught some of them. I'd ask, 'Have you ever learned anything about free speech or how free speech came about?' And they hadn't. It strikes me as a big gap in education that people don't understand. I mean, it was obviously important to me when I was editing for 20 years. People were always trying to stop you from publishing stuff. It's

completely central to the current debates around social media and public protests.

I think this government is very illiberal in trying to silence aspects of the press, trying to silence aspects of protest. The Online Safety Bill, which the government tried to pass, which tried to outlaw things that were what they called legal but harmful, speech that you don't like, that might cause offense, might cause harm. But how do you define harm? So, I think a lot of the freedoms that have been won over three or four hundred years in this country around speech are under threat. And we have a younger generation, some of whom want protecting from speech they don't like. That's why I added it to the list of things that I think ought to be better known, and I think probably ought to be properly taught about more in school.

IW: Did you find that the issue was generational and linked to the idea that people took free speech for granted, so they didn't maybe value it in the same way that others did?

AR: Yes, I did. As I say, I was sceptical about it when I first went to there. What I used to say to the students is, 'You're supposedly the brightest of your generation. If you can't defeat these arguments and want protection from them, then what hope is there for the rest of us?' So, I did my best to train them up with that in mind. Also, the feeling that once they left Oxford, there was no protection, there were no safe spaces from speech. So it's a very bad habit to get into, the need to be protected from speech. I think you need to take it on. Of course, it's immensely complicated by digital media and the proliferation of trolling and hateful speech and so on and so forth. But I welcome the fact that the government seems to have gone into retreat over this legal but harmful aspect. I also sit on the Facebook oversight board, which is trying to help Facebook think through content moderation and their policies over hateful speech, nudity, blasphemy, political speech, violent speech. I think it's so current. But the starting point, I think, in all these things has to be an innate presumption of free speech and that free speech is a good thing. And I'm not sure that people necessarily start from that presumption nowadays.

IW: I guess you would make the case that you only defeat negative, cruel or bad ideas by exposure and by arguing against them? If you cancel or ignore them, they don't go away?

AR: I do think that, and I think some people don't think that any longer. The John Stuart Mill best answer to speech is more speech. I know there are sophisticated arguments about the balance of power and who has the power and who has the megaphone. But in the world of digital, it's a relatively level

“

But I think in any good magazine, you need to find light and shade and find a way of writing about ideas through people, as well as just the ideas themselves.

”

playing field compared with any other time in history. You haven't got privileged people on the platform who own the printing presses and with no way of responding to them. I guess with each generation you have to rethink the framing and geography of speech. As I say, I would teach media literacy and speech and argument. I mean, all these things I think, ought to be more ingrained in the curriculum in schools.

IW: So why free speech matters should be better known?

AR: Yes, certainly should.

IW: Your final choice is *Prospect Magazine*, a current affairs magazine which you edit. It's been going since 1995. What attracted you to the magazine?

AR: There are two things that I think the media sometimes suffer from. One is it just gets faster and faster and faster and faster. So, we're all looking at our mobile phones thirty, fifty times a day for the latest headlines. It's quite nice, I think, having a monthly magazine that is the opposite. It's very slow and considered and has lots of context, and tells you where stories came from and where they might be going next. The second thing is, I think the modern world is terribly polarised, with people taking up positions and living in echo chambers. When I took over the magazine about a year ago, I thought, well, I would like *Prospect* to be read by anybody, left or

right. I've only written one piece for it where I said I'm not going to write an editorial telling you what to think. Make your own minds up. We will have people from all bits of the political spectrum. I've introduced a feature which is a conversation on what people agree about rather than what they disagree about. So, we've had extinction rebellion, people in the same room as oil executives, or George Monbiot, the noted environmentalist who believes that farmers are all terrible people, in the same room as a farmer. We're about to do one with Brexit and bring the two sides together instead of throwing tomatoes at each other. Where should this now go? So that's what we're trying to do with it. Almost literally nobody knows about *Prospect*. It's got 6 per cent brand recognition, I was told when I walked through the door. So, it definitely should be better known.

IW: Obviously that deliberate tactic of trying to hear both sides of the argument is pretty unusual and probably goes against the grain of trying to get more brand recognition, but it means that you're bringing something different to the debate. People don't pick up the magazine expecting to hear a point of view they already agree or disagree with.

AR: Yes, I think people would find it quite a relief. We know that social media algorithms are skewed to disagreement and controversy and sensationalism, because that drives traffic and keeps people there and keeps them angry. I think if you pick up a copy of *Prospect*, it will soothe you and you won't feel angry, but you'll definitely be better informed and maybe slightly amused.

IW: Its topics, I think, are more varied than people might think. So obviously there's a focus on current affairs, but recent articles I've come across, there was one on great unknown female artists. One by Mike Brearley on the England cricket team, one about 50 years since Watergate, so it's quite an interesting mix. Is that important to you?

AR: We've got Mike Brearley writing every other month. Sheila Hancock writes every month. I've got a farmer, I've got a woman priest, I've got an asylum seeker, who writes every month about what it's like living on £5.84 a day. A psychiatrist who is herself bipolar. It is primarily a magazine about ideas. But I think in any good magazine, you need to find light and shade and find a way of writing about ideas through people, as well as just the ideas themselves.

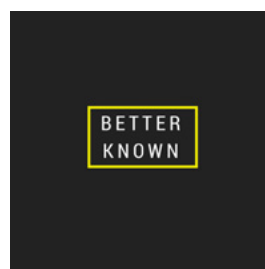
IW: Although, of course, it should be better known, presumably, if it was really well known, it would be harder to do some of that? Perhaps by having a relatively smaller audience you can do that more?

AR: No, I'd love to have bigger audiences. This is terribly self-serving of me, but I clearly want as many readers as possible, but I think there's an appetite for it. It's just marketing these things is expensive. I think it's a well-kept secret and definitely should be less of a secret.

IW: So today we've had bone-conducting headphones, audio sleep masks, the music of Billy Mayerl, electric bikes, why free speech matters and *Prospect Magazine*. So, out of your six choices, Alan, which one do you feel most strongly should be better known?


AR: Well, I'd have to say *Prospect Magazine*. The proprietor would thank me.

IW: Thank you very much to Alan Rusbridger for your choices.



Better Known

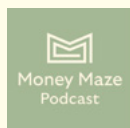
*Published in 2022 and transcribed for
The Podcast Reader.*

 [Read more @ podread.org](https://podread.org)

 [Listen @ Better Known](https://www.betterknown.co.uk)



On the history and causes of banking crisis



HUW VAN STEENIS
MONEY MAZE PODCAST, 2023

Interview by Simon Brewer

Huw van Steenis: History doesn't repeat itself, but human behaviour does. At the end of the day, this does remind you of some past banking crises where you had a prolonged period of stability, stability bred complacency, a huge shock in interest rates, and a couple of people who are reaching for the top shelf have toppled back over. Part of this is the sharpest increase in interest rates in our lifetime. But also, it's working through the excesses of the pandemic. So, to just put some numbers on it, in the States between 2020 and 2022, there were 5.2 trillion deposits put into the US banking system, only 2 trillion of which were insured. Now, loan growth as we remember was really tepid, at about \$700 billion of loan growth. So, what do you do with the rest? Well, banks put about \$1.9 trillion on deposit with the Fed, but bought two-and-a-trillion quarter of securities. And, obviously, the foolish ones bought long-dated securities. If we come on to Silicon Bank, effectively, it was a really dumb rates trade which went sour.

Simon Brewer: If we reference the history books and look back at the US savings and loans crisis, I believe between 1986 and 1995, a third of those 3000 institutions went bust. After that and following the great financial crisis of 2008, regulators made changes to establish more resilience. But under Trump, there was pushback. Is this where the seeds of this crisis were sown?

Huw van Steenis: I think it's good to go back to that early eighties, and probably even dial back a bit earlier. But there's a great book by Will Black called *The Best Way to Rob a Bank Is to Own One*. He was a bank regulator, and there are some phenomenal insights

in it. He said what made that debacle worse were the three Ds: deregulation, desupervision, and then, the decriminalisation of white-collar crime. Deregulation, desupervision was on show here in 2018. The Trump reforms, which were supported bipartisan, rolled back what was a systemic bank from \$50 billion to \$250 billion. In the Eurozone, a systemic bank is cut off at €30 billion. SVB was the sixteenth largest bank, and yet wasn't subject to Basel stress tests nor to liquidity coverage ratios. And so not being in the eye of the supervisor, they had some poor practices.

SB: Which, of course, allows us then to move over to Europe that has had, one might argue, better scrutiny and regulation. We have Credit Suisse, a poorly run institution for a long time. But suddenly, any remaining trust evaporates. Reminds me of Ernest Hemingway's quotation, 'How did you go bankrupt?' 'Two ways,' he answered. 'Gradually and then suddenly.' So, tell me, Huw, what goes wrong?

HS: It's tricky, because I think at one level, you say there's not much in common between the banks in the States and Credit Suisse. But if you unpick it, there probably is. In all the cases of the four banks which have failed, they were the banks of the billionaires and the VCs. What we've now found is these are hyper-connected people. They're all sitting on social media. They're all frenzied by what's going on on social media and whip their money out. I think that billionaire movement is something which regulators will have to come back and look at. These uninsured deposits moved very fast. And throughout history, it's the uninsured depositors who run. Even in 1984, with the largest bank crisis of that stage, Continental Illinois, it was foreign depositors who ran.

SB: Let's just stay with the concept of deposit and deposit protection, because there are siren calls that depositors shouldn't lose money under any circumstances. You have written in the *Financial Times* that banks need to be more sensitive to the threat of deposit flight and to pay up for funding, and that that will lead to tightening financial conditions. I wonder whether you could explain that.

HS: I think we've just seen the sharpest tightening of financial conditions in history following the fastest bank run in human history and the fastest increase in interest rates in our lifetimes. I think that's partly a function of this digital bank run. In the States now, if you're a mid-cap bank, you're really thinking about how many loans you want to give, what's the right price you want to give them, because you're not quite sure of the stability of your funding base. I think, in contrast, some of the bigger banks are actually benefiting. They're actually getting excess deposits, hundreds of billions of excess deposits, no doubt, but they are equally going to worry, are these easy come, easy go deposits? So, they may not want to make loans either. So I think the financial conditions will sharpen tightly. It comes down to the point of how sticky are those deposits, and then, what you do with it. If you've got daily liquidity money, you don't put it in 10-year illiquid securities, which is what the management of SVB were doing.

SB: I just want to stay with this sense of the stronger getting even stronger. Is there almost now an oligopoly of the super banks? And from that might flow super profits because you're worried about where you go, therefore, your choice has diminished to a very small number.

HS: Yes, but I think we were already creating a super league of banks anyway after the last financial crisis. Take the top US banks. Pre-financial crisis, they used to represent about 20 per cent of industry profits. They now represent over 60 per cent of industry profits. So, I think that partly is a function of the regulation giving them a moat, and therefore, they're viewed to be safe and secure. But I think there's something else, which is around the role of technology. The more that technology becomes an important source of competitive advantage for a bank, how you serve your clients, the more that 'winner takes most' characteristic you see in tech is starting to come to banks. So the profitability of the top three banks is way higher than even the next three or the three after that. So I think that 'winner takes most' is playing out in banking and I think that has really helped the super league in the States.

SB: I am older than you. When I started in the industry, banks were if not utilities, they were viewed through a lens of being rather dull. But now it seems that we have moved into a different environment where super profits might be available. I don't want to go from the money maze to the moral maze, but we're in a situation where the profits are being privatised but the risks are being socialised.

HS: I'm not sure I'd go quite that far. With the reforms

“

I think we've just seen the sharpest tightening of financial conditions in history following the fastest bank run in human history and the fastest increase in interest rates in our lifetimes.

”

of the last 15 years, a huge amount of the risk-taking activity has moved out to the banks and into the private equity world, hedge fund world, to specialist players. If you think about the value creation in the last five years, it's been in people on your show, the KKR's and Blackstones, the stock exchange groups, and so forth. The fact we've got a pluralistic system, you can go to get money from a bank or BlackRock or a private equity firm is obviously very helpful. What that means then is that what's left in the banks is a lower-margin business. Of course, that means you do need to scale. And I think that's the challenge. The more scale you need, the super league, the top three, top five, disproportionately benefit. So, in a way, the scale is also a function of public policy rather than just because of greedy bankers.

SB: If you were to look at your crystal ball and think about the next few years, we've got a tightening cycle that's still ongoing. We probably are headed into recession, I would say we are headed into recession in the US, it's just a question of how severe and its duration. Would you think that there are going to be a lot more air pockets within the financial sector as a result of that trajectory?

HS: Like you, I think we are moving to a recession. We can debate how long depending on where monetary policy goes. But what we've seen in the past is these initial problems turn into asset quality problems as people can't roll their debt. You mentioned the S&L crisis early on. One of the phrases there was a rolling loan gathers no bad debts. Obviously, if you can't roll a loan, then that's the issue. I think around parts of

private equity, parts of commercial real estate, and obviously, parts of the economy where low interest rates have allowed anyone to roll their debts, it will be gobsmacking if we don't see some asset quality problems. But going to this point, I think much more of the speculation around zero-interest rates came through in the US than it did in Europe or Japan. And therefore, I think a lot more of the focus needs to be on those areas.

SB: I'm struck that our conversation is focused on the US and Europe and it's left out that huge chunk, which is Asia. What are your observations about the Asian financial system? Big domestic institutions, population growth, etc. Are they learning anything from this and trying to do things differently? How should we be thinking about the opportunity set there as investors?

HS: Going back to 2008, I always find it fascinating that, for the West, a bank run was a surprise. However, if you take the IMF database since 1972, I think there's been 170-ish systemic banking crises around the world. We don't learn from the emerging markets. But in our lifetime, of course, we have had in Korea and Indonesia beyond major banking panics. As a result, FX mismatch is way lower. Liquidity management is better. So, I don't think they're seeing the same stresses that we are at the moment, although I still think this nature of a digital bank run is something we should worry about. I think the nature of the excesses and mismatches is always where we should focus. Obviously, China is in a different camp. There are some state-run banks there which obviously have got a series of bad debts which they're tromping through. But if I think about the private sector banks, whether it be the Singaporeans or the Australians, they look in somewhat better shape for the moment. But like everything in life, one needs to think about the individual bank, not just the system.

SB: On the issue of climate change, we had Nicolai Tangen as a guest on the show last year, a fantastic guest, and, you know him well, just an inspiring individual. You are working with Norges Bank on their climate advisory board. Tell us a little bit about what you and he are hoping to achieve.

HS: Nicolai and Norges Bank itself wants to be a world leader in thinking about investing in the energy transition around the world. I think there are three blocks to that. First is obviously understanding investment risks where there is energy transition or climate change. What does that mean for a portfolio? Particularly for fund, which is the largest investor in the world, they own 3 per cent of most European companies and about 1.4 per cent of every US

company. They're thinking about very long-dated risks. So, we want to understand what do twenty-first-century risks mean to companies they are a custodian of. Second is to seek out the opportunities. Who can be the biggest investor in climate tech? Who can be the biggest investor in the solutions and opportunities around this? And third, to be honest, I think this is where Nicolai and Carine are very inspirational is, can there be a standard setup? And this is the really tricky one. As a sovereign wealth fund, most SWFs (sovereign wealth funds), as you know, want to keep a little bit below the radar because of their political influence. They also want to be involved in standard setting bodies to enhance disclosure so that the investment community can make more informed decisions. So, I think it's a very inspiring mandate and he's very kindly invited myself and three others to join a climate advisory board. To think through not just the plan they've got, but how every year can they get that little bit shrewder around the investment risks and opportunities, and how can they improve disclosure standards in the market. All credit to Nicolai and team.

SB: Staying with interesting folks, Sandra Robertson, who is the CIO at Oxford University Endowment Management, who you know well. She wanted me to ask you, if you had \$100 billion to invest in the climate transition, how would you allocate it and what percentage return would you require?

HS: I think this is both a really important but actually quite difficult question. Where are the opportunities? First, climate tech climate solutions are incredibly exciting opportunity areas and they're very broad. This is from AI, to optimising how a grid is run, through to data standards and carbon capture. This is an area full of opportunity. It remains relatively capital light so you can earn returns. There's a lot of money trying to chase this, but I think that's the area where the climate solutions really have to be the standout and the category you want to put your money.

Second is what I think of as the industry leaders. How is McDonald's changing the agriculture value chain? How is Microsoft changing the tech value chain? These leaders can have disproportionate influence on emissions of not only themselves, but the entire industry. Again, getting behind these leaders and helping empower them to do that is very exciting, probably more in a public market equity portfolio. Now, here's the tricky bit. Ideally, we don't want to put a tonne of money behind renewables. Capital-intensive renewables have got a shockingly bad return on investment. If you want to go and invest around, I don't know, an offshore wind farm at the moment in Europe, you probably get 3 per cent to 4 per cent returns on capital. That's not what an



“

But what we've seen in the past is these initial problems turn into asset quality problems as people can't roll their debt.

”

endowment, a family office, a pension fund typically wants to have. It's probably more for an insurer. And that's the tragedy because that's where the money is needed. Because the challenge to national security from the Ukrainian war and beyond is that we now need more energy independence. So, there are plenty of subsidies going into renewables and I think we therefore need to think about pre-subsidy, that 3 per cent to 4 per cent just isn't enough. But can it be supercharged, or can you get behind the benefit of those subsidies would be a question I want to do the investment work on. But I think it's tough to put a lot of your \$100 billion into only getting 3 per cent, given inflation is still rampant.

And then last, the improvers. If you go back to the modelling of how emissions fall, half of the emissions falling is companies going from grey to green. It's getting Solvay to reduce its emissions and become cleaner. It is getting Volkswagen to go from dirty to green. I think the improvers have got a bit of a bad rap because, at the moment, that's where the emissions are. And I think what a lot of the NGO community have got wrong is it's not just about reducing financed emissions, it's about financing emissions reduction. We need to help those improvers get better and I think that grey-to-green opportunity is a huge one. I think that's one of the more exciting ones, both for private market opportunities as well as the public.

SB: If we could set up a dinner for you to sit with two of the great finance people or investors, from the present or the past, who would they be?

HS: That's a very good question. I haven't thought about that at all. So I am passionate about Keynes and what I've learned from reading him. I know it's

hideously obvious, but he nailed it in so many ways. Keynes was rich and thoughtful and interesting. Through Morgan Stanley, I was blessed to meet some of the most extraordinary minds. I used to go and have lunch with George Soros at his townhouse in London. I have to say, I'd love to carry on talking to George because I think his understanding of the macro and translating that through into trades and micro was extraordinary. And also, to think about risk appetite. When you've got hot hands, when do you really go for it and when do you not? Effectively, as a bank analyst, I'm looking at macro through a macro lens, and I think George, in some ways, personifies what makes a good bank investor.

SB: And what would you tell a 20-year-old Huw?

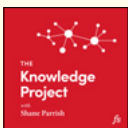
HS: I definitely would encourage him to read financial history. My good friend Niall Ferguson has a line that he wishes that more central bankers had read history rather than doing a PhD in economics, and I'm sure that's right. So, having a really good list of financial history, from I know what happened to Weimar Republic, Adam Ferguson's brilliant book *When Money Dies*, through to what happened more recently. So, one is financial history. Second, of course, get the technical skills. And third is just be voracious in listening to the signals. It can be just subscribing to *The Economist*. These days, you can get retail access to Bloomberg quite cheaply. That's phenomenal too. And obviously, of course, listen to your podcast.

SB: I didn't even prompt you. Thank you. Huw, this has been terrific. Thank you so much.

HS: Thanks, Simon. It was a real pleasure.

On technology, climate change and geoengineering

Photograph: Brecht Denil



NATHAN MYHRVOLD THE KNOWLEDGE PROJECT PODCAST, 2023

Interview by Shane Parrish

Shane Parrish: I want to start with your postdoc. You worked for Stephen Hawking. How did you end up working with him?

Nathan Myhrvold: Well, that wasn't a surprise, in a sense. I was a graduate student in physics. I got my PhD at Princeton. The kind of physics that I was interested in is the kind that Stephen did. It was about quantum theories of space and time. So, I applied to a

number of places ... and that was my top choice. And he said yes. That's how I got there.

SP: Where did you go after that?

NM: After that, I wound up taking a leave of absence from that position. I was supposed to be there for a couple years, and I was there for one year. It was a leave of absence working on a project that I'd started in graduate school with some friends. That project turned into a software company, and I became the CEO of the company. Then Microsoft bought the company and I joined Microsoft. Many, many years later, I retired from Microsoft and when that was announced, I got an email the next day from Stephen, saying, 'Shall we clear out the office? Are you coming back?'

SP: You also took a leave of absence from Microsoft, didn't you?

NM: Well, that was my second leave of absence in life, yes.

SP: That was to study ... cooking.

NM: Well, no, there were two leaves of absence at Microsoft. I did go to chef's school while I was at Microsoft. Bill hoped I would come to my senses after being bored at home, but I wasn't, so I never went back.

SP: What interested you in cooking?

NM: Well, I was interested in cooking long before I even knew there was a computer. I started, I think, when I was nine years old. I discovered cookbooks in the local library and I announced to my mother I was going to cook Thanksgiving dinner all by myself. So I got this arm's load of cookbooks, and I cooked Thanksgiving dinner. It wasn't that great, by my current standards, but it wasn't such a disaster that it turned me away from the topic. So I had always been very serious about food and cooking. And at one point, while I was working at Microsoft, I realised I had all these advanced degrees in topics I didn't actually spend that much time doing. And cooking is something I did spend time doing. Why don't I actually get some real education on that?

SP: So you took a leave of absence, and convinced the culinary institute to take you as a student?

NM: I had to convince the culinary institute in France to take me as a student. I had to convince Bill to give me a leave of absence. But it worked out on both sides.

SP: Take me back to Microsoft. You guys made a lot of decisions at Microsoft in a very fast-changing environment. How did you think about those decisions?

NM: Well, I think there are two things. One is that you try to be as analytical and as careful as you can about making a decision. But the second is that you also have to monitor how it's going and not be afraid to change your mind. The personal computer revolution was something that we take for granted today. Back then, not so much. At Microsoft, we used to have a slogan for the company: 'a computer on every desk and in every home.' And I took so much grief from people when I would say this in speeches, where people would say, 'I'm never going to have a computer in my home. That's just an absurd nerd fantasy.' Well, it worked.

SP: Now you have one in your pocket.

NM: Well, I actually wrote a memo about that in 1990 that you also would need one in every pocket, and ...

“

... on the day when
autonomous cars are
actually better than
humans at driving,
will we accept them?

”

we do have one in every pocket. And, by the way, our cars are basically computers with wheels. Obviously, you need engines and lots of other stuff, but a huge amount of computing goes into the automobile, into the navigation system, but even into how the engine operates – that's all a bunch of chips. When I was a kid, you could go to a speed shop and they would have all these fancy accessories that hot-rodders would put on. For a lot of cars these days, the simplest way to change performance is literally with a new chip.

SP: Where do you think we're going?

NM: Well, in the short run, you can make a whole lot of predictions that I think will turn out to be pretty close to correct. As you get further out, it gets much more difficult. In the 1950s, and actually up through the seventies, science fiction novels posited this idea of stellar travel and rockets and space. That was really cool. The moon landing got people inspired.

Even in Star Trek in both its original incarnation and the recent ones, they have a human driving the starship. How absurd! You'd never do that. There's no modern jet that doesn't have an autopilot, for God's sakes. And of course, that's because in that particular application, computers do a way better job. With self-driving cars, that's a little bit more difficult because we have traffic, and that traffic is other cars and it's kids running into the street and so forth. But if you're out there in two dimensions on a boat or three dimensions with an airplane, autopilots are great. Even auto-landing systems work pretty well these days.

SP: Why don't we have autonomous trains then? Because that's a closed network, closed loop, practically speaking.

NM: Well, many modern train systems are autonomous, for all intents and purposes. They will have humans on board as a safety or a backup feature. But when you have modern trains – by which I don't mean any train in the United States, I'm afraid – but if you have the high-speed trains that are going 300 kilometres an hour in Europe or, increasingly, in Asia, you can't rely on humans to make all of the decisions. We don't operate fast enough. We're not totally reliable. We can be distracted. We're great for some things. Don't get me wrong. But those systems are very much controlled with computers. The same thing is true for most modern subway systems. Again, automation is way safer than just doing it by human control. Now the humans are still involved, and that's for a variety of reasons. Some of them are good reasons; some of them are more holdovers from the past that we'll eventually get over. But for a whole set of things, why would you use a human?

SP: They are safer. But then we have to answer all these societal questions about who's at fault in an autonomous car?

NM: Well, at the moment, autonomous cars aren't really autonomous, and they don't really do as good a job as a human driver does because there are too many random things that can happen on the road that the autonomous systems have a hard time doing. But over time, that will change, and it wouldn't take very much of an improvement before they're better than we are.

So now you're asking a really important societal question, which is, on the day when autonomous cars are actually better than humans at driving, will we accept them? And for lots of things, there has been pushback. It's been true for a long time. Humans are the weakest part of a fighter jet. You should use drones. But the air force, which is composed entirely of pilots or former pilots, hate that idea. But drones are here to stay, and there will be aerial combat drones. Pilots are insanely good at what they do, but they can't take a hundred Gs, and missiles can. Plus, the other thing that is rude about human pilots is that you've got to call their mum when they die.

The manned space program is another absurd thing. It started off wonderfully that the only way we could explore space, for a certain class of things, was with humans. Well, it's long since passed. And today, the manned space flight is a reality TV show. In fact, the cutting edge of manned space flight is space tourism. That's what Blue Origin and SpaceX and others are gearing up for. There's just no reason

to take a fragile human up there.

The amount of cost and complexity that goes into keeping this fragile human alive in a context where it really shouldn't be is enormous. Mars exploration is a great example. People want to colonise Mars. Is Antarctica full? Antarctica is great compared to Mars. It's warmer. It has air. It has water. It's got all these great features. But to go to Mars, you have to be exposed to a certain amount of radiation from cosmic rays. It takes a really long time to get there. And there are lots of other things that are difficult. If you look at the history of Mars space probes, we've lost a ridiculous number of them. And you can say, 'Oh, well, that was a mistake here and a mistake there.' Well, yeah, because it's really complicated.

SP: Literally rocket science.

NM: It is literally rocket science. For the rovers and the drone that they have on Mars right now, they have the problem that you can't control them from Earth anymore because the amount of time it takes a signal, even travelling at the speed of light, to get to Mars, is minutes. If you were using that to drive a typical, radio-controlled car that way it would never work. So, they have to make those things semi-autonomous. And, over time, they'll become more and more autonomous.

SP: Do you see the need for interplanetary travel? Or would you focus on something else?

NM: Well, I'm a huge fan of science. I'm a huge fan of exploring the solar system and having both intra-solar system and interstellar probes. I think those are both ultimately fantastic things for us to do to explore the world. But it's vastly cheaper and easier and better to do it with machines than to do it with humans. In terms of a need to colonise the solar system, people believe that that's an imperative as a holdover from the ideas of the age of imperialism, when we thought we have to go conquer and colonise the whole world. But, at this stage, you have to say, 'Well, if we really need to leave this planet ... we must have screwed something up pretty badly.' Some people say, 'Well, that's why we should go to Mars, so that we have a second home for humanity.' Well, if you really had to do that, the moon is a lot closer and has a variety of advantages. Or you could build large habitats in space.

There was a physics professor named Gerard O'Neill at Princeton who was a big fan of this idea. But, so far, we don't have a need to do that. And it makes a lot more sense to try to take care of this planet and not screw it up than to treat Earth as being disposable. And so, 'It's okay, we'll have Mars.' Because if you keep thinking that way, you're going to screw Mars up, too.

SP: Where are we going wrong in taking care of this planet?

NM: Humans are not great at long-term things. So, we all know that we should eat right and work out. Lots of us don't. And that's because of the immediate gratification of 'I'm late for an interview this morning; I've got to leave. I don't have time to work out.' That was me this morning. Or 'Wow, that doughnut looks really great,' or 'There's an extra slice of pizza left.' We are such suckers for that. Which is why, if you take the standards of the past, the average American is overweight.

That's why we do a variety of self-destructive behaviours individually. As a society, we try to plan ahead a little more. And, for some things, we're able to do that. But if you take a problem like climate change, it's the worst-case situation. I like to compare climate change to ecological disasters. So in Spain, I don't know, 20 years ago, something like this, there was a flood. The flood washed over the banks of a river near where there was a gold mine. Gold mines use cyanide to help recover the gold from the dirt. So there's this cyanide spill into the river, and it killed a quarter of a million fish, which floated downriver and piled up in the city of Seville in Spain in August. Now, you've got a quarter of a million dead fish rotting in a river in the middle of hot August. Boy, you know you have a problem. And whenever we have a problem that is an ecological problem, that is localised in space, localised in time, and is easy to trace responsibility, we fix it.

Love Canal in the United States – there's a super fun site. Huge amounts of chemicals were dumped for many years by a company which made the ground horribly poisoned. It's a tragedy. It should never have happened. But it did happen, so we fixed it. We even fixed a slightly less localised issue with most air pollution in the United States, which is solved. It is not solved in Beijing. Beijing has horrible air. This summer in Seattle we've had horrible air because of wildfires. That's not a controllable thing and it is, in part, due to the climate thing. In part, it's due to forestry practices. It's a whole set of stuff. So as much as we can say we feel superior that Los Angeles used to look like Beijing in terms of air, and Los Angeles is way, way better, we did fix that. China will fix it, too, because eventually the people living in Beijing will say, 'Hey, this is screwed up. Let's fix this.'

The forest fire problem is more diffuse. It affects lots of things over a long period of time. The forestry practices that are mostly in government-controlled national forests are there in part because they valued timber and they wanted to stop all little fires. And it turns out, if you stop all little fires, you leave tonnes of fuel for a big fire. But you've got to change that mindset.

Now you get to climate change writ large, and you have the worst of all situations because, currently, we're doing the first order of approximation – exactly

nothing. You might say, 'What about all of the renewables? What about these other things?' Yes, we do have some very admirable work that's going on, but the trouble is that the world's also growing economically, which is a good thing, mostly. But as a result, the percentage of renewables, even if we grow renewables really fast, it's hard to grow them fast enough that it makes up for all the coal plants that are being built in China and India and other places whose economies are building. If you measure CO₂ up on the top of Maunakea in Hawaii – that's where the best measurements come from – every year it's higher. So, no, we made no progress.

But to make progress, we would have to invent a technology that was just totally superior to fossil fuels. Someday, we will. When fossil fuels came in, fossil fuels were superior to directly harvested fuels. So, coal pushed out wood. England used to be covered with trees. It's a barren tundra today. But if your fuel needs and your ship-building and house-building needs grow faster than trees can be replaced, you get a totally deforested land, which is the first approximation of what Britain is.

SP: So we have a long-duration problem. It's not right in front of us. What's really causing the problem?

Nm: There's another factor, which is jealousy and human nature. I think we understand what's causing the problem. It's a whole variety of things around modern life. Transportation – fossil fuels are great for transportation. The density and the convenience of a liquid hydrocarbon fuel are just enormous. We're making numerous electric vehicles, except today, most electric vehicles are filled up with fossil-derived electricity. If you drive a Tesla and you live on the east coast of the United States, statistically speaking, you're hurting the climate, not helping it. In Washington state, we can feel slightly better about it because we have a mix of nuclear and hydro. So, most of the power in Washington state is at least clean so if you charge your Tesla there, that helps.

But here's the worst part of it. Humans will typically put stuff off and put stuff off and put stuff off, and then we panic. Now we saw that with the pandemic. I had been publicly saying that we had to worry about both a natural pandemic or a bioterrorism thing. I've been saying that for 10 or 15 years. Other people have been saying it for longer than I have. In 1918, you had this horrible pandemic that killed 3 per cent of the world's population, something like that. That's a giant number. But no matter how much you pound the table and say, 'We need to take pandemics seriously, 1918 was a long time ago. And in the trade-off between our immediate gratification and the long term ... I had some very reasonable people who would argue with me on this pandemic issue. They would say,

'But look, I study heart disease' – or cancer or HIV or some other terrible disease – 'and we may have a pandemic at some point, but shouldn't we put our first priority towards diseases that kill people every day rather than disease or situations that occur once every hundred years?' Well, you have a mix.

The trouble with the pandemic is that if it happens, even though the probability of a pandemic happening in a year might not be that different than the probability that a building burns down, at least within a year, it does happen. And then, when it does happen, it affects society in such a profound way because we're so dependent on transportation, logistics, all of these things in modern life. Yeah, you should put some effort aside for pandemics, guys.

Now, it seems obvious today... The hope – I can't say as I feel terribly confident about it – the hope is that the world will remember this and will take more precautions because there'll be another pandemic.

Sp: Plus, it's asymmetric. If you're a nation and you want to cause harm to another nation, you might not be able to afford tanks and missiles and all these other things. But you can do cyber, you can do biological weapons, you can do other things.

NM: This will sound very strange, but thank God for nuclear weapons, in that most countries with ill intent, really, they see the United States or Russia, as superpowers because they own a nuclear weapon. It's really hard to make a real nuclear weapon. Bioterrorism is potentially worse because it's potentially much cheaper and much easier. It's also much harder to control. So, I think we'll see a bioterrorism event someday.

If just two years ago I had said there will be a war of conquest in Europe, you would've said I was nuts. The EU is already one uniform thing. And, as we speak, there's a war of conquest going on in what is geographically Europe. You can say, well, historically, Ukraine wasn't Europe. But it's right beside all these other places that are legally and treaty-wise part of Europe. And we don't know where that's going to end. We don't know anything about it.

I try to be an optimist about things. And it takes some willpower because pessimism is just so damn easy. With the pandemic, we did panic, and when we panicked, we did a lot of great things. I heard a speech by the CEO of one of the companies that developed a vaccine, Pfizer, I believe it was. And he said, 'If I had told everybody that we needed to do it in eight years, they would've said, "You're crazy. It takes us 10 to 12." So I told them we had to do it in eight months, because if I told them eight years, they would've taken the old process and tried to fix it in little ways, and that would never work.' This is brilliant. Whether he really thought all this upfront, I

“

I'm a huge fan of exploring the solar system and having both intra-solar system and interstellar probes. I think those are both ultimately fantastic things for us to do to explore the world. But it's vastly cheaper and easier and better to do it with machines than to do it with humans.

”

don't know. But it's a brilliant analysis, even if it's after the fact, because what you would have to do is break everything you know, yet you still care a tremendous amount about safety and all these other issues. And my God, the Pfizer vaccine and the Moderna vaccine – not only did they develop them in record time, but they are among the best vaccines we've ever made for anything in terms of their efficacy. Without these techniques from modern biology that allowed them to engineer these vaccines in a fundamentally different way, we never would've made it.

There will be a day when we really panic on climate, I predict. I don't know when that day will be. I don't know what set of effects or problems it would take to make us panic. But let's say that there's a day when we really panic. So suppose that we panic one day, when we say, 'Right, from tomorrow onwards, there will be no more emissions.' Well, unfortunately, the chemistry and physics of climate change are not very forgiving. So, when you emit CO₂, it stays in the atmosphere for a really long time – tens of thousands of years. It turns out, if you stopped all emissions on day one, the climate would continue to drift upwards for about 100 years. I've written a bunch of papers on global warming; I've done this calculation. Other people have done this calculation. You can quibble as to whether it's 85 years or 110 years, but it's something in that range. Then it'll slowly start going down. It will

take 145 years for it to equal the same temperature on the day that you panicked and shut it off. And that was assuming you did it instantly.

We have the problem that, at the moment, we're not developing alternative energy sources fast enough. But once we get there, just switching over – and of course you can't switch over the whole economy in a day – if you then get more realistic about it, then that says you probably are going to have increasing temperatures for more than 100 years after that day. Now it turns out there's a possible technological fix to that, which is something called 'geoengineering,' which gets people very riled up because they think that's terrible for X, Y and Z reasons.

SP: What is geoengineering?

NM: Geoengineering is a way to directly combat global warming. The earth is kept warm by sunlight from the sun. That is our primary source of heat. We would all be frozen on Earth if that wasn't the case. The problem with enhanced CO₂ in the atmosphere and other greenhouse gases is that light will come from the sun and bounce off the Earth. Some of it is absorbed by the Earth and warms the water or warms the land. Some of it reflects off as light. Then, of the heat that goes into, say, the ground, that makes the ground hot and gives off a little bit of infrared radiation. That infrared radiation used to just pass right through the atmosphere back out into space. It still does, mostly. But a tiny bit of it gets trapped. The CO₂ traps it. If you have a flask of CO₂ or other gases and you shine infrared light of the appropriate wavelength through it, it absorbs it.

SP: So the more CO₂ in the environment or in our atmosphere, the more we're going to retain.

NM: Yes. And the classic metaphor for this is the greenhouse. If you're in a greenhouse, the visible light of the sun goes through the glass or the plastic that makes the greenhouse. The infrared radiation from the things inside that are heated up doesn't pass through, it gets trapped. And that's why a greenhouse can be so warm, even if it's the middle of winter in Alaska; you can be much warmer in a greenhouse because it traps the heat. But the amount of heat that's trapped is tiny. It's about 1 per cent roughly of what the sun produces. So, if you average the sun's output here on Earth over the whole Earth, and over the whole year, it's about 300 watts per metre, rough enough. The amount of extra heat that's trapped in the atmosphere is a quantity called 'radiative forcing.' And that's a little less than 1 per cent.

SP: But is that compounding every year?

“

I try to be an optimist about things. And it takes some willpower because pessimism is just so damn easy.

”

NM: Yes. One per cent of a ginormous number is a lot. But here's the thing: if you could make the sun 1 per cent dimmer, well that would solve the problem, wouldn't it? Now there are natural things that effectively do this. One is certain kinds of clouds. It turns out the physical clouds are very complicated, but clouds up in the high atmosphere can reflect the sunlight so it doesn't hit the ground. That helps. Ice helps. So, one of the big problems in the Arctic is that you have all of this area covered with sea ice. Sea ice is white, it reflects a lot of light. If you replace that, sea ice melts, now you have water. First order of approximation, deep ocean water is black. It absorbs all of it. Big problem. Deserts actually reflect a lot. The Sahara Desert reflects a tremendous amount of heat. So, if you could reflect enough sunlight back into space, you'd solve global warming.

Now there are simple things which sound dumb, but they're a good idea. They wouldn't fix the whole problem, but they would go little ways, which is, we should have white roads. We should always have a white roof. Black roofs are terrible from this perspective. However, if you really wanted to affect this, what you'd have to do is put particles in the upper atmosphere.

There's a bunch of ways to do this. But one inspiration is volcanoes. So when volcanoes erupt, they put gases and particles into the atmosphere. A very strong volcano puts it in the upper atmosphere, the stratosphere. And Mount Pinatubo did that in



1991. It changed global temperatures by 1 to 1.8 degrees Fahrenheit for a year. After about a year, year and a half, the particles fall out of the sky. There are people that want to put these particles up there, and there are ways that you could engineer to put those particles there. That is one of the methods of geoengineering.

One concern is that if you offer geoengineering as a solution, people will gravitate to that solution rather than making hard decisions. Another concern is if you do that, you will get some other problem, that you're just trading one problem for another. And to that, I say, yes, exactly, just like all of human history.

SP: How do you think about the global cooperation that's needed between nations? As one nation you can have an impact, but if others defect in some cases they would get an advantage in terms of cost of manufacturing, cost of energy?

NM: A huge advantage. So of course, people are jealous of other folks that don't do the same thing. That's one of the problems with global warming ... it affects all of us. I'm involved in a nuclear company. For a while we had a deal to build one of our brand-new plants. We've invented a new type of nuclear plant that's way safer and it's got lots of other benefits. We were going to build our first one in China, and there were lots of people in the US who said, 'How dare you build it in China? Why would you let them

have it?' And I said, because we all share the same atmosphere, and they're the ones building new coal plants. We're not building new coal plants in the US, so actually that's where the problem is. Now, as it turns out, that project didn't go forward and we're currently trying to build our new plant in Wyoming. We have a deal to do that, which is very exciting. You have this issue that the developed countries of the world – the US, or the UK is an example – we've both been burning lots of fossil fuels for 150 years at least. And yet, we want to tell China and India that they should not do that. Those people say, 'What do you mean? We have people who are living way below your standard of living. You got rich this way; it's our turn.' And irrespective of what you think about the moral imperatives, we don't have a way to force them unless you want to go to war over it.

SP: Thank you for taking the time today, Nathan. I appreciate it.

NM: Well, thank you.



“I promise to help end homelessness.”

The Thomas Family, Melbourne, Victoria

Every night 116,000 people in Australia sleep homeless. While another 1.5 million live in serious housing stress. We have a housing crisis. We can't let this go on. And with Homes for Homes, we won't.

Make yours at homesforhomes.org.au

Homes
for
Homes

PERFECT DOESN'T GROW ON TREES.



But wonky does. And big.
And curvy. And dinky.

Around 40% of all food grown goes uneaten, wasting with it, a humongous amount of CO2e and water.

Join the fight against food waste
and make a difference to the
planet with every delivery.
Go to oddbox.co.uk

GET 50% OFF YOUR
FIRST BOX WITH CODE
CURVYCOURGETTE



ODDBOX



VISIT WWW.PODREAD.ORG TO PURCHASE BACK EDITIONS

Leighton
Fine Art

Leighton Fine Art

Dealers in important 19th and early 20th Century paintings.
Impressionist and Post-Impressionist specialists established in 2001.



Sunrise – Sint-Martens-Latem

by Leon De Smet

We offer one of the world's most important hand-picked collections

Over 400 paintings in stock.

View the collection at: www.leightonfineart.co.uk

Or contact us: **Tel: +44 (0)1628 523552** info@leightonfineart.com

Free shipping worldwide and no-quibble, 7-day money-back guarantee.

neat.

clean reimagined



Refillable Aluminium Bottle



Free-From
Single-Use
Plastic



Vegan &
Not Tested
On Animals



Powerfully
Plant
Based

Just add tap water

www.neatclean.com