VOL FOUR / *ISSUE* THREE

New Scientist THE COLLECTION

ESSENTIAL KNOWLEDGE

EVERYTHING YOU NEED TO KNOW TO MAKE SENSE OF THE 21ST CENTURY

£9.99

ARTIFICIAL INTELLIGENCE, *MIGRATION*, AIR POLLUTION, *FAKE NEWS*, YOUR IRRATIONAL BRAIN, *ENERGY*, CLIMATE CHANGE, *THE ALGORITHMS THAT RUN YOUR LIFE*, THE FUTURE OF WORK… *AND MORE*

NewScientist
THE COLLECTION

VOL FOUR / ISSUE THREE
ESSENTIAL KNOWLEDGE

NEW SCIENTIST THE COLLECTION
110 High Holborn,
London WC1V 6EU
+44 (0)20 7611 1202
enquiries@newscientist.com

Editor-in-chief Graham Lawton
Editor Daniel Cossins
Art editor Craig Mackie
Picture editor New Scientist Picture Desk
Subeditor Hannah Joshua
Graphics Dave Johnston
Production editor Mick O'Hare
Project manager Henry Gomm
Publisher John MacFarlane

New Scientist The Collection is published four times per year by New Scientist Ltd
ISSN 2054-6386

Printed in England by Williams Gibbons (Wolverhampton) and distributed by Marketforce UK Ltd
+44(0)20 3148 3333

Cover image
Leandro Castelao

Smarten yourself up

SCIENCE can be exhilarating and awe-inspiring. Who could fail to be moved by the warping space-time described by Einstein's theory of general relativity, or by evolution's breathtaking perspective on the story of life?

But it can often feel abstract and remote from everyday concerns. Grasping quantum mechanics won't help you understand the migration crisis, for instance, and knowing your neurons from your dendrites doesn't mean you can overcome your irrational brain's tendency to make bad decisions.

The good news is that there is another side to science, one that can help you make sense of our rapidly changing world. This issue of *New Scientist: The Collection* serves up the stuff you need to know to understand – and negotiate – life in the 21st century.

We start by going back to basics. Chapter 1 asks: What is knowledge? Why do we seek it? And can we ever really know the truth? These are pressing questions at a time when distinguishing fact from fiction is increasingly fraught.

Chapter 2 is dedicated to helping you make better decisions. You will get to grips with your hidden biases and how to overcome them, the pitfalls of misunderstanding probability, and the maths that can make you richer.

Chapter 3 gets existential. We start with an eye-opening look at how we calculate the value of a life, and then tackle its very meaning. We also delve into the depths of the human mind to reveal why we all skew the facts to fit our existing beliefs.

Chapter 4 takes you to the contentious frontiers of biomedical research. We introduce the revolutionary gene-editing technique CRISPR, and explore the opportunities and risks that come with the power to edit life. We ask how we can beat the next global pandemic, and we reveal the truth about air pollution. We also bust a few myths in our guide to misleading medical statistics.

Chapter 5 lifts the curtain on the maths that runs your life. You don't notice the algorithms running behind the scenes, but they control the world and, increasingly, you. On a similar tack, Chapter 6 is all about the rise of the machines. Artificial intelligence is here, and it is improving all the time. How did we get to this point? Should we welcome AI or fear it?

Chapter 7 clocks on to investigate the future of work. We ask whether we should worry about robots taking our jobs, and explore the prospects for a radical new way to run the economy.

Chapter 8 is all about notions of nationhood. We lay out the evidence about migration, from the economic impact to the evolutionary psychology of xenophobia. Then we ask whether the nation state has had its day, and, if so, what the alternatives are.

Chapter 9 is dedicated to the realities of climate change. It's happening, so how bad is it and what can we do about it? It also includes a handy guide to winning arguments with so-called climate sceptics.

Finally, in Chapter 10, we look at energy. We join the race to build a better battery and discover the emerging technologies that could revolutionise the way we power our lives. We also explore different ways to store and supply energy, and chart chemists' quest to crack the reaction that could give us unlimited juice, with none of the environmental impact.

At the end of your journey, you should have all the knowledge you need to understand what is really going on in the world. And knowledge, as they say, is power.

Daniel Cossins, Editor

CONTENTS

VOL 4 / ISSUE 3

ESSENTIAL KNOWLEDGE

CONTRIBUTORS

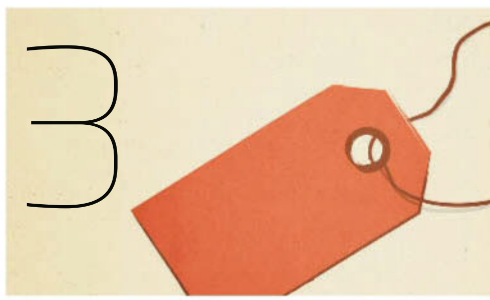
Sally Adee
is technology news editor at *New Scientist*
Gilead Amit
is a feature editor at *New Scientist*
Kat Arney
is a writer and broadcaster based in London, UK
Jacob Aron
is analysis editor at *New Scientist*
Michael Bond
is a consultant for *New Scientist*
Catherine Brahic
is a feature editor at *New Scientist*
Michael Brooks
is a consultant for *New Scientist*
Teal Burrell
is a writer based in Richmond, Virginia
Jon Cartwright
is a writer based in Bristol, UK
Nello Cristianini
is at the University of Bristol
Daniel Cossins
is a feature editor at *New Scientist*
Catherine de Lange
is deputy features editor at *New Scientist*
Richard Elwes
is at the University of Leeds
Shannon Fischer
is a writer based in Boston, Massachusetts
Nic Fleming
is a writer based in Bristol, UK
Marianne Freiberger
is co-editor of online maths magazine *Plus*
Mark Harris
is a writer based in Seattle
Douglas Heaven
is a writer based in London, UK
Olive Heffernan
is a writer based in Dublin, Ireland
Hal Hodson
is a reporter at *The Economist*
Bob Holmes
is a consultant for *New Scientist*
Joshua Howgego
is a feature editor at *New Scientist*
Dan Jones
is a writer based in Bristol, UK
Alice Klein
is a reporter at *New Scientist*
Debora MacKenzie
is a consultant for *New Scientist*
Paul Marks
is a technology journalist based in London, UK
Tiffany O'Callaghan
is a feature editor at *New Scientist*
Michael Le Page
is a reporter at *New Scientist*
Fred Pearce
is a consultant for *New Scientist*
Rachel Thomas
is co-editor of online maths magazine *Plus*
Richard Webb
is features editor at *New Scientist*
Jon White
is comment editor at *New Scientist*

The articles here were first published in *New Scientist* between February 2011 and June 2017. They have been updated and revised.

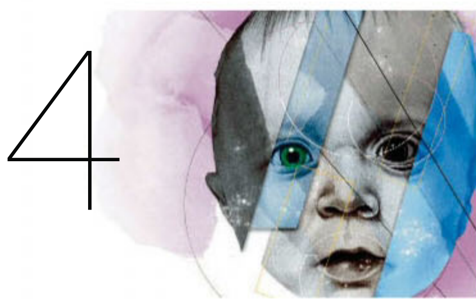
Epistemology

Intelligent life

Know thyself

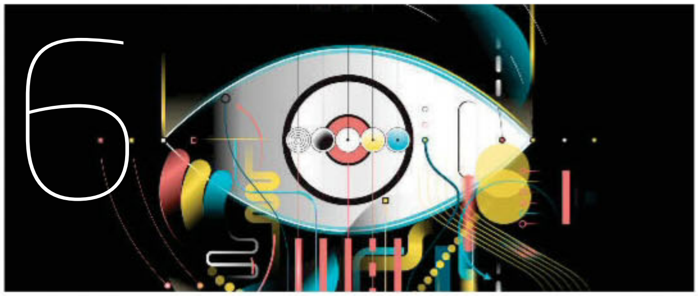


Medical frontiers



Slaves to the algorithm

Rise of the machines



Clocking off

World in motion

Warming planet

Powering the future

CHAPTER ONE

EPISTEMOLOGY

INSIDE KNOWLEDGE

The pursuit of knowledge underlies human progress. In many ways we know more than ever before – yet distinguishing fact from opinion and truth from belief can seem ever harder to do. Here we go back to basics, asking what knowledge is, what it is worth to us as individuals and as a species, and other central questions of epistemology – the science of knowledge itself

WHERE KNOWLEDGE COMES FROM

One way to classify knowledge is by how we acquire it

PERCEPTUAL
The direct evidence of our senses

TESTIMONIAL
Facts we acquire from other people and media

INNER SENSE
Awareness of our own feelings and states, such as pain and hunger

INFERENTIAL
Knowledge we stitch together ourselves from raw inputs

WHAT IS KNOWLEDGE?

I'VE won the lottery. I haven't checked my numbers yet, but I just know. You know what it's like, when you just know you know.

Knowledge is a slippery concept: what we know, how we know we know it, what we know others know, what others know of what we know, how what we or they know differs from what we just believe. You would probably argue that, until I see the winning numbers, I can't know I have won the lottery – I can only believe it. Yet how do you know that?

Most of us make our way through life without peering too closely under the bonnet of epistemology – the theory of knowledge. "We manage it intuitively, we don't have to reflect and calculate," says Jennifer Nagel, a philosopher at the University of Toronto, Canada, and author of *Knowledge: A very short introduction*.

JAMIE MILLS

> *"Knowing something is a far richer, more complex state than merely believing it"*

But it rewards closer inspection. For a start, the degree to which we know stuff and know what others know is quite possibly what separates humans from everything else on the planet, from rocks to chimpanzees. It is certainly the lubricant of all human interactions. "We can cooperate, communicate and compete better if we know what others know and don't know," says Nagel. "Tracking states of knowledge can help you in the course of the argument, brace yourself against fallacies."

Yet defining the basic currency of knowledge is surprisingly difficult. To know something you must first believe it, but that's not enough: to make factual knowledge, that belief must also be true. "That is the one thing we're all happy to accept," says Nagel.

"True belief" is insufficient, though. A belief can be true just by chance, or we can arrive at a right answer via a wrong route. So epistemologists have traditionally added another condition for a true belief to count as knowledge: it must also be justified in some way. In the lottery example, the perceptual evidence of the numbers on my lottery ticket plus the testimonial evidence of, say, a broadcaster reading out the winning numbers creates the inferential knowledge that I have won – or not (see "Where knowledge comes from", above right).

For a long time, the conception of knowledge as a justified true belief ruled the roost. But then US philosopher Edmund Gettier put forward a couple of devastating counterexamples in a short paper published in 1963.

An example of a "Gettier problem" is someone glancing at a clock that says 12, at midday. The catch is that the usually reliable clock is broken, and is showing the right time only by chance. Our clock watcher believes it is midday, that belief happens to be true, and the stopped clock provides justification. But in fact no one knows it is midday – they just believe they know it.

STOP THE CLOCKS

Various attempts have been made since to tighten up the standards of justification to get round this sort of problem and provide a definition of knowledge everyone can agree on. But no one has quite yet nailed this one down. "It's kind of an awkward question," says Nagel.

In the end, though, an answer might not be the point. All these epistemological investigations point us to one fact that we are wont to forget: that knowing something is a far richer, more complex state than merely believing it. The ability to distinguish between fact and opinion, and to constantly question what we call knowledge, is vital to human progress, and something we cannot afford to let slip (see "How can I know the truth?", page 10).

"Knowing something is a mental state that locks you on to the truth," says Nagel. What that lock is, though – well, we don't really know. Richard Webb

WHY DO WE SEEK

KNOWLEDGE?

IN 1969, Robert Wilson, the first director of Fermilab near Chicago, was asked by a US Congressional committee whether the new particle accelerator he was seeking funds for would aid the fight against the Soviet Union. "This new knowledge has all to do with honour and country," he said. "But it has nothing to do directly with defending our country except to help make it worth defending."

Wilson's full testimony is a robust and elegant defence of the pursuit of knowledge for knowledge's sake. But it raises its own questions. In what sense is knowledge "worth it"? And what motivates our urge to acquire it?

Knowledge is more than just information. Even the nematode worm *Caenorhabditis elegans*, owner of one of the smallest brains we know, forages to maximise information about its environment, and so its chances of staying alive and reproducing.

But as far as we know *C. elegans*, or indeed any species other than our own, doesn't ponder the universe's origins; they certainly don't publish papers on it or build particle accelerators to find out. Knowing as we understand it involves abstracting information and interpreting it for use at different times and in other contexts. "When you have knowledge, you can do lots of things," says epistemologist Duncan Pritchard at the University of Edinburgh, UK. "You can deal with entirely new situations in creative ways."

Reading this article – weighing up its beliefs, truths, justifications and perhaps misapprehensions – won't get you a square meal or make you more attractive to a potential sexual partner (or perhaps only indirectly). And yet brain-imaging studies show that when we answer trivia questions or look at blurry images designed to pique curiosity, areas associated with our response to food and sex light up. That suggests we treat knowledge as a similar primary reward.

The precise details of how we first came to love knowledge may always elude us. But it is easy to see how it would have spurred our success as individuals and as a species, furnishing us with the tools – often literally, if you think of cutting blades or fire – to survive and prosper.

In that case, we are in some way addicted to knowledge because it has served us so well in the past – as it still does now, in everyday life as well as at the frontiers of technological progress. As Abraham Flexner, founder of the Institute for Advanced Study (IAS) in Princeton, New Jersey, pointed out in a 1939 essay "The usefulness of useless knowledge", radio communication and all that came with it wasn't ultimately the invention of Guglielmo Marconi. It was down to James Clerk Maxwell and Heinrich Hertz, scientists who worked out the basics of electromagnetic waves with no practical objective in mind.

MAKING THE CASE

There are plenty of similar examples, says Robbert Dijkgraaf, current director of the IAS, who has written a companion essay to a reissue of Flexner's original this year. "The theory of general relativity is used every day in our GPS systems, but it was not the reason Einstein solved it," he says.

That doesn't mean science gets a blank cheque – and certainly not in a world where children are going hungry, as Wilson made plain in his testimony. There is such a thing as useless scientific research – it is just hard to say what it is, says Massimo Pigliucci, a philosopher of science at the City University of New York. "That is why scientists need to make a case to their peers and to the public for why what they do is interesting or important."

As to why they do it, it's no different for scientists than for anyone else. Seeking knowledge is what separates us from the worms. Daniel Cossins

> *"As far as we know, nematode worms don't ponder the origins of the universe"*

IS SCIENTIFIC KNOWLEDGE

SPECIAL?

NULLIUS *in verba*: "take nobody's word for it". The motto of the Royal Society, the UK's national academy of science, encapsulates the spirit of scientific enquiry. Do an experiment, record its outcome faithfully and objectively, and make that record available for doubters.

This way of working means that, if knowledge is defined as the route to the truth (see "What is knowledge?", page 7), science is an expressway to enlightenment. Thanks to what science tells us about human physiology, the universe's history, nature's forces and Earth's geology, flora and fauna, we know Earth isn't flat, the universe is nearly 14 billion years old, and that there are no dragons or unicorns. We live longer and in more comfort, and can send space probes to the edge of the solar system. Pretty darn special, huh?

But let's take a more sceptical look, starting with that "we". Some people do believe Earth is flat. Others say the universe is 6000 years old. Some doubt the theory of evolution by natural selection, or the reality of human-made climate change. We is not everyone.

UNICORNS IN LOCH NESS

It is tempting to say that's their problem, not science's. But science is also limited in what it can say. It can't prove a negative: there might be dragons and unicorns, a monster in Loch Ness, a God. It can't even be definitive about all the positives. "Our evidence may at times leave us able to make only probabilistic judgements – we may sometimes be restricted to saying that a certain outcome or theory is likely to be true," says Jennifer Nagel at the University of Toronto, Canada.

This weakness becomes greater as we extend the scientific method into more complex realms with more variables and so more uncertainty, such as social science or climate change. Science progresses legitimately through speculation and hypothesising, but until these speculations are tested by experiment, for a stickler any "knowledge" that emerges from them must strictly be labelled as provisional.

It is a weakness (or strength, depending on your point of view) exploited with gusto by climate change sceptics. But it points to a blunt truth: if scientific knowledge feels special to you, you are part of its in-group. As we grow up, we absorb beliefs from our cultural environment. For some that means accepting scientific knowledge; for others it means "revealed" knowledge, from the Bible, say.

And here's the thing. For all the bluster about "the evidence", if you are a scientific believer you too are taking almost all of it on trust. "In principle, everybody should be able to replicate scientific results given time, money and training," says Brigitte Nerlich at the University of Nottingham, UK. "But not everyone has a Large Hadron Collider or a climate-modelling computer." You are taking someone's word for it. Like other forms of knowledge, most of science comes down to trusting the source.

Not special, then? Perhaps – except that science also provides mechanisms to justify trust in the knowledge it generates. "Authority in science is earned – at least, when a scientific community is functioning well – by success at predicting, and more generally at analysing, empirical phenomena," says philosopher Edward Hall of Harvard University. Science's conclusions are accepted when they fit with our experience of the physical world, and are discarded when they cease to. That makes trust in science a justified true belief – and the knowledge that true science generates a cut above the rest. Just don't take my word for it. **Michael Brooks**

HOW CAN I

KNOW THE TRUTH?

POST-TRUTH was 2016's word of the year, according to Oxford Dictionaries. Not least in the furious debates surrounding the UK Brexit vote and Donald Trump's election as US president,there were claims and counter claims of fake news, dodgy experts and media mendacity.

For a hardcore of relativist philosophers, that's all a storm in a teacup – there's no such thing as objective truth that exists outside our minds. Nonsense, harrumphs Peter van Inwagen of the University of Notre Dame in Indiana. If a doctor says I have cancer of the gut, he says, "whether that is true depends on what is going on in my gut, and not on what is going on in my doctor's mind".

Accept that, and the challenge is to ensure that our inside knowledge is aligned as far as possible with outside truth.

That's hard, not least because in a complex society we rely on the knowledge of others, even when we don't realise it. Ask someone if they know how an everyday object such as a ballpoint pen works and they'll generally say yes, until you ask them to explain it. It turns out that our confidence in our own knowledge is often based on the certainty that somebody else knows.

That is often good enough; ballpoint pens exist and work. "As individuals we know hardly anything," says Steven Sloman of Brown University in Providence, Rhode Island, co-author of *The Knowledge Illusion: Why we never think alone*. "But most of us do very well and as a society we create incredible things. We sent a person to the moon. How is that all possible? Because of the knowledge of other people."

So how much should we trust people who actually do know stuff? "It's not that we want people to uncritically accept whatever experts say," says Timothy Williamson of the University of Oxford. A certain level of scepticism is healthy.

But it makes things difficult if we begin to mistrust expertise as a default. In philosophy, a true sceptic questions everything, so they have nothing left to build knowledge on. That's where we don't want to go as a society – while not losing sight of the fact that expertise differs in value and reliability (see "Felicitous falsehoods", below right). We should accept that, if we need a tooth pulled, going to see someone with a degree in dentistry just guarantees a dentist, not necessarily a good dentist, says Catherine Elgin of Harvard University.

How do we tread that fine line between healthy scepticism and destructive cynicism? First, think critically and assess the credentials, track record and potential bias of the sources we rely on. "If somebody is telling me this, what motives could that person have for wanting me to believe that, other than that it's true?" says van Inwagen. "Those are the practical questions."

Pose the same questions of yourself, too. "Ask 'How do you know?', 'How do they know?', all the time," says Elgin. Train yourself to ask whether your reaction to new knowledge is rooted in something trustworthy or something else, like wishful thinking. "Think about something like global warming," says Elgin. "To do something about that might require a certain amount of rather inconvenient stuff, so you'd really rather not believe it and start to make the sacrifices you would have to make."

And finally, avoid the seductive belief that you are privy to knowledge purposely being denied to others, or a warrior for truth when all others are peddling lies. Human beings are, in general, terrible at keeping secrets. "If you were sceptical you wouldn't be convinced by conspiracy theories," says van Inwagen. But these experts would say that, wouldn't they?

Tiffany O'Callaghan

> ***"If you are a sceptic, you won't be convinced by conspiracy theories"***

FELICITOUS FALSEHOODS

To get at the truth, sometimes we need to lie – for instance in building up scientific models that simplify an often complex world. The ideal gas law, for example, tells us how the volume, pressure and temperature of a gas are related, but assumes the individual molecules of the gas behave as perfect spheres that bounce off each other elastically.

Of course they don't. "But if you took into account the actual shapes of the actual gas molecules in a volume of gas, just the geometry, the problem of saying what is going on with this gas would be incalculable," says Catherine Elgin, a philosopher at Harvard University.

So these felicitous falsehoods help us dig at deeper truths – as long as we don't forget where they are and let them become the weight-bearing part of the structures we build.

Economists, for example, have traditionally created models of stable markets by assuming that buyers and sellers have perfect information and make rational choices – only to be constantly surprised when the irrationality of human decision-making creates a messier reality.

So simplify for science's sake – but sensibly. "The real problem isn't with doing it, it's not being aware of what you're doing," says Elgin.

CAROLYN DRAKE/MAGNUM PHOTOS

HOW MUCH CAN WE
EVER KNOW?

ULTIMATELY, the jumbo shrimp tells us why we are doomed to never fully understand reality.

But let's back up a bit first. We know we live in a universe where there are hard physical limits to what we can know. Light's finite speed restricts our ability to see in time and space, quantum uncertainty our understanding of subatomic particles.

So what? Our largest telescopes look back to a few hundred million years after the big bang, while our sharpest microscopes can spy on individual photons escaping from atoms. The universe is as it is, and we work quite well within its limits. True, we can't explain what happens at the big bang, or inside a black hole – but that's just a matter of devising better theories of nature and ways to test them.

So to know more we need to compute better. Easier said than done. Were we able to simulate the fine-grained movements of all the universe's matter, we might predict its evolution and fate. But with current computing power, that would take more time than the universe has to offer.

Computational power is a practical limitation we can blame for everything from unreliable weather forecasts to shoddy logistics: once you try to optimise an itinerary linking more than a few thousand destinations, it becomes impossible to compute. "There are so many parts to it, it's simply improbable that we can work it out," says Noson Yanofsky, an information scientist at the City University of New York.

But ultimately, that's just a fig leaf for a mega-sized limitation. However powerful we make them, computers ultimately rely on human input to program them – and human thought is a glorious mess. Statements like "this statement is false", hating someone yet loving them and yes, that small-yet-large jumbo shrimp, both compute and do not compute. "Language is an expression of the mind, and my mind and language is full of contradictions," says Yanofsky.

That flexibility allows us to think outside the box, while remaining firmly inside it. Because we are predicated on contradiction, we see contradiction everywhere. The defining feature of reality, however, is that it admits no contradiction. Take the way quantum objects apparently act as waves or as particles according to how we choose to measure them, a confusing duality physicist Richard Feynman called "the only mystery" of the quantum world. In all probability, the basic building blocks of reality are neither wave nor particle, but something else entirely. It's just something that we lack the experience or cognitive ability to express.

Logic, and the mathematics that builds on it, is supposedly our way out. That's all very well until you encounter the logical limitations of mathematics itself.

These start with well-known injunctions such as never to divide a number by zero. Why not? Because if you do, you can begin to do things like prove 1 = 2. If maths is the language of a flawless universe, we can't allow that – so we don't. "If you want mathematics to continue without contradictions then you have to somehow restrict yourself," says Yanofsky.

And sooner or later, we come to the end of the road. As Austrian mathematician Kurt Gödel showed in the 1930s, any system of logic containing the rules of arithmetic is bound to contain statements that can be neither proved nor disproved. It will remain "incomplete", trapped in the same inconsistency as we are (see "An unknowable problem", left). Gödel incompleteness is a mathematical expression of the logical-illogical statement "this statement is false". The fundamental truth is there is no way for anything, be it a simple sentence, a system of logic or a human being, to express the full truth about itself.

This problem of self-reference is endemic. Gödel's contemporary Alan Turing showed that you cannot ask a computer program in advance whether it will run successfully. Quantum mechanics sprouts paradoxes because we are part of the universe we are trying to measure.

So the sobering truth is that we can build the most powerful telescopes and computers we want, but we will never overcome the limitations of our minds. Our perspective on reality will always be skewed because we – and the jumbo shrimp – are part of it. **Richard Webb**

AN UNKNOWABLE PROBLEM

It is a statement whose truth is impossible to prove. Only we can't prove that.

This "continuum hypothesis" has to do with what sort of infinities exist. There are in fact infinite levels of infinity. The lowest is the "countable" infinity of the whole numbers: 1,2,3,4,5 and so on. Higher is the "continuum" infinity of the real numbers - all the countable numbers plus all others with any number of decimal places in between.

In 1878, Georg Cantor hypothesised that the countable infinity and the continuum infinity are neighbouring rungs in the ladder; there is no other infinity between them. But he couldn't prove it, and no one ever has.

Attempts have instead revealed the existence of a "multiverse" of different mathematical worlds, all producing the sort of logical structures that correspond to physical reality, but differing in whether the continuum hypothesis is true or false. That is perhaps an indication that mathematics is itself only part of a much larger logical structure we have yet to reveal.

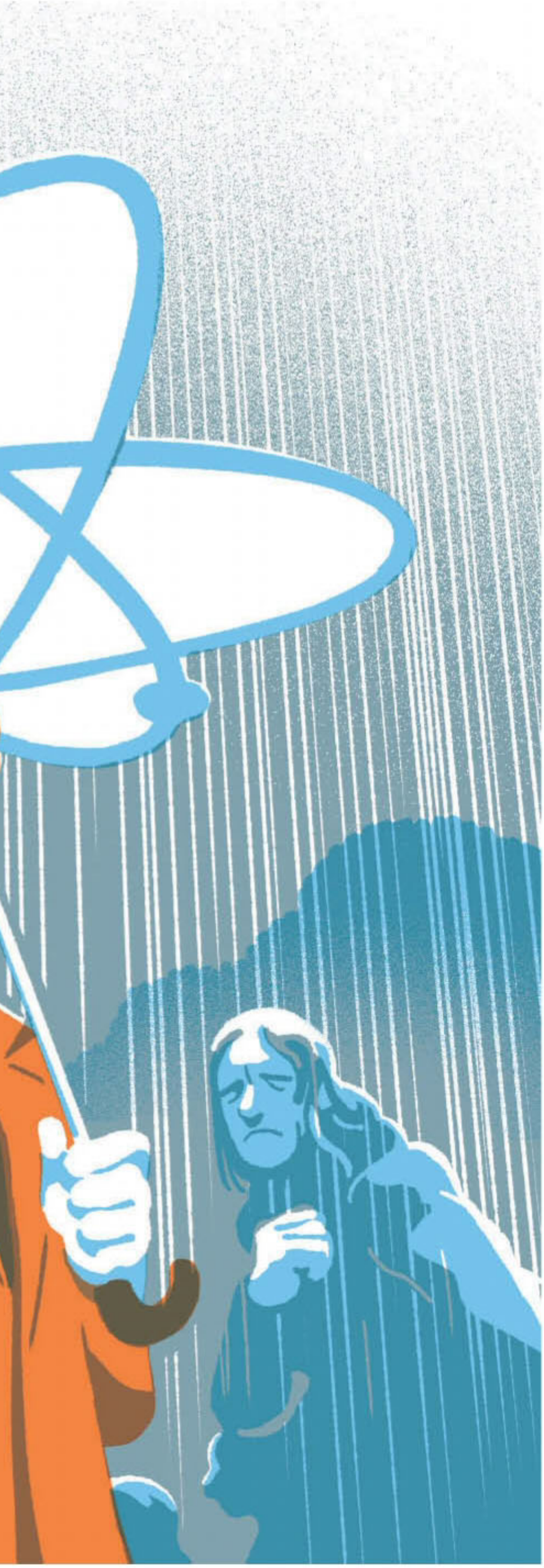
CHAPTER TWO

INTELLIGENT LIFE

GET *SMARTER*

Science has given us much to marvel at, from Newton's universal law of gravitation to Darwin's evolution by natural selection. It's wonderful, useful and often beautiful – but sometimes it can feel a little divorced from our everyday realities.

Beyond these great and grand theories, however, there is another canon of knowledge. Grasp these ideas and they can help us lead better, happier and smarter lives. This is the science you need to navigate the world



BRUNO MANGYOKU

NEED TO KNOW

COGNITIVE BIAS

Evolution has made our brains irrational

IT'S Sunday morning and I'm feeling a bit impulsive, so I head to a cafe near my home in London for breakfast. I open the menu and see the following:

MENU

Breakfast

Full English breakfast £9.95

Smoked salmon & scrambled eggs £5.95

Waffles with maple syrup £4.75

Boiled egg and soldiers £4.00

What would you have picked? I went for the smoked salmon and scrambled eggs. And a surprising number of you would have done the same.

Why so? Understanding the often irrational factors that affect how we make decisions has been a key aim of psychologists over the past few decades – and we're just getting to the stage where we can begin to apply their insights.

That menu first. The reason why many of us would plump for the smoked salmon (besides, perhaps, liking smoked salmon) is to do with something called relativity. That's not Einstein's theory, but rather our tendency to be awful at assessing an item's value without being given something to compare it with. As Dan Ariely of Duke University in Durham, North Carolina, puts it in his book *Predictably Irrational*, humans "don't have an internal value meter that tells us how much things are worth".

In this case, the presence of the rather expensive-looking full English breakfast makes the smoked salmon seem like good value. Take the full English off the menu, and more of us will choose one of the two cheaper options. Regardless of whether anyone actually buys the full breakfast, then, its presence means that we're shelling out more than we otherwise would.

A similar effect, known as anchoring, often kicks in when we're out shopping. Say we have £50 to spend on shoes. We see a pair we like for £100, way above our budget. Then we see a similar pair reduced to £75, and without much thought snap them up. The first price anchors the idea that the second price represents a bargain: we end up thinking we've made a saving when in fact we've spent more than we can afford.

Such foibles persist regardless of the size of a purchase. Another phenomenon called hyperbolic discounting means we tend to overvalue what's available now relative to what we can have later – one reason why many people find it hard to invest in a pension scheme (see "Exponential growth", page 16). Then there's the sunk-cost fallacy, the tendency to stick with something we have already invested in even if all the signs point to this being a bad idea. The classic example is the supersonic jet, Concorde, which never made any money in all the decades it was flying.

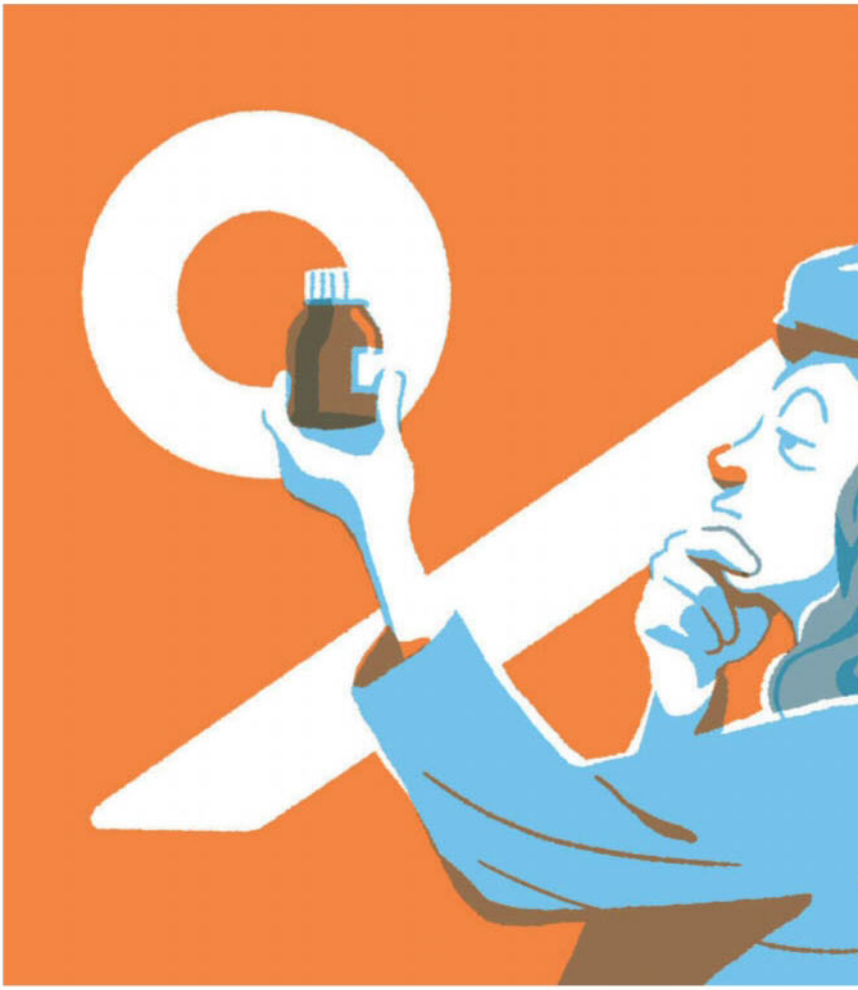
There are good reasons why such biases are embedded in our psyches. They have evolved to help us make quick decisions with limited information and difficult decisions with large amounts of hard-to-assess information; to continue foraging in an area that's becoming depleted of fruit or move on to another part of the forest where richer pickings are not necessarily on offer, for example.

But to make more logical, calculated decisions in the complex modern world, we › need different mental formulae. And there are some simple ways to construct these. A list might include getting yourself or someone else to play devil's advocate to any significant decision; identifying and discounting any sunk costs;

"Awareness of mental biases is good for more than just being smart with our money"

and working methodically to eliminate options one by one before coming to a final decision.

Awareness of our cognitive biases is good for more than simply being smart with our money. It might also help limit the scope for mental blips to cause major disasters, for instance. In the wider realm of economics, models that take better account of real humans' irrational decision-making are coming more into vogue to explain events such as the crash of 2008, the better to avoid them in the future. When it comes to improving our futures, it's all in the mind. **Joshua Howgego**



NEED TO KNOW

STATISTICS

Don't believe every number you read

AH YES, statistics. The temptation to start any discussion of this subject with the aphorism popularised by Mark Twain is almost overwhelming. "Lies, damned lies, and..." You know the rest.

We can't afford to be that dismissive. Statistics is the science of drawing informed conclusions from large amounts of data. In a sense, then, it *is* modern science. From trials of the latest wonder drug to the discovery of the Higgs boson, breakthroughs that advance human knowledge are these days seldom made without someone somewhere applying statistical reasoning. And as those bits of knowledge filter down to the rest of us, we are increasingly expected to make decisions – from the political to the medical – on the basis of numbers with that confidence-inspiring suffix "per cent".

Trouble is, few of us do that sure-footedly. Sample sizes, false positives and the difference between absolute versus relative numbers are among the factors that affect how we interpret statistics. Often, they are impossible to extract from a bare number.

It's a systemic problem. "There are large numbers of experts – not just laypeople – who have no training in statistical thinking," says Gerd Gigerenzer of the Max Planck Institute for Human Development in Berlin, Germany, and author of *Risk Savvy: How to make good decisions*. "Children are taught the mathematics of certainty: algebra, trigonometry, geometry and the like. That's beautiful but often useless."

For a complex and risky world, he reckons we need a different type of preparation. "We should be taught uncertainty," he says. And that needn't be so difficult. For Gigerenzer, there are a few golden rules we can apply to sharpen our reasoning.

The first is to understand that there is no such thing as certainty, and that looking for it is an illusion. "There are risks everywhere and you need to quantify them," says Gigerenzer.

The second is to look for statistics that encapsulate absolute numbers, not relative ones. Say you read that popping a certain pill will reduce the risk of having a stroke by 50 per cent. This relative number means nothing if you don't know how likely you are to have a stroke in the first place. If that absolute number is 3 in 1000, a 50 per cent reduction will take it down to 2 in 1000 – a puny decrease.

You still might consider the pill worth taking. But wait for Gigerenzer's third rule: always look for the other side of the story. If told about a pill's supposed benefits, for instance, also ask about its potential risks – and make sure both are presented in the same, absolute terms. "I don't want to know whether a drug reduces something by

50 per cent," says Gigerenzer. "I want to know if half take it and half don't, what happens five years later."

In the medical arena, the same reasoning should raise a mental red flag whenever you read of a test being so-and-so per cent accurate – a meaningless figure unless you also know the test's false-positive rate (see diagram, right). Similarly when "survival rates" for a certain condition are quoted or compared: this is a relative measure that can vary considerably depending on how a condition is diagnosed and tested. What you want to know is the mortality rate, an absolute figure that tells you what proportion of the population will die of the condition over a certain period.

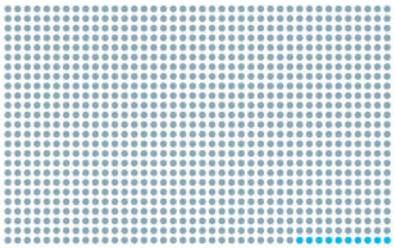
Such rules can help anywhere you see a number with statistical trapping. They have their limits – for example when someone has wilfully cherry-picked their data or otherwise massaged the figures. But they are a good start in sorting out a damned lie from a statistic. **Richard Webb**

Falsely positive?

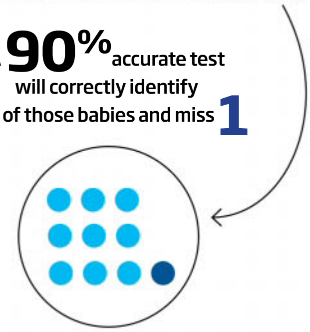
A prenatal test for Down's syndrome can be 90% accurate but also 85% wrong

The chance of having a baby with Down's syndrome varies with age. Assume it's 1%

Of **1000** babies, **10** will have Down's syndrome

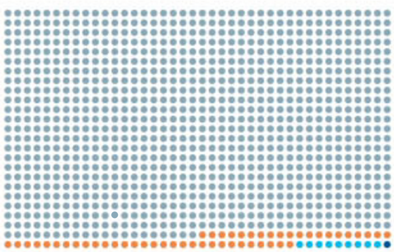


A **90%** accurate test will correctly identify **9** of those babies and miss **1**



But the test has a false positive rate of **5%**

Of the **990** babies who will never get Down's syndrome, about **50** will test positive for it



So in total **59** babies test positive, but only **9** of them actually have Down's syndrome

A positive result thus means the baby has a **15%** chance of having Down's syndrome

NEED TO KNOW

PLACEBO EFFECTS

Harness the power of positive thinking

MY MUM swears that reiki, a technique claimed to channel healing energy through touch, cured her painful frozen shoulder. And my sister promises me a homeopathic remedy will relieve my frequent stomach aches.

Such claims raise eyebrows among those who champion rational thinking. There is often no physiological mechanism by which these and other alternative therapies could work, and they regularly fail to pass the standard tests for efficacy in medicine. But if someone feels better after their chosen remedy, who are we to say it didn't work for them?

At the heart of such questions lies the placebo effect – the way that we tend to feel better just because we believe a medical treatment is going to work, even if the treatment itself is a sham. The power of placebos has been shown in many settings. In one study from 2002, 60 people were even given fake surgery to treat arthritic knees. An elaborate ruse involving doctored footage on a video screen convinced them that they had full surgery, whereas in reality they had only had the skin on their knees cut. Even so, their symptoms improved, and they recovered as well as those who had real surgery. The improvement lasted at least a year.

"We feel better if we believe a treatment will work – even if the treatment is a sham"

"It's hard to believe that sham surgery can produce a long-lasting effect," says Luana Colloca, who studies the placebo effect at the University of Maryland in Baltimore. But it can.

Pain seems particularly susceptible to placebos, but they can also improve the symptoms of other conditions, even asthma and Parkinson's disease. The effects are exceptionally strong in mental-health conditions such as depression and anxiety, says Irving Kirsch of the Beth Israel Deaconess Medical Center in Boston. In 2008, he and his colleagues found that the antidepressant Prozac and placebo were about as effective as each other.

The effect even works if we know about it. In a 2010 study, Kirsch and his colleagues gave an inert pill to people with irritable bowel syndrome. "We told them it was a placebo, but that it might make them feel better," says Kirsch. Even so, the volunteers › saw an improvement in their symptoms.

That doesn't mean alternative remedies like reiki or homeopathy are fine. The placebo effect might make people feel better, but that doesn't mean their underlying condition has improved. Harm might come from not seeking out proven treatments. And of course, any therapy comes at a price – hence ongoing squabbles in the UK over the public funding of alternative therapies such as homeopathy.

The positive message, though, is that by understanding the placebo effect, we can harness our minds to improve our own health prospects. Simply remaining optimistic when being treated helps, for example – as difficult as that might sometimes seem. So does maintaining a good relationship with your doctor, or surrounding yourself with people you feel comfortable with: studies have shown that hormones such as vasopressin, which are associated with trust, appear to boost the placebo effect.

A pleasant view can make you feel better too; a view of a park is known to improve recuperation compared with a view of a brick wall. If we can use what we are learning about the placebo effect to design medical treatments and clinical environments that are both physically and psychologically effective, we might all end up feeling a lot better. Jessica Hamzelou

ARENDA OOMEN/HOLLANDSE HOOGTE/EYEVINE

Feeling nice after reiki? That'll be the placebo effect

NEED TO KNOW

EXPONENTIAL GROWTH

The mathematical law that can make or break you

"THE greatest shortcoming of the human race is our inability to understand the exponential function." These are the words of the late Albert Bartlett, a physicist at the University of Colorado, Boulder, whose lectures on the subject became a YouTube hit. Arguably, he's right.

Take saving for retirement. "Start early" is the mantra, but it is easy to overlook just how much difference a few years can make. It all comes down to exponential growth – an often abused term that refers to anything that grows in proportion to its current value. It dictates that a forward-thinking 18-year-old can retire as a millionaire at 65 by investing around £250 a month with an average annual return of 7 per cent.

That figure might sound high by today's standards, but it's a rough average of the stock market return since 1960. The surprise is that when our saver reaches 55, the savings will amount to a little under £500,000. Thanks to the power of compound interest, however – exponential growth by another name – it will double to £1 million just 10 years later. Wait

NEED TO KNOW

PROBABILITY

Coincidences are more common than you think

IMAGINE you receive an envelope addressed in an unfamiliar hand. Enclosed are predictions for this weekend's football matches and an offer to invest in the sender's foolproof betting syndicate. What tosh, you think, shoving it in the recycling bin.

But come the weekend, you notice that those tips turned out to be correct. And then comes the really strange bit. The next week, an identical letter arrives with predictions for that weekend's games – and they turn out to be accurate too.

At this point, you send off your cash, convinced that whoever this person is possesses some genuine insight. (Either that, or you go to the police to report that you've uncovered the biggest match-fixing scandal yet.)

Or if you're familiar with the law of large numbers, you might be tempted to bide your time. This law, a facet of the perennially bamboozling subject of probability, states that, given a large enough sample size, any outrageously improbable thing is eventually bound to occur. If our sly soothsayer simply sent letters systematically to enough different people, each with a different set of scores, then at least one recipient is likely to get accurate predictions enough times in a row to make them bite. And even if just a few people hand over their money, it probably makes the scam worthwhile.

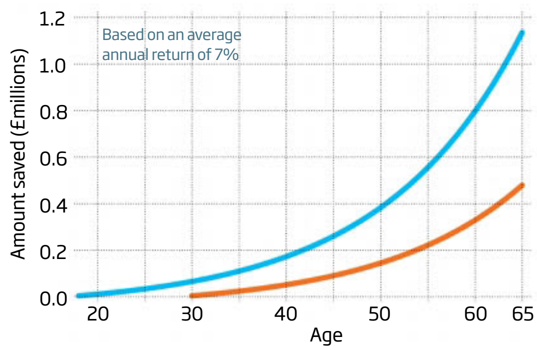
This sort of trick works so well because the existence of these hundreds of disappointed punters never occurs to us. "It's very difficult to count all the times something could have happened and didn't," says David Spiegelhalter, a statistician at the University of Cambridge.

Another way this plays out is in the birthday paradox. You're at a party with 23 guests and are surprised and delighted to find that two guests share a birthday – what a coincidence! In fact, a little calculation shows you the odds of this happening are better than 50:50. Again, the crucial point is the number of possibilities. One person in the room can share a birthday with any of the 22 others, and there are 365 days on offer in a standard year for this to happen. A second person can also share a birthday with any of the others, and so on. Continue adding up the possibilities for each person, and you end up with enough to make a shared birthday more likely than not.

We have this blind spot for large numbers for good reason. Being hypersensitive to "coincidences" was handy in our evolutionary past. "By following coincidences we make important discoveries," says Tom Griffiths, a cognitive scientist at the University of California, Berkeley. We first had to realise that the sun rises every day, and search for an underlying cause, before we could (eventually) conclude that Earth rotates. Equally the recurrence of stomach pains and worse taught us to stay away from certain berries.

Start 'em young

A short delay in saving for your pension can have a huge impact on how much you accrue



£1.134 M
Saving £250 a month starting at 18

£478 K
Saving £250 a month starting at 30

until you're 30 to start saving that £250 and you'll only reach about half as much (see diagram, above). Starting at 30, you would actually need to save more than £600 a month to make it to £1 million by 65.

Exponential growth's stealth factor is nicely illustrated by the story of the man who invented chaturanga, an Indian precursor to chess. He presented his king with a beautifully laid out board divided into 64 squares and when asked to name his reward, requested a grain of wheat to be placed on the first square, two on the next, four on the third, and so on. It sounded a modest reward, but had the king obliged across the board, he would have given away more than 18 billion billion grains. Fail to understand exponential growth, and our debts can rapidly spiral out of control too. This is an engine for creation and destruction wrapped up in deceptively simple maths.

In reference to the chaturanga legend, US futurist Ray Kurzweil refers to the sudden changes that spring from exponential growth as the "second half of the chessboard". The number of transistors that can fit on a electronic chip provides an example. Over the past few decades, it has roughly doubled every 18 months, a phenomenon known as Moore's law. The accelerating effect of exponential growth explains why we spent 25 years with bulky desktop computers before rapidly switching over to sleek smartphones. Kurzweil is famed for believing that this sort of technological growth will lead to an event called the singularity, when computers will become powerful and smart enough to improve themselves and outpace us all.

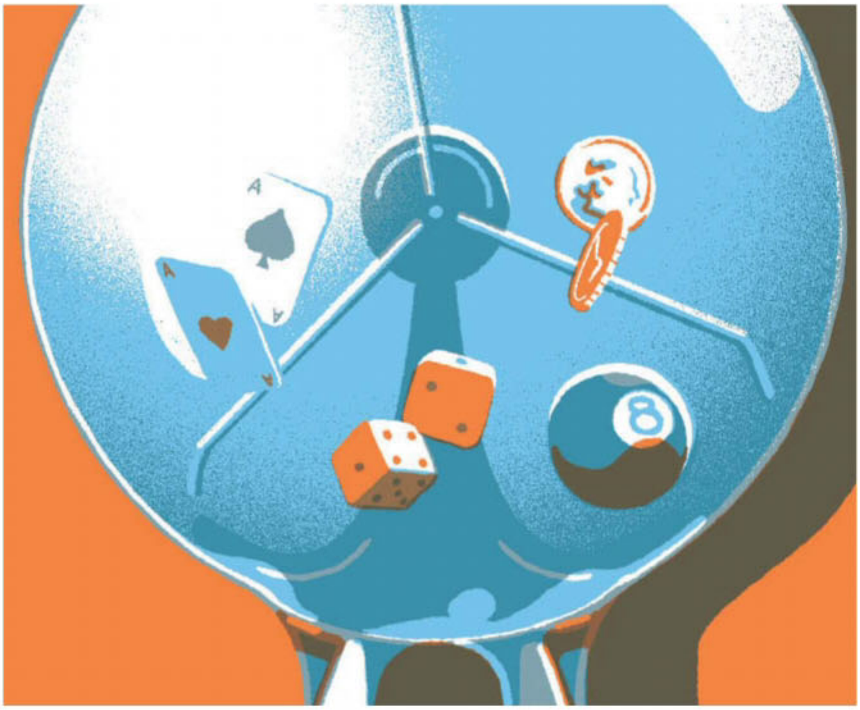
The spread of viruses often works in a similar way: one ill person infects a few others, who in turn each infect a few more, until we've got an epidemic on our hands. Immunisation acts as the limiting factor, which is why the world scrambled to treat last year's Ebola outbreak, which at one point saw the number of known cases doubling every few weeks.

When it comes to exponential growth, you can't trust your short-term instincts. Whether it's finance or technology, the largest changes won't happen for some time. But when they do, your whole world can be turned upside down in an instant. **Jacob Aron**

"We just happen to live in a world where most of the relevant causal relationships have already been discovered," says Griffiths. The result is we see patterns where none exist. Our problem with coincidences today, then, is a relic from a simpler world, like cognitive bias (see page 13).

Looking past what our gut tells us, at the often hard truths revealed by probability, calculations can also help us in balancing the risks we all face in our daily lives, from not taking an umbrella to not taking out an insurance policy. That probability is so useful is no coincidence. **Gilead Amit**

"Given a large enough sample, any improbable thing is eventually bound to occur"



NEED TO KNOW

GAME THEORY

The science of strategic thinking

IN THE film *A Beautiful Mind*, John Nash and his buddies, all of them graduate students in mathematics at Princeton University, are sitting in a smoky bar when a group of women walk in. As the men tease each other about their chances, Nash is struck with inspiration. Is there a logical, mathematical way of working out the best strategy for each man getting a date? Next thing you know he's shambling out of the bar, and spends the night furiously scribbling unfathomable-looking equations.

It sounds a little crass, and the episode probably never happened in reality. But in a ham-fisted, Hollywood sort of way, it does hint at how game theory, the branch of mathematics Nash helped to make famous, can apply to our everyday lives.

In fact, we use it all the time without even realising. "Every time you think about what you should do in terms of what someone else will do in response, you're doing rudimentary game theory," says Kevin Zollman of Carnegie Mellon University in Pittsburgh, Pennsylvania.

The trouble is, we are novices. When we need to think through situations several steps ahead or when they involve more than just a few people, we start to make mistakes. But delve into the theory just a little – there's no need to be a maths whizz – and you can harness some of the insights to make smarter moves in your own life.

Lesson one is that there are different sorts of games. Broadly speaking, there are zero-sum games, in which one player gains what the other loses, and variable-sum games, in which players have both common and opposed interests.

An example of a zero-sum game would be chess or poker. When you win, your opponent automatically loses and vice-versa. These sorts of situations don't crop up much in everyday life. Variable-sum games are more common and more complex. They are exemplified by what's known as the prisoner's dilemma, a scenario in which the punishment you receive for a crime depends on both your plea and that of an accomplice. You don't know how your accomplice will behave, but game theory organises the possible outcomes into a pay-off matrix that allows you to think through the various possible outcomes (see "Decisions in the frame", below).

It turns out, perhaps counter-intuitively, that your best option for both you and your accomplice is to confess. This decision is what's known as a Nash equilibrium because neither party can benefit from making a different choice while the other party's choice stays the same.

"There is another way to win at chicken: throw the steering wheel out of the window"

Nash died in 2015, but his contribution to game theory, including the equilibrium idea, helped him win a share of the 1994 Nobel prize in economics. Lessons from the discipline have been applied all over the place, from politics and diplomacy to economics and business. It helped the US formulate its nuclear deterrent strategy during the cold war, for instance. Today, broadcasters use it to jostle for the rights to air top-level sports fixtures.

But individuals can harness insights from game theory, too. One example is understanding the power of "credible commitment", says Rakesh Vohra, an economist at the University of Pennsylvania in Philadelphia. This concept is best described by a game of chicken. Think of two cars accelerating towards each other; the loser is the one who swerves. Here, the Nash equilibria are the two situations in which one player swerves and not the other.

But a game theory analysis shows there is another possible way out. One of the drivers can force an outcome by changing the rules of the game – for example by removing the steering wheel and throwing it out of the window. Then the other driver must swerve to avoid destruction. "You're making your opponent recognise that you have no choice but to take a particular action, which then forces them to do what you want them to do," says Vohra. "Paradoxically, limiting your options can sometimes make you better off."

The same principle can be applied to buying rather than crashing a car. Do your research on prices and make a take-it-or-leave-it offer. "By committing yourself, you force the seller to make a choice: either sell at that price or make no sale at all," says Vohra. This reasoning applies to any situation in which two competing parties have to negotiate a price, including agreeing a salary for a new job, for instance.

Be warned, however: even the greatest game theorists won't always get it right. The problem is that game theory assumes we act rationally all the time – and we don't (see "Cognitive bias", page 13). Even the experts will sometimes be thrown off by the quirks of human behaviour. **Daniel Cossins**

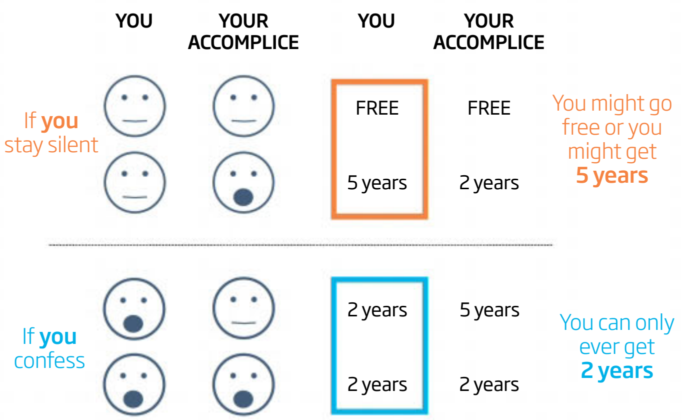
Decisions in the frame

Game theory provides us with a framework to make decisions when we have incomplete information – such as in the famous prisoner's dilemma

You and an accomplice are being held separately for a crime. The maximum penalty is 5 years – or 2 if you confess

The police can only prosecute if one of you confesses – if neither does, you walk free

Should you stay silent or should you confess?



TRENT PARKE/MAGNUM PHOTOS

NEED TO KNOW

HIDDEN ENERGY

Say goodbye to greenwash

SOMETIMES you want to make a decision that helps the planet. Maybe you're selecting an electricity supplier, choosing whether to sign a petition against wind farms or wondering whether to install solar panels. The right option might seem obvious: who could argue with the green credentials of a solar panel, for instance? But such decisions are often harder than you think.

Frequently the problem comes down to hidden energy. Energy is most obvious when it is kinetic, producing a visible effect such as when you kick a ball or wrench open a door. But energy is also needed to make things, and is locked up, or "embodied", in all sorts of manufactured matter, from a metal pipe to a slice of pizza.

Take buying a car – a big purchasing decision that could have a more significant impact on your carbon footprint than most. Say you've decided to swap your old petrol-powered banger for a new, fuel-efficient version. If your current car was manufactured a decade ago, it might typically pump out about 1.8 tonnes of carbon dioxide a year, if you drive an average sort of distance of 12,000 kilometres. If you were to buy the most fuel-efficient model you can find on garage forecourts today, and drive a similar sort of distance, you would emit about 1 tonne annually, according to company-declared emissions at least.

That seems like a worthwhile saving, but crucially we've yet to consider the energy embodied in the car. It takes energy to run the machines that built it and to produce the materials that form it, and this also generates CO_2. According to manufacturers' figures, the process of making a car typically takes between 600 and 800 kilograms of CO_2. Factor in making the steel for the body itself, and you can add in another tonne of emissions. Add in the carbon footprint of the aluminium components – which require five times more energy to smelt than steel – plus upholstery, glass, rubber and electronics, and your new car clocks in at around 6 tonnes of CO_2, according to the UK's Carbon Trust. You'll have to run it for at least eight years to recoup its carbon cost.

Faced with such figures, you might consider going the whole hog and buying an electric car. Here the same sorts of embodied emissions will be involved as with the petrol car, but at least there are no emissions from the exhaust, right? Sure – but if, when you plug the car into the mains, it is sucking power from a carbon-belching coal-fired power station, it has merely shifted its emissions elsewhere.

> "They might seem easy, but few green choices are as straightforward as they first appear"

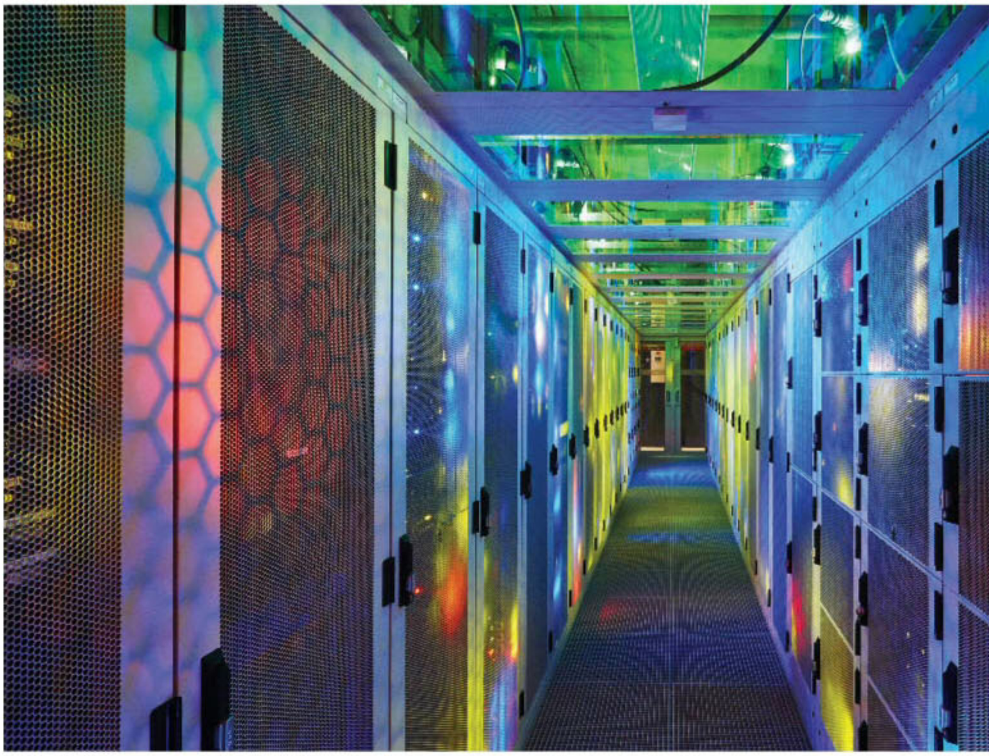
So beware: manufacturers will often try to play on your sense of environmental responsibility, but few green choices are as straightforward as they first seem. With a solar panel, what is the emissions cost of making the slab of highly purified silicon it's probably made of? Or of the huge magnets that help harness the energy from a wind turbine?

Asking such probing questions can give you a whole new perspective on a range of decisions, for example what you eat. Say you live in the UK, and you have a choice of a locally grown tomato or one grown in Spain. It's common to think the green option is to limit the food miles travelled and go for the local produce. But the embodied emissions associated with heating the hothouse in which the UK tomato was grown might easily – depending on how the Spanish tomato was freighted in, by air or by lorry – trump the embodied emissions associated with transport.

Such considerations rarely make our decisions easier. If you opted to ditch the new car entirely and cycle to work instead, for instance, what is the carbon cost of shipping in all those extra veggies to power your leg muscles? But then no one said saving the planet was easy. **Fred Pearce**

INTERNET ARCHITECTURE

Where your data goes determines how much you can hide



MARK HORN/GETTY

Around the world in 80 microseconds

You use Google's email service Gmail to send a cute baby pic to your mother-in-law a few miles away, who then posts it to Facebook. That data will be stored on server farms like the one pictured above – but where? Starting from *New Scientist*'s London, Boston and Sydney offices, it goes something like this

FROM LONDON...
The Gmail pings around London then speeds under the sea to be stored in Dublin, Ireland, where Google has a huge data centre. When your mother-in-law posts the picture, it lands at Facebook's European data centre in Lulea, Sweden – a country that gives itself permission to investigate any data crossing its borders. It is also backed up at Facebook's original data centre at Prineville, Oregon.

FROM BOSTON...
After ping-ponging across the US – but never leaving it – the photo lands at Lenoir, North Carolina, the closest of Google's seven US data centres. It is also stored at Prineville, and backed up at a couple of other locations. Gmail encrypts data, so the average hacker is probably not a problem on this circuitous route; the US National Security Agency, which maintains "backdoors" to US technology company servers, perhaps more so.

SNAP. You press the shutter icon on your phone and capture a photo of your baby daughter. With a couple of swipes, you attach it to an email in your Gmail app and fire it off to your mother-in-law.

As personal data goes, it doesn't get much more innocuous. But the truth is that spraying around any private information is risky. You might think that's overblown. As long as you have nothing to hide, you have nothing to worry about.

It's not that simple. Just look at the 2015 hack that exposed the data from Ashley Madison, a site catering for people looking for an affair, and imagine if the same happened with all your emails stored by Google, or your photos on Facebook. Even if you've done nothing illegal or immoral, faced with a database of every photograph and comment you've ever shared privately, friendships and business deals could dissolve the world over.

And there are plenty of vulnerabilities. The material displayed on the web is stored, often in central server farms. Whenever you upload text or pictures, they are ferried to these farms by cables. Although there are safeguards, data can in theory be hacked, stolen or altered at many points along the way.

Let's return to that photo, and imagine that once your mother-in-law receives your email, she immediately uploads the baby pic to Facebook. The likelihood is that even such an everyday occurrence will send information pinging on unexpected routes around the world, often leading it to be stored in places with unfamiliar privacy laws (see "Around the world in 80 microseconds", right).

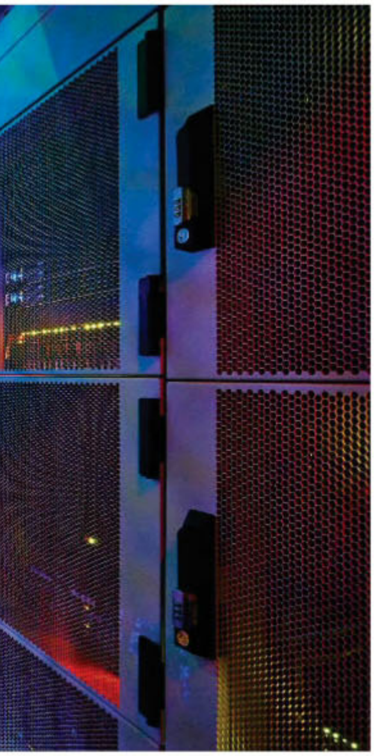
Bear in mind, too, that companies such as Google and Facebook will often store many copies of your data. Andreas Olah, an analyst at GlobalData Technology in London, says Facebook copies your data into formats readable by all sorts of devices and creates backup copies. If your mother-in-law had recently been on holiday to the US, for example, Facebook would probably have sent the photo to a US data centre too, just so that she wouldn't have wait for it to load were she to visit again.

In short, one simple share can create reams of potentially hackable data. This state of affairs demands caution from all of us, says Judith Lewis, a consultant based in London who previously worked on email security. "Just pretend you have a stalker who sees everything."

Studies show that people with more complete mental models of the physical internet also have a fuller understanding of privacy risks. Unfortunately, the same research found that this understanding leads to "security fatigue" – we get bored with worrying.

What can we do with this knowledge, anyway? A common refrain is that you should encrypt everything. Encryption has traditionally been hard to learn and required enthusiastic adoption not just from you, but everyone you communicate with. "If we want privacy to be protected, the only way to do it is collectively," says Alessandro Acquisti, a cyberprivacy researcher at Carnegie Mellon University in Pittsburgh, Pennsylvania.

Encrypted web communication protocols such as the https system commonly used for internet shopping and banking, and those that underlie some popular messaging apps, have made that easier. But a better solution might be to include protection in the core design of devices and online services. The Blackphone, released in



FROM SYDNEY…
Neither Google nor Facebook maintain data centres in Australia. The email is stored in Singapore or Taiwan; the western US is again the destination of the Facebook photo. Many nations don't like data straying outside the protection of their own privacy laws. A 2015 European court ruling might mean US firms will no longer be able to whisk data across the Atlantic – perhaps paving the way for countries to insist on having data centres on their turf.

TRACE YOUR OWN DATA

Data you produce online will take various routes. Follow these steps to start tracking

1. Open the terminal application on your computer
   Mac: Applications > Terminal
   Windows: All programs > Accessories > Command prompt
2. Type "tracert yyy.com" and hit enter where yyy.com is the website you are visiting
3. Watch as your route through the internet is traced

The long numbers are the IP addresses of the computers that are routing your data. The strings of letters are the machines' hostnames, indicating which company is running the computer. Buried in the hostname are airport codes, indicating the cities each machine is in.

September 2015 and powered with Google's Android software, allows you to use any app you like but feeds it blank fields instead of the data it expects.

At launch the Blackphone cost $799, and Google still gets your data, even if third parties don't. Plus, it doesn't allay the worry that others might not take care with your data. "No matter how securely you share something, the recipient can still be stupid with it," says Lewis. A basic understanding of how the internet works should give us pause for thought in how much we share. **Sally Adee**

NEED TO KNOW

EPIGENETICS

Your lifestyle can change your genes

DID you hear the one about how the giraffe got its neck? Aeons ago there was an animal that walked along a dusty path to a watering hole every morning. Halfway, she would spot a patch of trees with the tastiest, most succulent leaves on the savannah. Stretch as she might, she couldn't quite reach them. Then one day the stretching paid off, and she suddenly had a mouthful of juicy fronds. Years passed, and the giraffe had babies. Over the generations they became spindlier and spindlier, reaching ever higher into the treetops.

It could be one of Rudyard Kipling's *Just So Stories,* and you can see its charm: keep studying, training and eating healthily, and you can change yourself for the better. Not only that, but your efforts will endure and future generations will benefit, too.

It's just a shame that scientists from Darwin onwards have said that this picture is flat-out wrong. Random mutations in DNA, corralled by the forces of natural selection, fuel evolutionary change.

There's no change in that basic picture, but some research suggests that elements of our giraffe story might not be so wide of the mark after all. Far from being a rigid instruction manual, our DNA is flexible and responsive – and we might be able to change much more than we thought.

Almost all your cells hold the same 20,000 or so genes, but each type of cell uses a unique suite of them. Genes have to be turned on and off at the right time in the right place. The mechanisms by which this happens are referred to as "epigenetic", acting over and above the genetic code.

Epigenetic information is written into our genes through a series of biological marks that don't affect the underlying DNA sequence. A dazzling array of chemical tags can be stuck onto the proteins that package DNA, or even DNA itself, masking certain parts of our genomes from the cell's gene-reading machinery or making them more enticing. Crucially, our cells can rub out and rewrite these epigenetic marks to turn genes on or off.

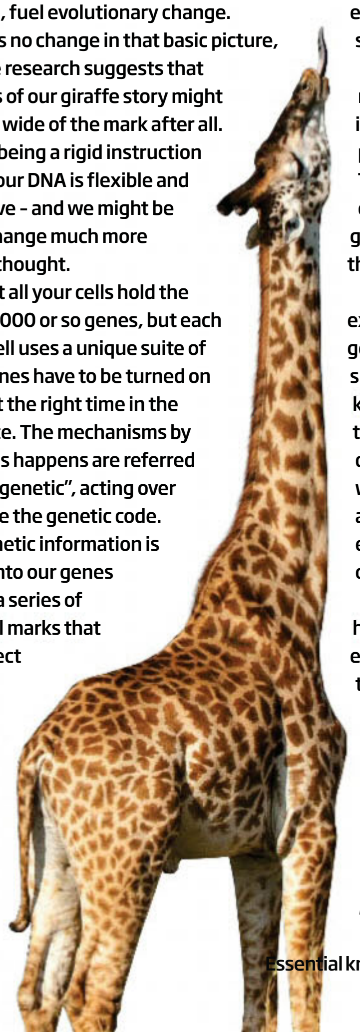
Munching on broccoli, supping on green tea and exercising are just a few activities that have been linked to changing our epigenetic marks for the better. And it works the other way too: some researchers lay the blame for ailments ranging from allergies to cancer to Alzheimer's disease at the door of dodgy epigenetic switches.

Things start to get really interesting when we ask whether epigenetic changes can be passed down the generations. If true, it would mean that the behaviour of a parent could affect how their babies, and even those of their children, develop, and even influence the diseases they might succumb to later in life.

For this to happen, epigenetic marks would need to be written into the DNA of eggs or sperm and preserved in the resulting offspring. There are mechanisms to wipe the epigenetic slate clean between generations, but some studies hint that this doesn't always work perfectly.

And there are some intriguing examples of the experiences of past generations seeming to affect subsequent ones. In a famous case known as the Dutch Hunger Winter, the children of women who conceived during the famine caused by the second world war were born smaller than average – as were their grandchildren, even though they were living at a time of post-war plenty.

For now, solid examples of this happening in humans are rare, but the evidence is growing. It's certainly one to keep your eye on: if your gym habit, penchant for pizza or caffeine addiction is going to have cascading effects on your unborn children, you ought to know about it.



DENIS-HUOT/NATUREPL.COM

Kat Arney is the author of *Herding Hemingway's Cats: Understanding How Our Genes Work*

WHAT ARE YOU WORTH?

Each life is priceless.
Except when it's not, finds Shannon Fischer

IF YOU had to put a price on your life, could you do it? What would it be? Where would you even start?

We may think valuing human life this way is the stuff of darkest history, now confined to the malevolent underworld of human trafficking. We look with shame to an era when a human being could legally be bought and sold, their worth tied solely to the profit their work would yield. In the mid-19th century, before slavery was abolished in the southern states of the US, a "prime male field hand" could be purchased for about $1100 – roughly $30,000 in today's money. Other human beings were bought and sold for far less.

Our repulsion at the idea of putting a monetary value on people is consistent with the modern principle, outlined in documents such as the UN's Universal Declaration of Human Rights, that all human lives are equal – and, we like to believe, equally priceless.

Yet we routinely trample on those exalted ideals. The scientific literature and the news are both rife with examples of how unequally we value life – young over old, those like us over those who are different, the identified victim over the faceless masses.

We don't just value lives differently in a moral sense, but in real money. It is how we divvy up limited resources – from deciding how much to invest in building safer roads to setting compensation for families of soldiers and civilians killed in war, or those who were wrongfully imprisoned. And depending on who is doing the pricing and why, those numbers can vary, a lot. Life, it turns out, doesn't have a price. It has a hundred.

One of those numbers is a calculation of how much should be spent to prevent your death. To decide which potentially life-saving interventions are worth it, government bodies look at a quantity called the value of a statistical life (VSL), or as the UK Department for Transport (DFT) puts it, the value of a prevented fatality. "This is not the amount of money people would accept in exchange for certain death," says W. Kip Viscusi at Vanderbilt University, Tennessee, who helped introduce the VSL to US agencies. "It's really just a reflection of their attitude toward a very tiny risk of death."

Put simply, it's the type of calculation we make when deciding whether it's worth spending extra to buy the car with more safety features, just on a grander scale. Take the risk of dying from salmonella infection, for instance. If people are on average willing to pay $7 to reduce that risk by 1 in a million, then the VSL is $7 million. This would then be the figure used by the US Food and Drug Administration (FDA) to justify the cost of efforts to prevent salmonella outbreaks.

The VSL a country adopts tends to vary with its wealth (as a benchmark, the Organisation for Economic Co-operation and Development recommends that member nations use a


PATRIK SVENSSON

The biggest **life insurance policy** on record, bought by a mystery Silicon Valley billionaire

$201 million

figure between $1.5 million and $4.5 million).

Then there is the matter of how you determine what people would be willing to pay for a given reduction in their risk of dying. In the US, where the average VSL works out at about $9 million, economists make the calculation mostly by looking at what people do – wages someone accepts to take a risky job, for instance. In the UK, where the preferred technique is simply to ask people what they would be willing to pay, the DFT uses £1.8 million ($2.3 million). A 2009 report that compared the two approaches in Canada found that when the VSL is based on wages and risk, it's worth about a third more. Based on reviews of both methods, Australia's Office of Best Practice Regulation recommends a VSL of A$4.2 million ($3.2 million).

VSL also varies with the cause of death being considered. In the US, it has ranged from $200,000 up to more than $13 million: reducing the risk of workers dying in a coal mine was worth more than reducing the risk of death in a fire caused by your flammable couch upholstery.

When you get into that other form of death prevention, healthcare, things only get more complicated. To decide whether medical interventions are worth it, healthcare providers and insurance companies consider how much decent-quality life you might get for the money. The measure they use is called a Quality-Adjusted Life Year (QALY): if 1 is perfect health and 0 is dead, then four years in middling health equals two QALYs.

In the UK, one good year of life is worth around £20,000 to £30,000. That threshold is set by the National Institute for Health and Care Excellence (NICE), which oversees what new drugs or treatments the UK's National Health Service can provide. To do this, it looks at the cost per QALY of new treatments compared with that of existing care.

If a new drug offers one extra QALY for every additional £20,000 spent, that's well within budget. "NICE focuses on that increment," says Karl Claxton, a health economist at the University of York, UK, and a member of the NICE appraisal committee from 1999 to 2012. "What are the additional benefits, what are the additional costs, are they worthwhile?"

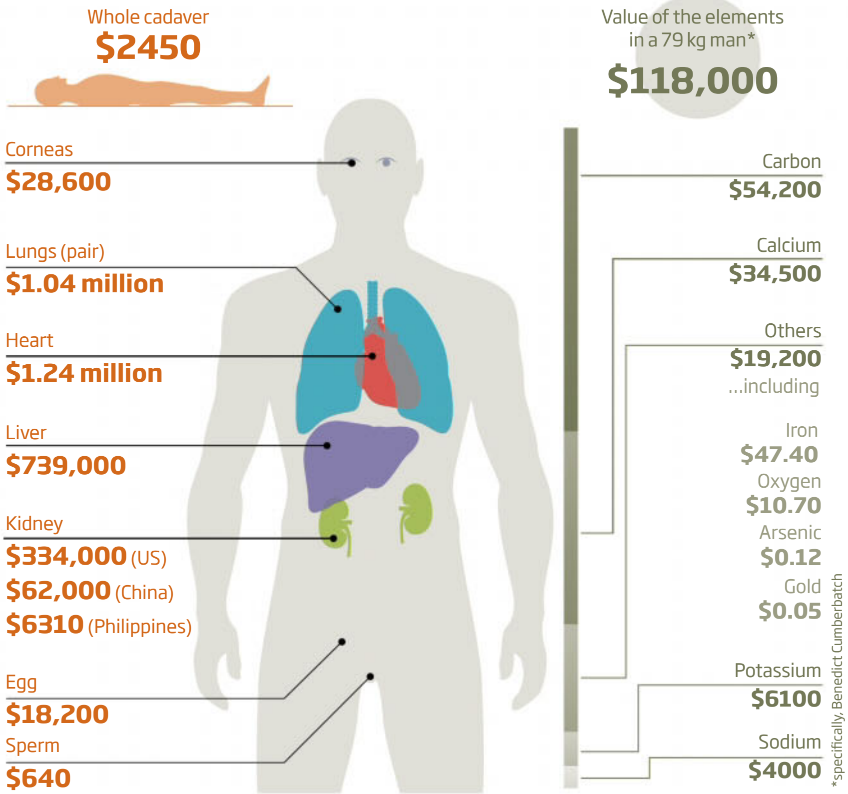
Although NICE does make exceptions for particularly innovative drugs or end-of-life care, if costs start to push past £30,000 approval becomes much less likely. Sofosbuvir (sold under the name Sovaldi), the wonder drug for patients with hepatitis, made the cut. But bevacizumab (Avastin), a drug that could give certain cancer patients about three more months when given together with chemotherapy, did not. It cost at least £82,000 per QALY.

The threshold NICE uses provokes controversy from time to time. That is in part because, as former NICE chair Michael Rawlins once conceded, it is not based on "empirical research" so much as "the collective judgment of the health economists we have approached across the country". Since it was first put into use in 1999, there have been efforts to pin it to more robust research. But it remains as is, unchanged even for inflation.

In countries such as Canada and New Zealand there is no explicit threshold. But when you analyse resource allocation decisions, in practice, it works out at roughly $15,000 per QALY. Contrast that with the situation in the US where, so long as an intervention is deemed "reasonable and necessary", then the government-run health plan Medicare will not

The sum of your parts

Valuing a human life is one thing – but what about bodies? Whether it's a cadaver for medical research, a heart for an organ transplant or the elements we are made of, it's possible to work out a price



SOURCES: ANATOMICAL ASSOCIATION OF ILLINOIS; MILL MAN; THE RED MARKET BY SCOTT CARNEY; ROYAL SOCIETY OF CHEMISTRY – VALUES OF ELEMENTS BASED ON PRICES FROM HIGH QUALITY CHEMICAL SUPPLIERS

Price per day to defend the lives of US presidential candidates

$40,000

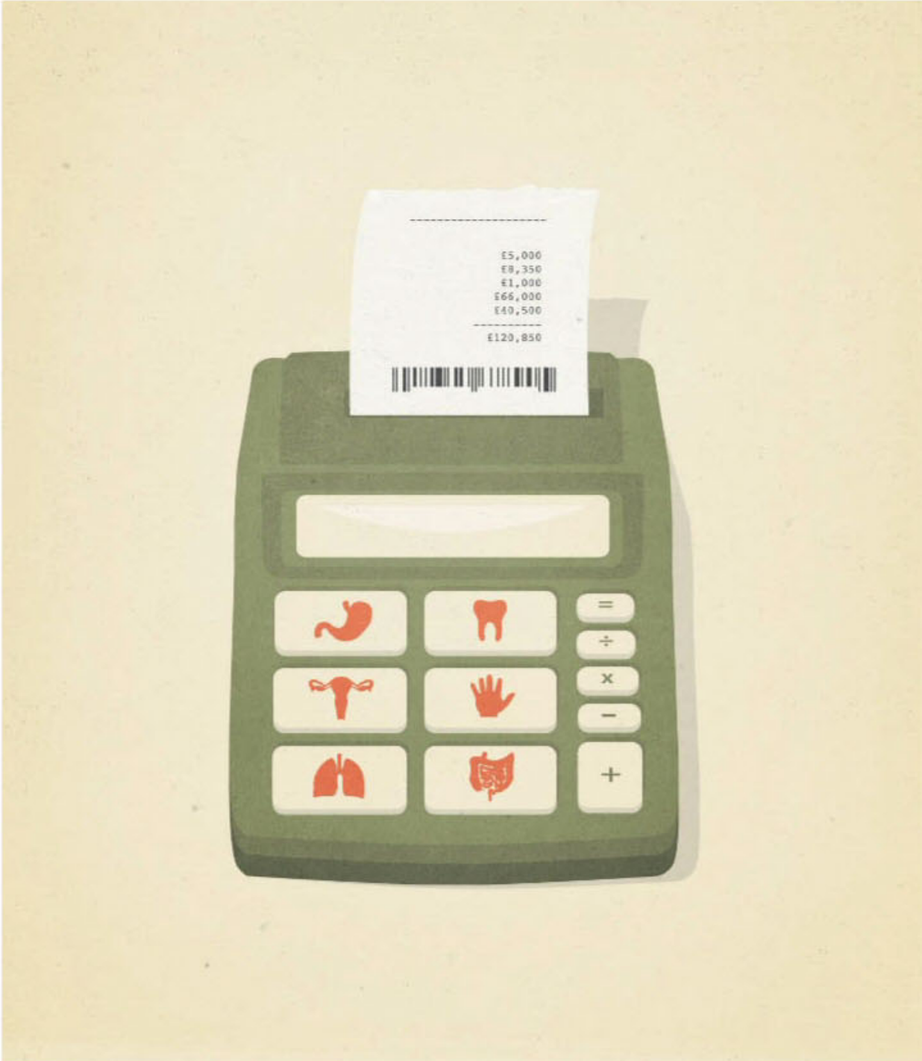
consider the cost. When you take into account that this could happen with a third-line cancer treatment costing $900,000 per QALY, it's no wonder that the US spends nearly a fifth of its GDP on healthcare. Private insurers are not held to the same standards, and many clearly do consider expense – the more cost-effective a drug, the less their customers may have to pay out of their own pocket, for instance. But these systems are not always transparent and Medicare is still by far the largest healthcare funder in the country.

Medicare does try to trim costs – or at least maximise benefits – in some ways, says James Chambers at Tufts Medical Center in Boston. They may only pay for drugs and devices for the sickest patients, for instance. But, legally, there is no straightforward way to do it. "There has been a view that you cannot, you cannot set a price on life," says Louise Russell, an economist at Rutgers University in New Jersey who specialises in healthcare policy. "Which really comes down to, you can't admit you have to."

Valuing someone's life gets even harder once they are dead. Some guidelines exist when mortal danger is part of the job, but a huge number of variables remain. Members of the US military who die on active duty are entitled to a tax-free "death gratuity" of $100,000, a life insurance payout of up to $400,000, and a host of other benefits from burial costs to money for their children's healthcare and education. By some estimates, the total can range from $250,000 to more than $800,000. The breakdown is similar in the UK, although much depends on the person's salary and age at death.

For police officers and firefighters, it is a similar patchwork of pensions, workers' compensation, life insurance, union benefits and dedicated state and federal funds. One programme run by the US Department of Justice gives families a sum of $339,881.

Outside of the line of duty, unexpected deaths reveal even more inconsistencies in how we value life. If compensation is settled in a wrongful death lawsuit, there's a purely economic component that works out logically enough, based on estimates of the victim's ›



PATRIK SVENSSON

INHUMAN BUSINESS

There are many legal ways we put a value on life (see main story), but there are plenty of illicit ones too. In cases of human trafficking, prices are often set based on the risk involved, transportation and labour costs and also market conditions. A child purchased in rural Africa for about $200 might later be sold for a far higher price, says Benjamin N. Lawrance, a historian and anthropologist at the Rochester Institute of Technology in New York. That will depend on how many hands the child passed through on the way to say, mine shafts in another country, how many documents may have been forged for transport, and how many other desperate children arrived around the same time. Philosophical notions of human value have nothing to do with it.

Similar factors influence kidnapping ransoms. According to estimates from Terra Firma Risk Management, which advises families and employers of kidnap victims on conducting negotiations to free the hostages, a hasty abduction on the streets of Venezuela by criminals who put in the bare minimum of planning might net just a few hundred dollars. A more elaborate operation by syndicates and corrupt officials targeting a very wealthy individual might land a six-figure sum.

"They're looking at the hostages as commodities, like a businessman would look at something on his shelf," says a spokesperson from Terra Firma, who asked to be kept anonymous. "The basic procedure is, they call and make a demand, then the other side makes an offer – a lot less – and all of a sudden, that victim is worth a certain amount of money."

lost lifetime earnings, medical costs, funeral expenses and so on. But when it comes to grief and lost companionship, it's all over the place. English courts limit bereavement awards to £12,980 total, far below NICE's lower threshold for the value of one good year of life. This is derived from an early figure of £3500 set by Parliament in 1982. "It's very arbitrary," says Laura Hoyano, a human rights law specialist at the University of Oxford. "And it's the very low level of it that's probably the most insulting."

In your absence

In the US, compensation can vary by all sorts of factors, including the nature of the death, the amount of insurance held by those responsible, and even whether the sum is set by a judge, jury or two lawyers across a table.

There is a lot of inconsistency, says Mark Geistfeld at the New York University School of Law. "How do you figure out how much my spouse suffers by way of loss of companionship if I die prematurely?" he asks. "Is that $100,000 or $100 million?"

He says that judges tell juries there is no set way to do it, so jurors tend to look for some reference figure. "If somebody suffered $100,000 in medical expenses, maybe we'll triple it, we'll use that for an anchor, we'll say pain and suffering will be $300,000."

Local laws also make a big difference. The parents of 6-year-old Brandon Holt were compensated $572,588 after he was shot dead by another child in 2013. This was in New Jersey, where the wrongful death statute doesn't allow juries to take the family's emotional distress into consideration.

In contrast, the wrongful death of 12-year-old Tamir Rice, who was shot by police in 2014 while holding a toy gun, was settled at $6 million. That case was decided in Ohio, where juries may consider mental anguish. The social context mattered here, too: Rice's death was part of the broader controversy about black lives and the police.

Ultimately, though, these cases did not attempt to value the lost life itself – nor does the law require it. Anthony Sebok studies these types of lawsuits at the Benjamin N. Cardozo School of Law in New York. "There's damages for pain and suffering before you die, there's damages for loss of income to your family after you die. But for the life itself that was lost, it's worth nothing."

That's why the context makes so much difference. After 9/11, the 2007 Virginia Tech massacre and the Boston Marathon bombing in 2013, attorney Kenneth Feinberg was tasked with distributing funds to survivors and families of those killed. The total amount available was different in each event, but individual payments weren't meant to reveal anyone's fundamental worth. Instead they were an exercise in demonstrating patriotism, strength and the compassion of a people.

After 9/11, the US government established a fund for injured victims and the families of the nearly 3000 dead. By Congressional mandate Feinberg had to follow certain aspects of wrongful death law, so he awarded funds partly based on the victims' incomes; the next of kin of CEOs received more than those of janitors. But there was also a non-economic portion, based on a flat rate: $250,000 per death, plus $100,000 for each surviving spouse and dependent. Payouts ranged from $250,000 to $7.1 million.

"Congress wanted to demonstrate to the world its empathy and support for the victims," Feinberg explained in his book, *Who Gets What: Fair compensation after tragedy and financial upheaval*. The programme was "proof positive that Americans stood together, a single community ready to help one another in our collective hour of need".

The money for the Virginia Tech and Boston Marathon victims was donated by individuals and businesses – "evidence of citizen compassion", as Feinberg put it. So different rules applied: all lives were treated equally. Families of victims of the Virginia

FOREVER YOUNG

When it comes to what philosophers call negative rights - the right not to be killed, chief among them - studies show that at a fundamental level we do believe all humans are equal. But for positive rights, such as the right to be saved, we aren't so even-handed. (See main story.)

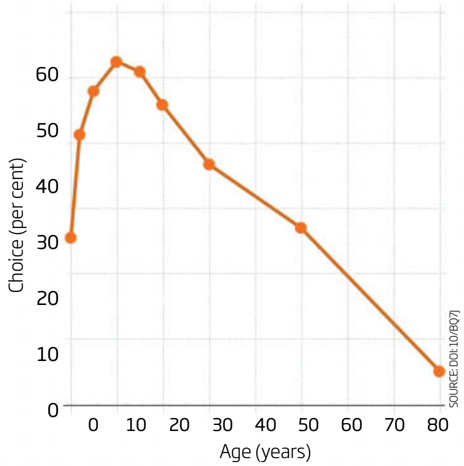
If there is only one dose of a life-saving treatment but there are two people who need it, how do you decide who gets it? University of Pennsylvania psychologist Geoffrey Goodwin and his then student Justin Landy set out to answer that question. When volunteers were forced to make a series of choices between people of different ages, a clear, if unsurprising, trend emerged: we tend to favour the young over the old (see graph right).

Why? When asked to justify their choices, people made what is known as the fair-innings argument. "The basic concept is that everyone should have their fair shake at life," says Landy. "The 50-year-old has gotten 50 years already, so we should save the 10-year-old." Another way to frame it is the years-left argument - the younger person has more years of life ahead.

But that reasoning only goes so far, because it turns out that our most highly valued humans aren't newborn babies. That is because you develop more - and more profound - social connections as you age. There is also something of a "sunk cost" effect: "Older children have had more work put into them by their societies - the payoff has not come yet," says Landy.

The net effect is that we view adolescents or young adults as most worth saving. And, as they say, it's all downhill from there.

In an experiment, volunteers made choices about pairs of people who were different ages - which of the two should receive a life-saving organ, or whose death would be more tragic, for example. The graph shows the percentage of instances in which each age was prioritised



PATRIK SVENSSON

Compensation for **27 years** unjustly spent behind bars in British Columbia, Canada

US$6 million

Tech massacre each received $208,000; families of those killed in the Boston Marathon bombing each got $2.2 million.

In the UK, the Criminal Injuries Compensation Authority handles payments to victims of violent crime. Their baseline number for a death is £11,000 to one family member or dependent, or £5500 each if multiple people deserve payment, plus extra for funeral expenses, income loss and lost parental guidance. There's a £500,000 ceiling, but according to an investigation by the *Financial Times*, for victims of the London bombings in 2005, the highest award was nowhere near that: £141,050.

Then there is the matter of people who spend years of their lives unjustly behind bars. In the UK, there is no guaranteed compensation. Individual cases are assessed according to previous criminal record and lost income, among other things, but payment tops out at £500,000, or £1 million if more than a decade was served. That's if it's ever awarded. "They're extremely strict," Hoyano says. It's not enough for a conviction to be overturned – people essentially have to find new facts to prove their innocence to the Ministry of Justice. The evidence bar is so high that many receive nothing at all. When Victor Nealon was released from prison in 2013 after DNA evidence exonerated him from a conviction of attempted sexual assault, he had served 17 years. He got £46 in discharge money.

In New Zealand, there is no legal right to compensation, but awards are granted based on the merits of individual claims. In those cases, the starting point per year spent in prison is NZ$100,000 ($72,000).

In the US, different states have different standards. In New Hampshire, it's $20,000 total, no matter what. In Florida, exonerated people get $50,000 per year served, to a maximum of $2 million. Robert Norris, a wrongful convictions expert at Appalachian State University in Boone, North Carolina, says that part of the value of this kind of payment isn't the money itself – it is recognition from the state that a mistake was made, even if that doesn't amount to an apology.

"The view that you cannot set a price on life really comes down to, you can't admit you have to"

Western governments that compensate families of civilians killed by their armed forces in Afghanistan and Iraq use a similar rationale: the payments are not apologies so much as expressions of sympathy and regret. The UK Ministry of Defence paid £5600 to an Afghan man who lost his wife and son in a mortar bombing mishap; the German government paid €3800 to each of 102 Afghan families after a deadly bombing; the US paid $10,000 to the family of a brother and sister shot at a checkpoint in Iraq. "It's hard to digest that the value of a human life is a few thousand dollars," the retired US Army brigadier general Arnold Gordon-Bray said in an interview in 2013. "But you know that in their economic situation, it is the equivalent of much more, and you feel better."

It may be unsettling to think that the value we place on a human life shifts with political priorities, national boundaries and social context – that it differs depending on whether you are considering the cost of medication or safety belts. But just because it is difficult to place a fair value on human worth doesn't mean that the attempt itself is unworthy.

Consider Benjamin Franklin's "moral algebra". Nearly 250 years ago, Franklin wrote a letter to a friend facing a difficult decision. He recommended making a list of all of the pros and cons, and then striking out those on either side that seemed of equal importance. This early cost-benefit analysis was not meant to downplay the gravity of the decision, just the opposite. When "the whole lies before me, I think I can judge better, and am less likely to make a rash step," Franklin wrote.

The same idea applies today. Our reluctance to consider these most difficult calculations means we are left fumbling when we inevitably need to. "You make choices about what you spend on, there simply isn't enough to do absolutely everything," Russell says. "You can either make those choices with your eyes wide open, or you can do it with your eyes shut." ■

Why am I here?

Having a purpose to what you do could help you live longer – and better, finds Teal Burrell

SOMETHING to live for. This simple idea is at the heart of our greatest stories, driving our heroes on. It is the thread from which more complex philosophies are woven. As Nietzsche once wrote, "He who has a why to live for can bear almost any how".

As human beings, it is hard for us to shake the idea that our existence must have significance beyond the here and now. Life begins and ends, yes, but surely there is a greater meaning. The trouble is, these stories we tell ourselves do nothing to soften the harsh reality: as far as the universe is concerned, we are nothing but fleeting and randomly assembled collections of energy and matter. One day, we will all be dust.

One day, but not yet. Just because life is ultimately meaningless doesn't stop us searching for meaning while we are alive. Some seek it in religion, others in a career, money, family or pure escapism. But all who find it seem to stumble across the same thing – a thing psychologists call "purpose".

The notion of purpose in life may seem ill-defined and even unscientific. But a growing heap of research is pinning down what it is, and how it affects our lives. People with a greater sense of purpose live longer, sleep better and have better sex. Purpose cuts the risk of stroke and depression. It helps people recover from addiction or manage their glucose levels if they are diabetic. If a pharmaceutical company could bottle such a treatment, it would make billions. But you can find your own, and it's free.

The study of how purpose influences our health largely began with Viktor Frankl, an Austrian psychiatrist who survived four Nazi concentration camps. He noticed that some of his fellow prisoners were far more likely to survive than others. "Woe to him who saw no more sense in his life, no aim, no purpose, and therefore, no point in carrying on. He was soon lost," he later wrote. After the second world war, Frankl dedicated his work to understanding the role of purpose and developed a therapy based on his findings.

Beyond happiness

Today, researchers define purpose as a sense of direction in life – a long-term goal set around one's core values, that makes life worth living, and shapes daily behaviour. It is a component of broader measures of subjective well-being or happiness (see "How do you measure purpose?" page 30), in which there has been a surge of interest in the past two decades. That's why, in 2012, then United Nations secretary-general Ban Ki-moon commissioned the first ever *World Happiness Report*, which has been updated annually since.

Measures of happiness can reflect broader social issues such as inequality, but when researchers look at the individual elements that make up well-being, they find purpose on its own has a unique influence on health.

Of course, teasing out whether it is actually purpose itself, and not the fact that purposeful people may exercise more or eat better, can be difficult. But over the past 10 years, the findings about the health benefits of purpose have been remarkably consistent – revealing that, among other advantages, alcoholics whose sense of purpose increased during treatment were less likely to resume heavy

CARLO GIAMBARRESI

drinking six months later, that people with higher purpose were less likely to develop sleep disturbances with age, and that women with more purpose rated their sex lives as more enjoyable. These findings persist "even after statistically controlling for age, race, gender, education, income, health status and health behaviours", says Victor Strecher, a public health researcher at the University of Michigan in Ann Arbor and author of the book, *Life on Purpose*.

In an analysis of 7000 middle-aged people in the US, even small increases in sense of purpose were associated with big drops in the chances of dying during a period of 14 years. A study of more than 9000 English people over 50 years old found that – even after adjusting for things like education, depression, smoking and exercise – those in the highest quartile of purpose had a 30 per cent lower risk of death over nearly a decade compared with those in the lowest quartile. Other studies show higher purpose cuts risk of heart disease by 27 per cent, stroke by 22 per cent and Alzheimer's disease by half.

The only reason purpose isn't a top public health priority, says Strecher, is because it somehow feels too vague or ephemeral. "It's not a construct that feels scientific enough," he says. "If this were a physical issue or a new drug or a gene, you would see lots of funding going into it."

Some of the scepticism has to do with concerns that purpose is merely a stand-in for opportunity in life, or wealth. Indeed, in recent research, Patrick Hill, now at Washington University in St Louis, did find that people with a stronger sense of purpose tended to have more money to begin with, and earn more over the period studied.

Health benefits

But a 2007 Gallup poll of 141,000 people in 132 countries found that, even though people from wealthier countries rate themselves higher on measures of happiness, people from poorer nations tend to view their lives as more meaningful. Shigehiro Oishi at the University of Virginia in Charlottesville, who analysed the poll data, suspects this is in part because people in developing countries have more concrete things to focus on. "Their goals are clearer perhaps: to survive and believe. In rich countries, there are so many potential choices that it could be hard to see clearly," he says.

Could it be that purpose is just another term for religious faith? Oishi's study did find that nations with the highest ratings of meaning in life were also the most religious. And religious people do tend to report having more purpose. But efforts to disentangle the two have revealed differences. Religiosity doesn't predict a lower risk of heart attack or stroke, for example. And certainly many non-religious people have high levels of purpose.

In fact, few of us rank on the very low end of the scale. "We tend to focus on the utter meaninglessness of the world," says Samantha Heintzelman, also at the University of Virginia. But, "for the most part, people feel like their lives are pretty meaningful".

So how does that meaning, that sense of purpose, actually improve your health? In part, it may be because greater purpose makes people more conscientious about maintaining their health. But Steven Cole at the University of California, Los Angeles, thinks there's more to it. "If people are living longer, there's got to be some biology underpinning that," he says. Cole has spent years studying how negative experiences such as loneliness and stress can increase the expression of genes promoting inflammation, which can cause cardiovascular disease, Alzheimer's or cancer.

In 2013, Cole examined the influence of well-being instead. He focused on two types: hedonic, from pleasure and rewards, and eudaemonic, from having a purpose beyond self-gratification. These two aspects were measured by having participants note down their well-being over the previous week, how often they felt happy (hedonic) or that their life had a sense of direction (eudaemonic), for example. Although scoring highly in one often meant scoring highly in the other and both correlated with lower levels of depression, they had opposite effects on gene expression. People with higher measures of hedonic well-

HOW DO YOU MEASURE PURPOSE?

To determine whether purpose affects health and longevity, you first have to measure it. To do this, many researchers turn to a set of scales developed in the 1980s by the psychologist Carol Ryff at the University of Wisconsin in Madison.

Ryff's scales measure six different aspects of well-being: autonomy; environmental mastery (the feeling of being in control in your everyday environment); personal growth; positive relations with others; purpose in life and self-acceptance. For each item, people read a series of statements, and select one of six responses ranging from "strongly disagree" to "strongly agree". These types of scales are often used to assess national levels of well-being.

For purpose, the statements include things like, "My aims in life have been more a source of satisfaction than frustration to me", or, "In the final analysis, I'm not sure that my life adds up to much".

Higher scores are based on stronger agreement with purposeful statements such as: "Some people wander aimlessly through life, but I am not one of them", and disagreement with remarks such as: "I live life one day at a time and don't really think about the future".

People who score in the bottom 25 per cent are considered to have low levels of purpose. A person with a high degree of purpose – someone who falls into the top 25 per cent – is characterised as someone who "has goals in life and a sense of directedness, feels there is meaning to present and past life, holds beliefs that give life purpose and has aims and objectives for living".

BURT GLINN/MAGNUM

ALEX WEBB/MAGNUM

Full circle: having goals that benefit others may provide particularly strong benefits for you

being had higher expression of inflammatory genes and lower expression of genes for disease-fighting antibodies, a pattern also seen in loneliness and stress. For people scoring highest on eudaemonia, it was the opposite. "There were surprises all around," Cole says. "The biggest surprise being that you can feel similarly happy but the biology looks so notably different."

"If people with purpose live longer, there must be some biology underpinning that"

Cole suspects eudaemonia – with its focus on purpose – decreases the nervous system's reaction to sudden danger that increases heart rate and breathing and surges of adrenaline. Over-activation of this stress-response system, as you see with chronic stress, causes harmful inflammation. "There may be something saying 'be less frightened, or less worried, anxious or uncertain'," says Cole.

That something could be a brain region called the ventral striatum, an area activated when people are told to focus on things of value. Cole has found in as-yet-unpublished research that people with more activity in this area show similar patterns of gene expression to those with high levels of eudaemonic well-being. Focusing on something positive and bigger than yourself may activate the ventral striatum, which can inhibit areas like the amygdala, which usually promotes the stress response. Another indication of this comes from research showing that higher scores on a scale of purpose correlated with less amygdala activation.

And one study indicates that people with higher eudaemonic well-being have both increased activity in the ventral striatum and lower levels of the stress hormone cortisol. "Things that you value can override things that you fear," says Cole.

An alternative theory for how purpose could affect biology is by preserving telomeres, caps on the ends of chromosomes that protect DNA from damage, but that shorten with age and stress. A study on stress reduction through meditation has found that it could defend telomeres. But close analysis showed that the benefit was down to a change in sense of purpose, not the meditation directly: the greater a person's purpose became, the more of the protein telomerase they had to protect their telomeres.

Because of findings like these, some researchers think purpose should be more of a public policy priority, shifting away from traditional measures of economy like GDP, and narrowing the focus of happiness campaigns. Doing this would reduce early mortality, give us better overall health and cut the need for medical help, says Michael Steger at Colorado State University in Fort Collins.

It may also help us all get along. Hill has found that people who report higher levels of purpose are less distressed in situations where they are in the minority. It makes sense: people on a mission must accept that achieving their aims requires getting along with others. "Whether goals are focused on helping others or not, it's very rare that our life goals don't involve others at all," he says.

This is all well and good if you're already brimming with direction, but how can people boost their sense of purpose if it is lacking? There are several different strategies. As the study on telomeres indicates, meditation can have an effect. And other research has shown that eudaemonic well-being is strengthened by carrying out random acts of kindness. Cole has found that having a purpose that benefits others may be particularly helpful. But striving for something that isn't necessarily constructive, like climbing a mountain, may be enough to create the health-boosting biology he sees in his studies.

To identify or strengthen your sense of purpose, Steger suggests starting small, by focusing first on making work more meaningful (see "I work therefore I am", page 82) or becoming more invested in relationships. Strecher recommends setting a different purpose for each of four domains in life – family, work, community and personal – and acknowledging that your focus will shift between them over time, and the goals themselves can shift too.

Purpose pills

Strecher says to consider what you would like to be said about you at your memorial, or to identify people you would like to emulate. He is also developing an app called Jool that he hopes can eventually serve as a kind of "purpose pill". Users begin with an assessment, and then get encouragement and guidance as they go on. It is currently being tested by companies to help employees hone their sense of purpose – and boost productivity. His team has followed an initial group of users for over a year, and since then they have begun a series of randomised studies.

There are also more formal therapies that foster purpose and meaning in life for people with conditions such as depression. For example, Dolores Gallagher-Thompson at Stanford University in California, has found that cognitive behavioural therapy can promote meaningfulness. She encourages patients to consider their legacy and how they might provide a good example for children and grandchildren.

Purpose isn't a fixed entity – it waxes and wanes with changes in life. Many people experience a drop in purpose following retirement, for instance, but can regain it by engaging in the community, helping others and remaining sociable. And, as Hill found, the health effects of purpose are apparent whether someone is 20 or 70. "To me, that's evidence suggesting that whenever one finds a purpose it can still imbue benefits," he says. In other words, it's never too late to start seeking the meaning of life. ■


MOD. K 1572 47-49 / 21
4409 / RAINBOW
8392 / CAM.
3455 / GREEN
193 - 194 / VAC
1547 / FEM.
535 - 536
2121
12 / STOP

SEEING REASON

Human brains skew facts.
How can we change our minds, asks Dan Jones

IN 2016, Donald Trump defied the pollsters to be elected the 45th US president. A few months earlier, UK voters decided to end their country's 43-year membership of the European Union. Throughout Europe populist movements are prospering. In every case, opponents have cried foul: these campaigns, they argue, win support by distorting or flagrantly disregarding the truth.

Politicians spin and politicians lie. That has always been the case, and to an extent it is a natural product of a free democratic culture. Even so, we do appear to have entered a new era of "post-truth politics", where the strongest currency is what satirist Stephen Colbert has dubbed "truthiness": claims that feel right, even if they have no basis in fact, and which people want to believe because they fit their pre-existing attitudes.

In recent years, psychologists and political scientists have been revealing the shocking extent to which we're all susceptible to truthiness, and how that leads to polarised views on factual questions from the safety of vaccines to human-caused climate change. The fact is that facts play less of a role in shaping our views than we might hope for in a species whose Latin name means "wise man" – and the problem seems to be getting worse. By figuring out when and why we have a partial view of factual information, however, researchers are starting to see how we can throw off the blinkers.

FRANCESCO BONGIORNI

Let's just establish one fact first: facts are good (see "Inside knowledge", page 6). They may be uncomfortable, or inconvenient, but only by embracing rational, fact-based solutions can we hope to prosper as a society. "We need to have discussions that are based on a common set of accepted facts, and when we don't, it's hard to have a useful democratic debate," says Brendan Nyhan at Dartmouth College in New Hampshire.

In a world of rational empiricists, facts and a careful weighing of the evidence would determine which claims we accept and which we reject. But we are biased. In the real world of flesh-and-blood humans, reasoning often starts with established conclusions and works back to find "facts" that support what we already believe. And if we're presented with facts that contradict our beliefs, we find clever ways to dismiss them. We're more wily defence lawyer than objective scientist.

What's my motivation?

Psychologists call this lawyerly tendency motivated reasoning. Take climate change. The science here is unambiguous: climate change is happening and human activity is driving it (see "Living with climate change", page 108). Yet despite this, and the risks it poses to our descendants, many people still deny it is happening.

The major driver, especially in the US, is political ideology. A Pew Research Center survey released a month before the 2016 US election showed that, compared with Democrats, Republicans are less likely to believe that scientists know that climate change is occurring, that they understand its causes, or that they fully and accurately report their findings. They are also more likely to believe that scientists' research is driven by careerism and political views.

Many liberals like to think this is a product of scientific illiteracy, which if addressed would bring everyone round to the same position. If only. Studies by Dan Kahan at Yale University have shown that, in contrast to liberals, among conservatives it is the most scientifically literate who are least likely to accept climate change. "Polarisation over climate change isn't due to a lack of capacity to understand the issues," says Kahan. "Those who are most proficient at making sense of scientific information are the most polarised."

For Kahan, this apparent paradox comes down to motivated reasoning: the better you are at handling scientific information, the better you'll be at confirming your own bias and writing off inconvenient truths. In the case of climate-change deniers, studies suggest that motivation is often endorsement of free-market ideology, which fuels objections to the government regulation of business that is required to tackle climate change. "If I ask people four questions about the free market, I can predict their attitudes towards climate science with 60 per cent certainty," says Stephan Lewandowsky, a psychologist at the University of Bristol, UK.

But liberal smugness has no place here. Consider gun control. Liberals tend to want tighter gun laws, because, they argue, fewer guns would translate into fewer gun crimes. Conservatives typically respond to that by › saying that with fewer guns in hand, criminals can attack the innocent with impunity.

Despite criminologists' best efforts, the evidence on this issue is mixed. Yet Kahan has found that both liberals and conservatives react to statistical information about the effects of gun control in the same way: they accept what fits in with the broad beliefs of their political group, and discount that which doesn't. And again, it's not about IQ: "The more numerate you are, the more distorted your perception of the data," says Kahan.

We are blinkered on other contentious issues, too, from the death penalty and drug legalisation to fracking and immigration. In fact, the UK's Brexit vote provides another compelling case study in the distorting power of motivated reasoning.

Drawing on responses from more than 11,000 Facebook users, researchers at the Online Privacy Foundation found that while both Remainers and Brexiteers could accurately interpret statistical information

> "The most numerate people are better at distorting the data to fit their beliefs"

when it came to assessing whether a new skin cream caused a rash, their numeracy skills abandoned them when looking at stats that undermined rationales for their views – for example, figures on whether immigration is linked to an increase or decrease in crime.

As a result, the facts they encountered didn't lead them to update their beliefs in line with the evidence – a weakness the Leave campaign exploited. As Arron Banks, co- founder of the Leave.eu group said in a 2016 interview: "The Remain campaign featured fact, fact, fact, fact, fact. It just doesn't work. You've got to connect with people emotionally. It's the Trump success."

Lewandowsky points to another problem: the lure of conspiracy theories. When it comes to climate change, "you can say 'All the scientists have made a mistake', which is a hard sell, but it's much easier to say 'They're all corrupt'," says Lewandowsky. His work shows that many people do in fact reject climate change as a conspiracy, and they tend to endorse a wide range of other conspiracy theories (see "It's a cover-up!", right).

Political ideology doesn't explain everything. The bogus link between autism and the vaccine for measles, mumps and rubella, while often portrayed as a liberal obsession, cuts across politics. "Opposition to vaccines is a diverse phenomenon, and resists easy generalisations," says Nyhan. "There's no demographic factor that predicts who is most vulnerable to anti-vaccine claims."

It's clear, then, that many of us, if not all, are stuck with blinkers. But how did we get to a point where facts have almost no value? It could be down to how we get our news. In the immediate aftermath of Trump's election, Facebook CEO Mark Zuckerberg came in for criticism for effectively running a media machine – perhaps the world's biggest – without the due care that should come with such a responsibility. In the US, nearly two-thirds of people get news through Facebook, which is programmed to bring you news similar to what you've already seen – often what the most ideological and politically active people in your feed have shared.

It's not hard to see how that could have an amplifying effect on motivated reasoning, and the rise of social media might well explain why our problems with facts seem to have grown more acute. These days, it's easy to drift into echo chambers reverberating not only with news and views that confirm your biases, but also falsehoods, rumours and conspiracy theories jostling with stories from reputable sources. So if we want to restore the power of facts, perhaps it is time to rethink how news is delivered on the largest scales.

But even if the social media "filter bubble" is burst and everyone is exposed to inconvenient

We can't see past our biases on immigration and vaccination risks

ELLIOTT FRANKS/EYEVINE; BOTTOM: SUSANNAH IRELAND/THE TIMES

It's a cover-up!

Why we're drawn to conspiracy theories

NASA

Estimates suggest that half of the US population believe in a conspiracy theory

Were the moon landings faked? Was the US government behind the 9/11 attacks? Is human-caused climate change a liberal hoax? The power of conspiracy theories has never waned – in fact, according to a recent estimation, at least half of the US believes in one or more of the common ones. And to some extent, we're all susceptible because conspiratorial thinking stems from universal aspects of human psychology.

There is our propensity to see threats lurking everywhere and to make links between coincidental events. But according to Joanne Miller, a political scientist at the University of Minnesota in Minneapolis, belief in conspiracy theories is also fuelled by politically motivated reasoning – a tendency to skew factual information according to our pre-existing beliefs and political allegiances (see main story). "Both conservatives and liberals are prone to accept conspiracy theories that make the other side look bad," says Miller. But she has also found that conservatives, especially those who are knowledgeable about politics but distrust mainstream authorities, are most likely to endorse conspiracy theories.

This reason, suggests Miller, is that conspiracy theories are most attractive to those who feel they're on the losing end of politics. Indeed, Miller has found that inducing this feeling of losing out increases endorsement of conspiracy theories across the political spectrum, though again the effect is more pronounced among conservative Republicans – which means Donald Trump's claims that the election was rigged made perfect sense as a campaign tactic.

Now that Trump is president, we might see a reversal, says Miller. "Liberals and Democrats might become more likely to believe conspiracy theories that make the other side look bad now that they find themselves the political losers."

truths, it may not be enough. A study of 1700 parents done by Nyhan and Jason Reifler at the University of Exeter, UK, reveals that fact-based messages of the sort often used in public health campaigns don't work – and sometimes have the opposite effect to what was intended. So while messages debunking the claim that the MMR vaccine causes autism, for example, did reduce belief in this misconception, they actually decreased intent to vaccinate among parents with unfavourable attitudes towards vaccines. Similarly, images of children suffering from the diseases that MMR prevents led sceptical parents to be less likely to vaccinate than they were previously. Nyhan and Reifler call this the "backfire effect".

That is not to say that debunking myths, which became an Olympic sport during the 2016 presidential election campaign, is a waste of time. Nyhan and Reifler found that during the 2014 midterm elections in the US, fact-checking improved the accuracy of people's beliefs, even if it went against ingrained biases. Democrats would update their beliefs after having a claim made by a Democrat debunked, and Republicans did likewise.

Work by Emily Thorson at Syracuse University in New York paints a similar picture. She found that misconceptions on issues like how much of the US debt China owns, whether there's a federal time limit for receiving welfare benefits and who pays for Social Security could be fixed by a single corrective statement.

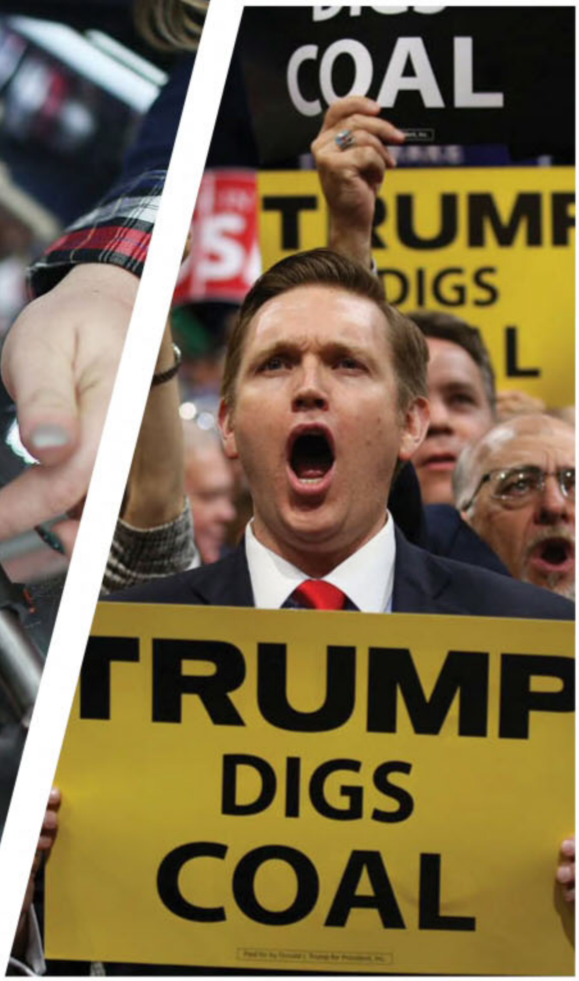
The bad news is that myth-busting loses its power on more controversial or salient issues. "It's most effective for topics that we're least concerned about as a democracy," says Nyhan. "Even the release of President Obama's birth certificate had only a limited effect on people's belief that he wasn't born in this country." And Thorson has found that even when corrections work – say, getting people to accept that a fictional congressman accused of taking campaign money from criminals did no such thing – the taint of the earlier claim often sticks to the innocent target, in what she calls "belief echoes".

Yet Thorson remains upbeat. "It's easy to become pessimistic when we focus on really frustrating cases like 9/11 conspiracy theories or Obama's birthplace," she says, "but there's still a lot of room to use facts to change attitudes."

Changing minds

In some cases, the power of facts to persuade might turn on the way they're presented. For example, Nyhan and Reifler have found that information presented graphically leads people to form more accurate beliefs about the topic in question – the effectiveness of Bush's troop surge in Iraq in 2006/2007, say, or the state of the economy under Obama – than simply reading text about the same topic. And this is true even when the people looking over the graphs have political reasons to reject the conclusions they encourage. For Nyhan, it is a simple way of re-packaging information that journalists and the broader media could take into account when reporting stories.

Another avenue draws on the idea that people reject facts because they threaten the identity built around their world view. If so, buffering self-esteem might reduce that threat. When Nyhan and Reifler got people to reflect on and write about values that are important to them, an esteem-enhancing intervention called self-affirmation, they found that it can do the trick – but its effects are not uniform. For instance, for Republicans whose identity is not strongly tied up with their party, self-affirmation makes them less likely to reject claims about climate change, but among Republicans who strongly identify with the party, the intervention either has no effects, or reinforces their beliefs. ›



LEFT: JOE RAEDLE/GETTY; RIGHT: DAMON WINTER/NYT/EYEVINE

From gun control to climate change, our existing beliefs skew how we see the facts

Likewise, Miller has found that self-affirmation increases endorsement of conspiracy theories among conservatives, but not among liberals. Combining graphical information with self-affirmation also produces mixed results, depending on who you're dealing with.

Until recently, researchers had found no personality trait that mitigates motivated reasoning. But in 2016, Kahan discovered something intriguing about people who seek out and consume scientific information for personal pleasure, a trait he calls scientific curiosity. Having devised a scale for measuring this trait, he and his colleagues found that, unlike scientific literacy, scientific curiosity is linked to greater acceptance of human-caused climate change, regardless of political orientation. On a host of issues, from attitudes to porn and the legalisation of marijuana, to immigration and fracking, scientific curiosity makes both liberals and conservatives converge on views closer to what the facts say.

Perhaps even more encouragingly, Kahan's team found that scientifically curious people were also more eager to read views that clashed with those of their political tribe. So finding ways to increase scientific curiosity, perhaps by increasing the influence of people with this trait, could take the heat out of partisan disputes more effectively than promoting scientific literacy.

Kahan sees other glimmers of hope. One might be to exploit what he calls "cognitive dualism", the ability to hold two seemingly contradictory beliefs at the same time.

> "Asking about climate change is akin to asking 'whose side are you on?'"

It's a phenomenon at play in the recent Pew survey on climate change: just 15 per cent of conservative Republicans agreed that human activity was causing climate change, but 27 per cent agreed that if we changed our ways to limit carbon emissions it would make a big difference in tackling climate change.

The same cognitive dualism is evident among US farmers. A 2013 survey of farmers in Mississippi, North Carolina, Texas and Wisconsin found that only a minority accepted climate change as a fact (see "Living with climate change", page 108). Yet a majority in each state believed that some farmers will be driven out of business by climate change, and the rest will have to change current practices and buy more insurance against climate-induced crop failures. By buying crops genetically engineered to cope with climate change and purchasing specialist insurance policies, many of them already have.

The psychological underpinnings of this "quantum mental state", in Kahan's words, are mysterious, he says, but it's important because it suggests that people can think about factual issues at very different levels, depending on the extent to which the issue is bound up with their identity. Kahan thinks that asking people about human-caused climate change is akin to asking "Who are you, and whose side are you on?", which is why political identity makes such a difference to their answers. But when you start talking about climate change as a local, personal issue, it loses its political edge and becomes a more pragmatic concern.

"When issues are wrapped up in national electoral politics, they have a resonance that divides people," says Kahan. "So you want to depoliticise things along one dimension to facilitate action at another level."

Taking poisonous partisan politics out of factual issues like climate change is part of what Kahan calls "detoxifying the science-communication environment". A major pollutant of this ecosystem, argues Lewandowsky, is the influence of dark money in politics. A 2013 study by Robert Brulle at Drexel University, Philadelphia, found that between 2003 and 2010, $558 million was funnelled through third-party "pass through" organisations, which hide the source of money, to climate-denial groups. "We have to talk about these anti-democratic influences and how they affect public discourse," says Lewandowsky.

So is there any hope for facts? Restoring their power is not going to be easy. But despite the challenges, Nyhan cautions against despondency. "It's important not to overstate what's different about today from the past, when there were other ways of circulating misinformation," he says. Although slower than today's instant-access 24-hour news and all-consuming social media, they still allowed politicians to introduce false claims into the national debate.

"There was no Golden Age of democracy when facts dominated public opinion or political discourse," says Nyhan. "But we've survived nonetheless". ■

the LIFE EDITOR

Biology is undergoing a revolution. **Michael le Page** reports on the gene-editing technique that is changing everything

THE food on your plate. The pets at your feet. The plants in your garden. The mosquitoes whining in your ear at night. The cells in your body. And perhaps even the brains and bodies of your children. All of these could be transformed by a new gene-editing technique – starting in your lifetime.

Terms like breakthrough and revolutionary are much abused. But when it comes to CRISPR gene editing, they are probably understatements. "The technology is unbelievable," says Kamel Khalili of Temple University in Philadelphia, who thinks it could clear viruses like HIV from the body.

Major impact

The pace of innovation is breathtaking. Just a few years after its invention, CRISPR gene editing is already having a major impact on biomedical research. It makes it easy to "turn off" genes one at a time, to see what they do. It can introduce specific mutations, to find out why they make cells cancerous or predispose people to diseases. And it can be used to tinker with the genes of plants and animals, to create drought-resistant maize, more muscular police dogs and much more.

In the not too distant future, CRISPR-based research could bring drugs for tackling obesity, more powerful gene therapies and plentiful supplies of transplant organs. "CRISPR is evolving incredibly fast," says Waseem Qasim of University College London, whose team recently used an older form of gene editing to save the life of a baby with leukaemia. "We can't keep up."

Then there is the most controversial application: it could be used to permanently alter the genomes of our descendants, in order to eradicate disease-causing mutations or even to enhance children by adding beneficial gene variants that both their parents lack.

We have been talking about the possibility of genetically engineering humans for decades, says Debra Mathews of the Johns Hopkins Berman Institute of Bioethics in Baltimore. "But we've never had a technology that had a reasonable chance of doing what we want to do without causing harm before."

This so-called germline gene editing first hit the headlines in 2015, when the results of the first attempts to modify human embryos with CRISPR were published by a team in China. Rumours about such attempts had already led to calls for a voluntary ban on editing genes in human embryos. That was one of the issues on the agenda at an international meeting on gene editing hosted by the US National Academy of Sciences in 2015.

Those in favour of such research say there might be good reasons for allowing germline gene editing, and that it is a powerful tool for understanding human embryonic development, which may reveal why some people are infertile or miscarry, for example.

In the following pages we look at the potential of CRISPR gene editing to transform medicine – and also its dangers.

GENE EDITING DECODED

GERMLINE EDITING Altering the genes of sperm and egg cells, or early embryos, so that changes are passed to subsequent generations
SOMATIC CELLS All cells in the body except sperm and eggs. Changes to these cells are not passed on to offspring
CRISPR Gene-editing technique derived from a mechanism that bacteria use to fight off viruses. Cheaper, faster and more precise than earlier methods
GENE DRIVE Gene-editing technique that allows traits to spread faster through a population than they would normally
PGD (preimplantation genetic diagnosis) Screening embryos fertilised through IVF for genetic diseases before they are implanted in a woman's uterus

PAUL RYDING

Will this lead to
DESIGNER BABIES?

Genetically modified superhumans. Babies born with made-to-order characteristics. The idea has been explored in everything from academic journals to movies. CRISPR technology could make it a reality.

The ability to alter human genes in a way that can be passed onto offspring, called germline engineering, has long been possible. But until recently the methods available to genetically modify animals were so inefficient and crude that no sane biologist would dream of using them on humans. Tinkering with the genes inside people has been limited to gene therapy, where the changes don't get passed to the next generation (see "Will gene therapy go mainstream", overleaf).

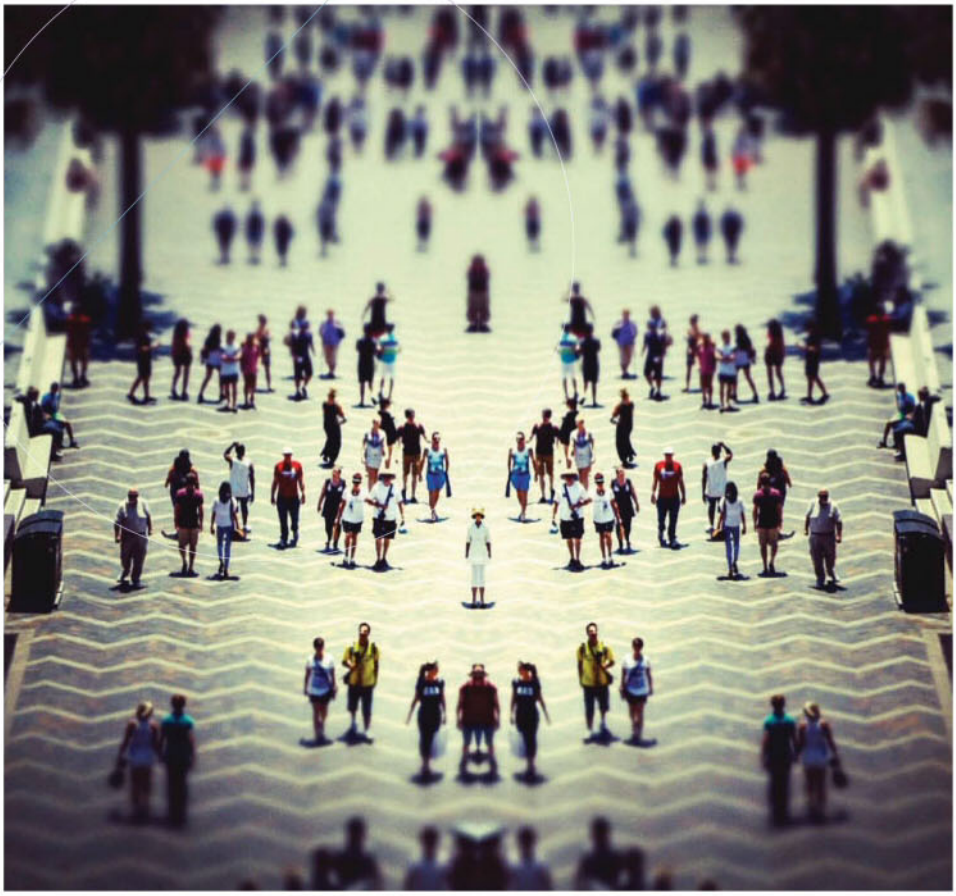
Now the precision and efficiency of CRISPR has reopened the debate about human germline engineering. But why do it? The most compelling reason would be to prevent the inheritance of genetic diseases, yet this is already being done without gene editing.

Dialling out disease

One approach is prenatal testing, which involves screening for the disease-causing mutation during pregnancy, giving parents the option of abortion. Another is preimplantation genetic diagnosis (PGD), in which prospective parents who undergo IVF have their embryos screened. Only those that won't develop the disease in question are then implanted. PGD can already be used to prevent thousands of serious genetic diseases.

But PGD is impractical if a child is at risk of inheriting two or more disorders. For two disorders, three-quarters of embryos might be unsuitable – and with couples getting only a handful of embryos per IVF cycle, rejecting three-quarters of them would make conceiving far less likely.

In fact, even with single disorders some couples get so few embryos ›



STEPHEN ISIS/EYEEM/GETTY

that none of them would be free of the disease-causing mutation. In these instances gene editing could be used to fix their DNA – an option some regard as ethically preferable to discarding embryos, says Robin Lovell-Badge of the Francis Crick Institute in London.

Similarly, while PGD is impractical for getting rid of the harmful gene variants that increase the risk of common conditions, such as heart disease, diabetes, Alzheimer's or schizophrenia, it might be feasible with gene editing. Eliminating dozens of these harmful variants could make a huge difference, allowing people to live longer, healthier and even happier lives.

"Even if gene editing were safe, we are a long way from being able to engineer intelligence"

This has huge potential for reducing the disease burden, says Chris Gyngell, who studies the ethics of human enhancement at the University of Oxford.

But the technology isn't there just yet. Nor do we know enough about the gene variants involved to start eradicating them. All this means that when it comes to reducing the risk of diseases, there is currently no compelling reason to attempt germline gene editing.

What about the far more controversial idea of enhancing children, by giving them gene variants both their parents lack? Many of the variants controlling skin, hair and eye colour have been identified, so in theory these kinds of cosmetic traits could be tweaked. But characteristics such as intelligence seem to be determined by hundreds of different gene variants, with each one having only a tiny effect. This means we are a long way from engineering intelligence into children, even if gene editing were safe enough to attempt it.

And that isn't yet clear. The most serious issue is that gene-edited embryos are often a mix of cells with the desired genetic change and cells without it – a phenomenon called mosaicism. The Chinese team observed this in their first attempt to edit human embryos.

Everyone in the field agrees it is far too soon to attempt to alter children. A few have gone much further and

HOW DOES IT WORK?

The first forms of genetic engineering involved adding extra bits of DNA to the genomes of plants and animals, with no control over where they ended up. One method involved shooting bullets coated with DNA at cells.

Gene editing, in contrast, adds DNA to precise spots in a genome, or alters a specific sequence, so is far superior. While a few methods of gene editing had been developed, until the advent of CRISPR, it was usually slow, difficult and very expensive.

CRISPR targets a particular DNA sequence using a piece of RNA that's complementary to that DNA. Linked to it is a protein derived from bacteria, called Cas9. The RNA finds the right bit of DNA and binds to it, then Cas9 cuts it (see diagram, right). The cell's repair mechanisms will re-join the two pieces, but in the process the DNA sequence gets slightly altered. This is how genes can be disabled.

Donor DNA

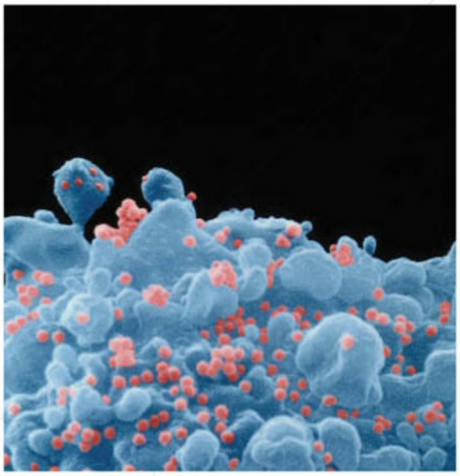
If, however, donor DNA with ends that match the DNA on either side of the cut segment is added to cells too, the cell thinks it is a fragment of broken DNA and will splice it into the genome exactly where the cut was made – adding DNA to a precise spot.

The Cas9 protein can also be modified so that instead of cutting DNA, it controls the activity of the local gene or genes – boosting or blocking their activity.

The effects of this will be short-lived: things return to normal once the Cas9 protein breaks down.

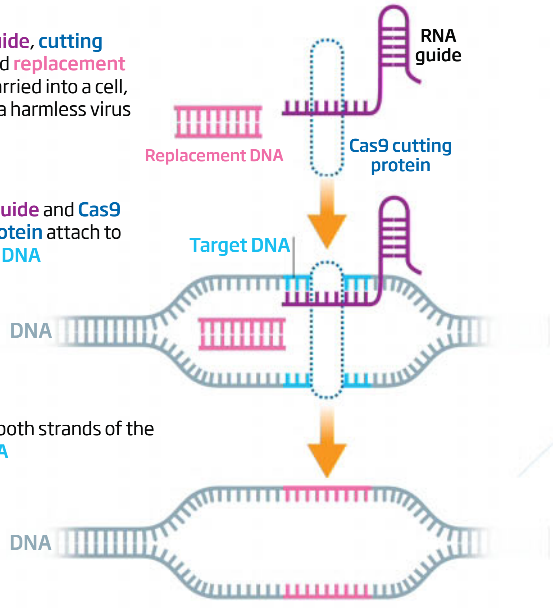
But it may be possible to produce longer-term changes in gene expression through epigenome editing, which alters gene switches, rather than genes themselves. This could one day help treat the wide range of disorders thought to involve epigenetic changes, including addiction and depression.

It could be possible to target HIV hiding inside immune cells



Cut and paste

CRISPR allows one or more genes to be edited far more precisely than ever before



The cell's own DNA repair mechanism splices the replacement DNA into position. With the right RNA guides and DNA replacements multiple gene changes can be made in a single step

called for a voluntary worldwide ban on any work involving gene-editing of human embryos, even if there is no intention of allowing them to develop. "At this stage the question is whether to go ahead with research," says Gyngell.

A ban appears unlikely. Several UK research organisations and funding bodies have declared their support for research in human embryos. So, too, has an international group of stem cell researchers and bioethicists, called the Hinxton Group, after holding one of the first meetings on it in 2015.

It concluded that not only should we leave the door open to germline gene editing, but there is much to be gained from research, including understanding embryonic development and finding out why some women miscarry. "We are all in agreement that this research has tremendous value," says Debra Mathews of the Johns Hopkins Berman Institute of Bioethics, who was at the meeting.

But while germline editing is still decades away, it might not be that long before we start using CRISPR to tweak the genes of adults...

Will gene therapy go MAINSTREAM?

Gene therapy – using genes to treat or prevent diseases – is already saving lives. But it's still very much an experimental treatment used on a few individuals, rather than a routine therapy. CRISPR will help change that.

For starters, it will be much cheaper and easier to develop treatments to the point where they are ready to test in animals and people. And these treatments should be safer. "It's a new age in gene therapy," says Oskar Ortiz of the German Research Centre for Environmental Health in Munich.

Unlike germline engineering (see "Will this lead to designer babies?" page 39), changes made using gene therapy can't be passed on to children. The conventional technique involves adding extra DNA to cells, but there is no way to control where it lands – and if it lands in the wrong place it occasionally results in cancer. With CRISPR gene editing, DNA can be added to a precise spot. Even then it can sometimes be added to the wrong place, but ways to minimise this have already been developed.

What's more, treating some diseases requires altering existing genes rather than adding new ones – and gene editing excels at this. If it proves safe, CRISPR could be used to modify cells in the body to treat a wide range of diseases.

The tricky part will be delivering the gene-editing machinery and new DNA to cells inside the body. Fortunately, biologists working on CRISPR therapies can take advantage of the decades of work spent creating tools for delivering conventional gene therapy. The most popular method is to use harmless viruses called AAVs to carry the new genes into cells.

Viruses carrying CRISPR components have already been used to target genes inside the brains of mice, to find out what those genes do. The drawback is that AAVs can only carry 4700 bases of DNA – and the gene for the key CRISPR protein, called Cas9, is nearly this big (see "How does it work?", page 40). That's OK if the aim is just to disable an existing gene. But it won't work for adding genes; there isn't enough room in AAVs to carry both Cas9 and a gene.

Unexpected problems

There are already some ways around this size limit. Smaller alternatives to the standard Cas9 protein are being tested, for instance. It is also possible to split Cas9 in two, and deliver each half in a separate virus, leaving more room for the rest of the payload. This approach is likely to be less efficient, though, as at least two AAVs have to deliver their DNA to each cell.

Years of animal tests will be needed to ensure that CRISPR-based gene therapy is safe enough to try in people, and Bryan Cullen of Duke University in Durham, North Carolina, cautions that unexpected problems will almost certainly arise when human trials begin. Nevertheless, he is confident that the approach will work: "It will lead to treatments within a decade," he says.

Could it help
CURE DISEASES
like AIDS and herpes?

Every human has viruses lying low within them. It's highly likely that you have been infected with a human papillomavirus at some point, for example. HPV inserts copies of its DNA into the genome of cells, allowing it to hide away for decades, ready to activate and infect more cells if your immune system slips up. The consequences can be lethal: HPV can cause neck, throat, anal and cervical cancer.

The herpes simplex virus does a similar thing, adding one or more copies of its DNA to sensory nerve cells around the mouth or genitals. And several other common viruses also exploit this dastardly trick to hide away within you.

Until recently, the best we could do was help the immune system suppress any viral activity. With the development of the first gene-editing tools, researchers began to explore the possibility of destroying the viral DNA inside our cells, but progress has been slow and results mixed.

Huge demand

Now, with CRISPR, the field is racing ahead. Several groups have shown that it is possible to target and destroy viral genes in human cells growing outside the body.

"We are all brimming with excitement," says Bryan Cullen of Duke University, whose team is working on treatments for several viruses, including herpes. "There will be a huge demand if a cure is possible."

It should be easiest to target viruses that integrate themselves into specific tissues in a small area of the body, such as herpes simplex. Could it also work for HIV? The problem, Cullen says, is that HIV hides in memory T-cells, which are scattered throughout the body.

But it might not be necessary to eliminate HIV in all cells in one go, says Kamel Khalili of Temple University, whose team has had success in animal tests. Khalili hopes that a series of treatments, in combination with existing antiretrovirals, will allow the immune system to recover to the point that people with HIV can eventually be cured. His team's aim is to make the treatment simple and cheap enough that it can be used in the resource-poor countries that have the most HIV-positive people.

PETER STEFFEN/AP/PA

PROFILE

Emmanuelle Charpentier is a molecular biologist at the Max Planck Institute of Infection Biology in Berlin, and the Helmholtz Centre for Infection Research, Germany

INTERVIEW

"IT'S AN ETHICAL RESPONSIBILITY"

Gene editing could transform life, but we need to discuss the ethics of how it is used, says **Emmanuelle Charpentier**, co-discoverer of the Cas9 CRISPR technique

Now that CRISPR has arrived, should there be a moratorium on editing human embryos?

There have been technologies before that could target specific gene sequences. So the idea of having a technology that would allow biologists to manipulate the human germline isn't new. But the previous technologies were less efficient. CRISPR is very powerful and so easy to use that it is being harnessed for different purposes. It's like any technology: there's a good side to it and there is an ethical responsibility with regard to how to use it.

How do we decide how to use this technology?

By discussions. There is perhaps a misunderstanding by the public about what the technology does and how it works. It allows more precise genetic changes than all the breeding technologies that have been used before – so the organisms are genetically much cleaner. I think all those around the table discussing the ethics – which includes scientists, clinicians, ethicists and the public – should first understand the technology and that it allows us to accelerate the understanding of the functions of genes. That will be important for the development of biotechnology and biomedicines. Then after that manipulation of the human germline needs to be discussed.

Is there ever a case for germline editing – altering DNA that can be inherited?

There would be reasons with regard to certain types of diseases. I hope that using the technology with the idea of changing human characteristics will not be pursued. When it comes to using it for therapeutic and preventative purposes – not to change traits that could be inherited throughout the population – then the debate will be for certain kinds of diseases for which maybe the manipulation of the human germline will be considered. But then the question is whether society wants to go there. Philosophically and sociologically speaking, I have lots of issues with this.

Did you expect your work to cause this debate?

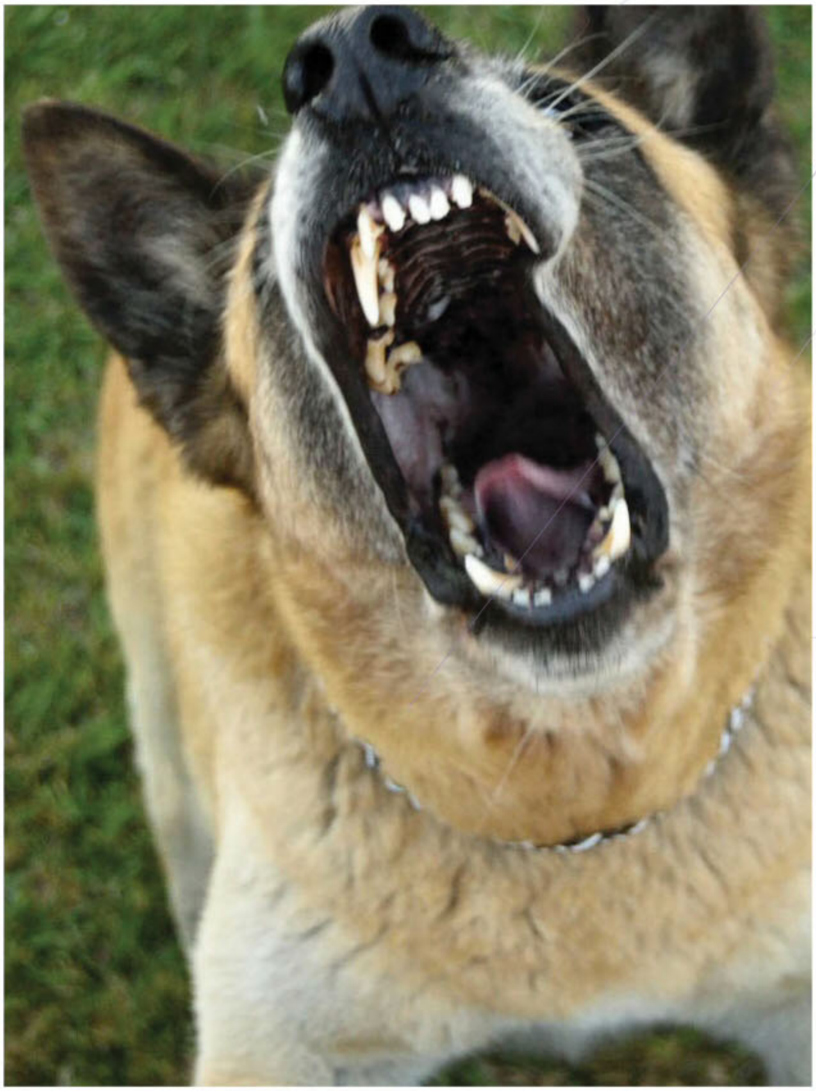
It was unexpected. The findings came a little bit all of a sudden. But it comes down to basic science – CRISPR is a very nice example of how basic science on an obscure immune system in bacteria can potentially lead to powerful technology.

What do you hope will have come out of your research in 10 years?

I hope that technology will be developed to the point at which CRISPR Cas9 can be delivered in cells and tissue to treat severe human genetic disorders.

Interview by Catherine de Lange

This is an edited version of an interview given at the Falling Walls conference in Berlin in November 2015. To read the full version visit bit.ly/NS_GeneEdit



LUIS DÁAZ DEVESA/GETTY

Should we be
WORRIED?

After Dolly the sheep was born, several groups announced they were going to clone people. A bizarre religious cult and a maverick fertility doctor even claimed success in the 2000s, but these claims have never been taken seriously. As far as we know, no clone of an adult human has yet been born, not least because we have struggled to create cloned human embryos.

CRISPR gene editing, by contrast, is relatively easy. It's not the sort of thing anybody could do in their kitchen, but with sufficient money a small team of rogue biologists and IVF doctors could create the first gene edited baby right now. "This is the thing that scares me the most," says Robin Lovell-Badge of the Francis Crick Institute.

In fact, there is nothing to stop IVF clinics trying germline gene editing in many countries, including the US. "You can easily imagine clinics trying to boost their revenue by offering this," says Lovell-Badge, who points out that unregulated clinics offering unproven stem-cell treatments are springing up all over the world.

Such irresponsible behaviour might be disastrous for the health of children – and the purses of their parents – but for now it poses no wider issues. We don't know how to create superhumans even if we wanted to (see "Will this lead to designer babies?", page 39).

The biggest impact from CRISPR will come from the enormous range of genetically altered plants, animals, fungi and bacteria it will be used to create. The technique has already been used to create extra-muscular dogs for police work, hornless cattle for farmers and micropigs for pets.

So far fears about genetically engineered plants and animals – that they will harm our health or the environment, for instance – have proved largely unfounded, but with CRISPR making it much easier to tinker with genes the odds of things going wrong will be greater. It's possible, for instance, that plants given traits such as drought resistance, salt tolerance or faster growth will start spreading and become invasive weeds. Then again, other human activities such as introducing exotic species have already created many invasive weeds and pests.

Another risk comes from something called "gene drives", which CRISPR is making both easier to create and more powerful. Normally a genetic variant in an organism has a 50 per cent chance of being inherited by offspring. But a gene drive can insert a copy of itself to the DNA inherited from the other parent. That guarantees it gets passed to all of the organism's offspring, meaning it can spread very rapidly through a population. In theory gene drives could be deliberately unleashed to wipe out unwanted species such as disease-carrying mosquitoes. But there are fears they could also spread uncontrollably in the wild as a result of lab accidents.

> "There are fears that gene drives could spread in the wild as a result of lab accidents"

"We need to be careful," says Austin Burt of Imperial College London, who works on gene drives. But the risk needs to be kept in perspective: we are already causing a sixth mass extinction because our activities are wiping out so many species. Gene drives would affect only one species at a time, and in species that reproduce slowly – like us – they would spread extremely slowly.

The worst-case scenario is that CRISPR is accidentally or deliberately used to engineer a pathogen that infects people or crops – a biological weapon, in other words. But it is already possible to do this in other ways.

The power of CRISPR means it could have huge benefits, allowing us to produce more and healthier food even as the climate changes, and to improve the health and welfare of ourselves, our pets and farm animals. But much depends on this power being used wisely. Or as the uncle of a fictional transgenic creation says: with great power comes great responsibility. ■

The coming plague

A killer pandemic is now more likely than ever. Where will it come from and how can we beat it, asks **Debora MacKenzie**

IN 1347, an epidemic of unimaginable ferocity struck Europe. People first experienced flu-like symptoms, but within days painful swellings developed, which turned black, split open and oozed pus and blood. The Great Pestilence, later dubbed the Black Death, swept across the continent within four years, killing up to half the population. The disease persisted in Europe until the 1700s, always circulating somewhere, killing people off.

We speak of it nowadays as history. In fact, it is more like natural history: infectious disease is part of the ecology of our species. Until 1900, and despite considerable competition from violence and starvation, it was our biggest killer, causing half of all human deaths. Now, it accounts for fewer than a quarter of all deaths worldwide, most of them in poor, tropical regions. In rich countries it is only a few per cent. And the toll is falling.

But we shouldn't be complacent: plagues will return. The 1960s notion that infectious disease was on the way out ended when HIV appeared in the 1980s. Since then, many infections like bird flu, SARS and Zika have caused alarm. But it took a near-disaster – the worst ever outbreak of Ebola – to scare the inertia out of governments. As a result, we are at last preparing for the inevitable. A clutch of new programmes will improve our grip on microbial killers. And the world now has an emergency medical response team – which, astonishingly, it never had before. But we aren't there yet. If a novel virus struck now, we would still be in trouble.

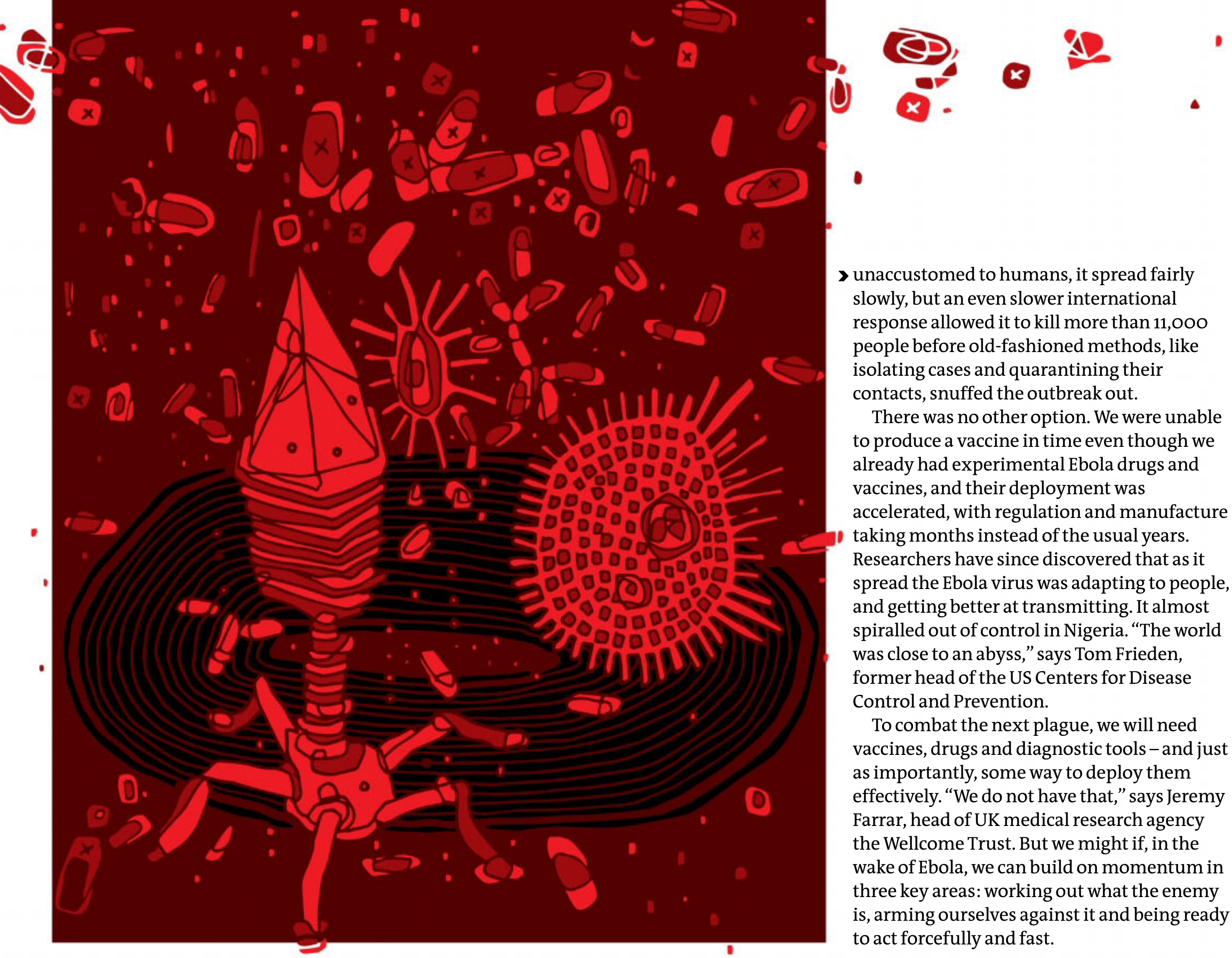
BRIAN LAROSSA

For all our high-tech modernity, and in many ways, because of it, the risk that new infectious diseases will evolve is actually getting worse. Pathogens began circulating regularly among humans only after we started farming and settled in towns. One reason was that we caught infections from our livestock: flu from ducks, tuberculosis from cows. But crucially, there were enough of us in close proximity that a germ could always find a new host and keep spreading, persisting among people and adapting to us.

Now we are crowding into cities and travelling more, especially within the tropics where pathogen diversity is highest. That plus globalised trade, migration and climate change leads to a reshuffle of wildlife, people and pathogens. Farms and towns invade the habitats of animals with viruses that can jump to us, or to our densely packed livestock, also booming as demand for animal protein soars.

Public health experts have been warning for years of "emerging" diseases, which can go from unknown to epidemic if the pathogen mutates or the ecology of its hosts changes to make its spread easier. And it is viruses that epidemiologists are most worried about. Bacteria can be deadly, and antibiotic resistance could mean diseases from gonorrhoea to ordinary bladder infections become incurable, but work has at least begun on new drugs. In contrast, viruses can evolve and spread faster, there are thousands we know nothing about, and we have few drugs against them. The worst emerging infections since 2000 have all been viruses.

None is more alarming than the 2014 outbreak of Ebola in West Africa. The virus infected 50 times more people than any previous outbreak, and reached big cities for the first time. As a bat virus still ›



unaccustomed to humans, it spread fairly slowly, but an even slower international response allowed it to kill more than 11,000 people before old-fashioned methods, like isolating cases and quarantining their contacts, snuffed the outbreak out.

There was no other option. We were unable to produce a vaccine in time even though we already had experimental Ebola drugs and vaccines, and their deployment was accelerated, with regulation and manufacture taking months instead of the usual years. Researchers have since discovered that as it spread the Ebola virus was adapting to people, and getting better at transmitting. It almost spiralled out of control in Nigeria. "The world was close to an abyss," says Tom Frieden, former head of the US Centers for Disease Control and Prevention.

To combat the next plague, we will need vaccines, drugs and diagnostic tools – and just as importantly, some way to deploy them effectively. "We do not have that," says Jeremy Farrar, head of UK medical research agency the Wellcome Trust. But we might if, in the wake of Ebola, we can build on momentum in three key areas: working out what the enemy is, arming ourselves against it and being ready to act forcefully and fast.

1 KNOW YOUR ENEMY

First, what should we prepare for? "Spotting the next HIV or SARS before it strikes is virtually impossible," says Ab Osterhaus, head of the Research Center for Emerging Infections and Zoonoses in Hannover, Germany. "There are too many viruses in too many species, interacting with humans and evolving in unpredictable ways." To narrow the field, he says, we need "a detailed understanding of when, where and how viruses are moving from wildlife to people". Because, like the historical plagues, the next big disease is likely to be one that has made the leap from other animals to us.

In 2016, Mark Woolhouse and his colleagues at the University of Edinburgh, UK, reviewed what we know about such viruses. They identified 37 already able to spread from human to human, though poorly, that could become more contagious. These range from virtual unknowns like o'nyong-nyong, an African virus that causes debilitating joint pain, to Rift Valley fever, a common livestock illness.

That's just the viruses we know. A project called PREDICT, funded by the US Agency for International Development, is looking for others. In places, mostly tropical, where humans and wild mammals interact, the project screens people, their food and their rodent, bat and primate neighbours, looking for genetic sequences of viruses in families known to spawn human pathogens. As of February 2017, they had found 984 viruses, 815 of them new to science. In the process, they have mapped hotspots of viral diversity and trained and equipped local labs to test for viruses and watch for disease.

Predicting risk

But which of these viruses should we focus on? Some are obvious, such as a Chinese virus closely related to SARS but different enough that prototype SARS vaccines won't work against it. Others might be identified using a clue discovered by Kevin Olival of the EcoHealth Alliance, who works with PREDICT.

He has statistically analysed all the available data on the flavivirus family, a troublesome lot carried by mosquitoes and ticks that includes yellow fever, Zika, dengue and West Nile. In 2016, he reported that the more species a flavivirus regularly infects, the more likely it is to infect humans as well. That makes the riskiest flaviviruses a clutch of virtual unknowns: Usutu – a bird-borne virus invading Europe – Ilheus, louping ill, Wesselsbron and Tyuleniy.

The Global Virome Project wants to go further in learning about the enemy, genetically sequencing and mapping most of the estimated half-million so-far undiscovered viruses in families we know can infect humans. It reckons it will need $3.4 billion to do that over the next decade, and this year it will start canvassing for funds. The hope is that knowing what viral diversity exists and where could provide unexpected insights and spur investment in disease control.

2 ARM YOURSELF

Once we know what we are fighting, we have to arm ourselves. Finding weapons won't be easy, though. The vaccines that defeated so many infectious diseases in the 20th century were mostly made by government-owned firms that didn't have to turn a profit and produced what was needed as a "public good". In the 1980s, everything was privatised. That was good for spurring profitable medicines for chronic conditions. But much medical innovation is now done by small, start-up biotech firms, which can't afford to shepherd their products through the "valley of death" – the long, expensive process of testing for safety and efficacy, and establishing manufacturing processes and formulations for licensing. Only big pharma companies have the know-how and the $1 billion or so needed to bring a new vaccine to market. But vaccines for common diseases offer little profit; those against a virus that might or might not go epidemic are a commercial non-starter.

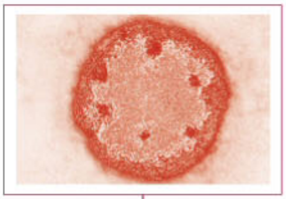
There have been efforts to bring public good back in. Since the 1990s, new treatments for diseases of poverty, like the meningococcal vaccine for Africa, were developed by public-private partnerships between big pharma, governments and large philanthropies like the Bill & Melinda Gates Foundation. "But the momentum fell," says Farrar. In 2013, government and private research on "neglected" emerging diseases amounted to only 1.6 per cent of the $195 billion spent on health R&D. Of that, only a fifth was private.

Ebola provided fresh momentum. In May 2016, the World Health Organization set out an "R&D blueprint for action to prevent epidemics", which aims to develop responses before the next plague strikes. Committees were set up to look for solutions to problems that emerged during the Ebola outbreak, from agreed protocols for quickly testing and licensing experimental drugs and vaccines, to liability insurance for using experimental products, to contracts ensuring information and biological samples are shared.

But the most important goal is to accelerate R&D on nine priority pathogens (see "The nine viruses of the apocalypse", right). Using an approach pioneered for malaria vaccines, the WHO is finding out what research is being done, get participants talking and push progress towards vaccines, drugs and diagnostic tests. Any products must be

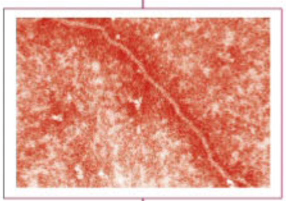
THE NINE VIRUSES OF THE APOCALYPSE

These are the diseases the World Health Organization thinks we should find remedies for, fast. The first six are its highest priority.



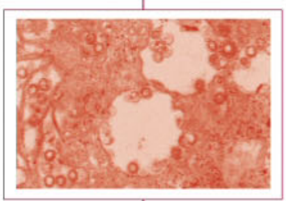
Lassa fever

This West African virus, carried by the common Natal multimammate rat, infects 300,000 people a year. Most have no symptoms, but it can cause diarrhoea and vomiting, then internal fluid accumulation, bleeding from orifices, shock, seizure and coma. It kills some 5000 people annually. Initial symptoms resemble other local diseases, making diagnosis tricky – one reason West Africa was slow to spot Ebola.



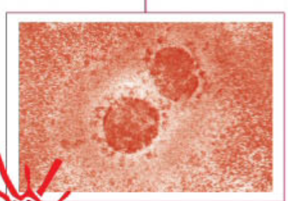
Nipah

This bat virus started killing people in 1999 in Malaysia after pig farms were built near fruit bats, which dropped half-eaten fruit into pigsties. People get it from pigs and bats, but it can also spread between humans. Nipah breaks out sporadically in and around densely populated Bangladesh, causes inflammation of the brain and has a high fatality rate.



Rift Valley fever

Widespread across Africa, this virus invaded the Arabian Peninsula in 2000, and could go further. It mainly infects cattle and is spread by mosquitoes; people can get it from mosquito bites or by eating infected beef. Symptoms are usually mild but it can cause haemorrhagic fever, which kills in half of cases.



SARS, MERS and emerging coronaviruses

These related bat viruses infect a range of mammals and have already emerged in humans twice, resulting in severe pneumonia: SARS in 2003 and MERS in 2014. Both spread from human to human.

Crimean-Congo haemorrhagic fever

Found across Africa, Asia and south-east Europe, the virus is invading new territory as its tick hosts capitalise on global warming. It appeared in western Europe in 2010. Infected people generally have a mild fever but some strains cause severe haemorrhagic disease, with bleeding internally and from orifices, from which 30 per cent of people die.

Chikungunya

A virus spread by *Aedes* mosquitoes between monkeys and small mammals in East Africa, Chikungunya started causing large epidemics around 2000 and exploded into Asia in 2005, after mutations made it better adapted to a new mosquito host. In 2014, it invaded the Americas and has occurred in Europe. It rarely kills but causes debilitating joint pains, which can persist for months.

Zika

A monkey virus that has infected humans in Africa and Asia for decades, Zika suddenly entered the Americas in 2013. In 2015, it was linked to a wave of severe birth defects including microcephaly. Companies are already working on vaccines but the WHO wants extra research into the virus's effects on fetal brains.

Severe fever with thrombocytopenia syndrome

Flies under the radar – possibly because of its name. The virus, discovered in 2011, can cause fever and multi-organ failure, killing 12 per cent of people it infects. It has been found in east Asia, seems to be carried by farm animals, and is spread by ticks. A nearly identical virus, called heartland, has turned up in the US.

Novel agent

Given the rate at which previously unknown or obscure infections have suddenly emerged in humans and other animals, the WHO is leaving a slot on its list for a germ we don't yet know. Research here may include looking for agents that might explode.

affordable. That means their prices will be "delinked" from the cost of developing them, by making sure companies are recompensed in other ways. So far no one knows how that will work, but it is already being discussed for new antibiotics.

The WHO is not alone in trying to encourage the forging of weapons. In January 2016, the Coalition for Epidemic Preparedness Innovations (CEPI) was launched at the World Economic Forum in Davos, Switzerland, to help get experimental vaccines through the "valley of death". CEPI, which is backed by Norway, India, the Gates Foundation and the Wellcome Trust, has commitments of $540 million and, say organisers, is "on track" to get $1 billion for the next five years. By then it hopes to have vaccines against Nipah, MERS and Lassa viruses tested for safety and effectiveness in phase II trials. It even wants to have small stockpiles of the promising vaccines for fast response to outbreaks.

However, no one can afford phase III trials on larger numbers of people. And no one can test whether a vaccine works until there is an outbreak. Those tests may have to be done in a hurry once an epidemic starts.

MONEY MATTERS

As global economies become more interconnected, contagious diseases and their knock-on effects spread more rapidly. "Nowadays the biggest risk from epidemics is economic," says Ramanan Laxminarayan of Princeton University. The 2003 SARS epidemic killed 800 people, for example, but cost the world $54 billion in quarantine measures and lost trade and travel. The World Bank estimates that a flu pandemic as bad as the one in 1918 would lop 5 per cent off world GDP and cause an $8 trillion recession. The faster we respond to an epidemic, the less expensive it will be. So we must be prepared – and that costs. Who will pay?

One answer may be novel funding models. In 2016, the World Bank launched something new: plague insurance. Rich countries are at risk from epidemics that start in poor countries. So under the Pandemic Emergency Financing Facility they can buy insurance against severe flu, coronaviruses like SARS or MERS, filoviruses like Ebola, and diseases that pass between animals and humans like Lassa. Premiums are based on risk, calculated by epidemiological modelling company Metabiota. If such a disease strikes a poor country, money to contain it is released quickly from the insurance pot. The bank also sells "catastrophe" bonds to fund response to a wider range of epidemics.

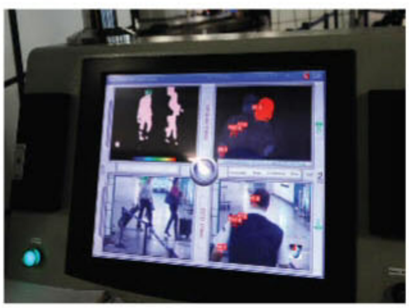
3 BE READY TO ACT - FAST

With potential mass killers identified, and drugs in hand, we will be on the right track. But we must also be ready to act fast on a large scale. Paradoxically, that means getting more familiar with what is normal, so we can spot ominous changes.

One problem is that contagion is exponential: case numbers rise very slowly at first, then sky-rocket. "First people complain that you are putting too much effort into a small problem. Later they say you were too slow," says Sylvie Briand, head of the pandemics department at the WHO. To better predict which outbreaks might take off, the WHO now has teams looking at the use of "big data", such as combining existing data sets on climate, vaccination and population immunity. It has also set up networks of social scientists and anthropologists to explore ways to improve communications among people swept up in plagues – a major roadblock to rapid response during the Ebola outbreak. The first and fundamental problem there, however, was surveillance: no one spotted the first few cases of Ebola before it spread widely.

"To get ready for the big one, we need health workers close to the entire population, everywhere, who know where to go if something funny is going on – then labs to test samples, and response teams," says Seth Berkley, head of GAVI, a global alliance that helps poor countries get routine vaccines. Under a 2005 treaty called the International Health Regulations, all 192 WHO member states must set up enough surveillance to tell the WHO about any outbreak that is serious, unusual or could trigger international travel or trade restrictions. However, not one world region, even Europe, has done everything the treaty requires. Africa, home of many worrying viruses, has done least.

An international collaboration called the Global Health Security Agenda is trying to help countries fill the gaps – and Ebola has scared many into listening. "There has been a change of mindset," says Briand: watching existing health risks more closely will help countries spot new ones.



OGNEN TEOFILOVSKI / REUTERS

Going global: international flights spread pathogens

Emergency responders

In addition, the WHO, which has always been a technical agency, setting policies by slow consensus, has reinvented itself to respond faster in an emergency. Instead of independent offices in different countries spotting emergencies – or not – according to their own criteria, the WHO now has dedicated staff worldwide who can do standardised assessments of unusual events, deploy emergency teams within 72 hours and scale up quickly. To aid coordination, they are answerable to the head office in Geneva, a first for the WHO.

The agency is also working with the World Food Programme to set up global supply chains for equipment such as masks and syringes. It will launch an online course to train emergency responders. And it is working with the Inter-Agency Standing Committee, a Geneva-based ›

Plagued by plagues

Infectious disease used to account for half of all human deaths before the rise of modern medicine, now globalisation is renewing that threat

Athens
430 BC
75,000-100,000
Plague of Athens
Typhus/smallpox?

Europe, Asia
AD 541-542
25-50 million
Plague of Justinian
Bubonic plague

Europe, Asia, Africa
1347-1351
75-200 million
Black Death
Bubonic plague

Mexico
1545 and 1576
17 million
Smallpox

Europe
1665-1666
100,000
Great plague of London
Bubonic plague

Asia, Europe
1817-1824
100,000+
First cholera pandemic
Cholera

Worldwide
1918-1920
20-50 million
Spanish flu
Influenza

Worldwide
1980s-
35 million
HIV pandemic
HIV

West Africa
2013-2016
11,300+
Ebola

1300 1400 1500 1600 1700 1800 1900 2000

body that coordinates the world's emergency responses to war and natural disasters, which last year expanded its remit to epidemics.

But, no matter how fast you detect outbreaks, or how many drugs or vaccines you invent, you still face the problem of producing and deploying enough of them to make a difference. "You can't build a vaccine factory and only switch it on in an emergency," says Martin Friede at the WHO. Like standing armies, production lines and staff need honing and updating.

A possible solution for limited manufacturing capacity comes from ongoing efforts to control flu – one pandemic we know for sure will come. The flu vaccine is made of a standard, benign flu virus with two new proteins from whatever strain is circulating that year stuck onto it to induce immunity to that strain. The vaccine changes every year, but doesn't need new plants or regulatory approval as the package is well tested. "We can produce safety-tested vectors at scale, then drop in antigens of interest if a new disease emerges," says Berkley. "That way, you can build vaccine capacity for a pathogen you don't even know."

That isn't happening yet. Nor is it clear if the WHO will get enough funds to continue any of this work, especially with a new US president who has opposed UN

ROBERTO CACCURI/CONTRASTO/EYEVINE

funding. "The really big problem is appreciating what is at stake," says Berkley. He says a pandemic is an "evolutionary certainty". "If people understood the risk, they would want to be sure systems are in place to deal with it. The costs of doing that are trivial compared to the cost of ignoring it."

We have been jolted out of our complacency, but there's still a lot to be done. "With Ebola the world recognised that the largest unmanaged risk to the global economy and security is infectious hazards," says Bruce Aylward, assistant director-general at the WHO. "Are we prepared for pandemics? Definitely not! Are we more prepared? Definitely." ■

Even supplying equipment in a crisis will be a challenge

HOW YOU CAN REDUCE THE RISK OF A PANDEMIC

In our increasingly crowded, urban, globalised world, a virus will eventually get out of control. There are things we can all do to reduce the risks.

Bear witness: Inform yourself and do what you can to spread awareness of the risks, and of the responses being devised that desperately need support. Politicians control purses, so get tweeting.

Stand up to denialists: Some will say warnings about pandemics are a hoax, because SARS/bird flu/swine flu was supposed to kill us all and didn't. Here's your riposte: a lot of people worked hard to keep SARS contained; bird flu hasn't gone rogue yet but it's a few mutations away; swine flu did kill and the next flu could kill far more.

Prepare: You needn't be a survivalist to prepare for the panic and disorder likely to attend a pandemic. Most countries have guidelines that recommend stocking a few weeks' worth of water, food, medicines, flashlight batteries and such. Learn about the best ways to avoid people who might be contagious. If you run a business, have a continuity plan. If you are a public official, check whether your administration has a pandemic plan. If not, check out the WHO's guidelines. If you speak for a health body or organisation, learn about communications in a pandemic because mistakes can be deadly. Hint: trust people with the truth.

Keep watch: Countries don't like to admit they have infectious diseases: it's bad for business. The ProMed global reporting site revealed SARS and MERS before the governments involved did. It has since helped launch Epicore. Medical and veterinary workers sign up to it, then when ProMed gets wind of something it asks them what's happening. Replies appear on a web platform that can be set to partial or total confidentiality. Wherever you are, if you meet the criteria, sign up to Epicore. You could be the first to spot something amiss.

PLAINPICTURE/ROBERTHARDING/JON PURCELL

Cutting through the smog

Is air pollution in the West really as bad as it seems? Nic Fleming investigates

THE bad news on bad air seems to get worse by the day. Air pollution causes one in nine early deaths, according to the World Health Organization. It is a major health concern in both rich and poor countries. In the US, it is estimated to be behind 200,000 untimely deaths each year. The UK, where the annual death toll is reported to be in the tens of thousands, was one of five countries threatened with legal action for repeatedly breaching the European Union's nitrogen dioxide limits in February 2017.

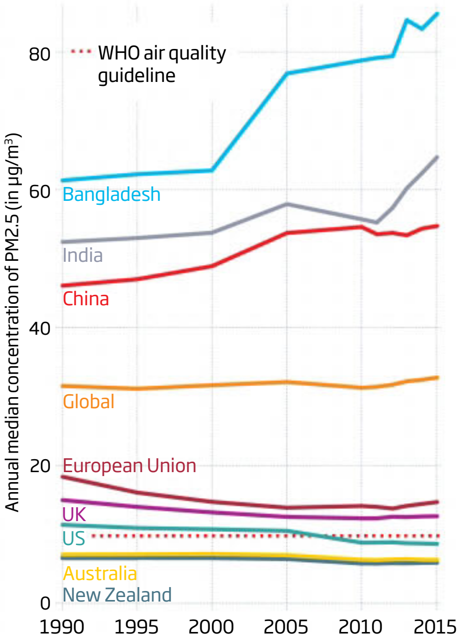
But all is not as it seems. Quantifying the impacts of polluted air is a more complicated and uncertain business than many headlines would suggest. The growing pressure for politicians to take action raises important questions. Is air pollution really getting worse? How bad is it? And what should be done about it?

Q.IS AIR POLLUTION REALLY GETTING WORSE?

In rapidly growing economies, the amount of pollution in the air is undeniably rising, but it is a different story in most rich countries. Take, for instance, PM2.5 particulates – believed to account for most of the health burden of air pollution (see "What's in the air", opposite). Worldwide, average concentrations rose 11 per cent between 1990 and 2015, according to a report by the Health Effects Institute and the Institute for Health Metrics and Evaluation, both in the US. The trend reflects large increases in India, Bangladesh and China: concentrations in the US, the European Union, Canada and Australia fell over the same period (see graph, right).

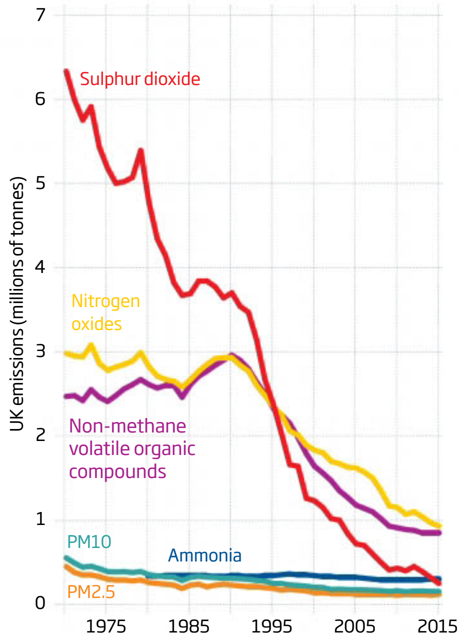
Media reports on air pollution in the West frequently don't mention the major improvements made since the 1950s. But the rate of progress has slowed and Europe, including the UK, is showing no signs of meeting WHO guidelines for clean air any time soon. "The data from monitoring sites across western Europe shows PM2.5 levels are going down," says Gavin Shaddick of the University of Bath, UK, who develops air pollution models for the WHO. "But they are not falling quickly enough."

Air pollution levels vary widely by country



SOURCE: SHADDICK ET AL

In the UK, emissions are dropping



SOURCE: DEFRA

Cycle or drive?
Cycle, and take the back streets. It's win-win: less pollution and more exercise. The same goes for walking. (see "Leave the car at home", page 53)

Q.IS IT KILLING ME?

Figuring out how many people die because of pollution is a tricky business. Widely quoted numbers vary enormously and mask a great deal of complexity, uncertainty and misunderstanding (see "How many deaths", opposite).

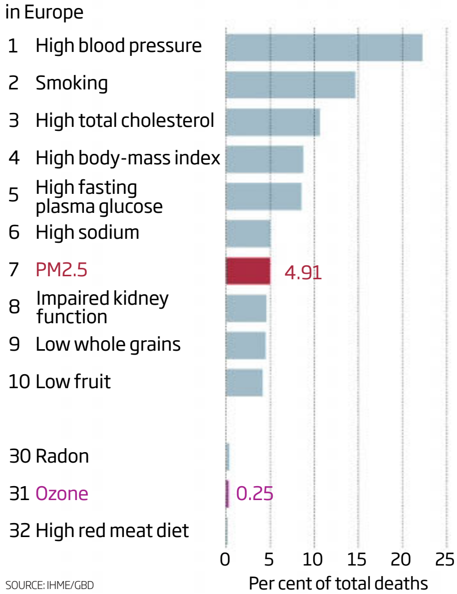
First, it's important to realise that nobody drops dead from walking down a polluted street. Rather, air pollution aggravates other things that are likely to kill you, cutting months off your life. The UK Committee on the Medical Effects of Air Pollutants (COMEAP) estimates that anthropogenic PM2.5, released at 2008 levels, would shorten the average person's lifespan by six months. By totting up all this lost life, the group worked out that outdoor air pollution would cause the equivalent of almost 29,000 deaths. COMEAP stressed that PM2.5 was shortening the lives of many more people rather than causing that number of deaths. Yet the media often reports that air pollution kills 29,000 a year in the UK.

COMEAP's calculations were based on a number of assumptions. One question is whether all PM2.5 has the same health effects regardless of its source. "I think everybody who studies this believes there are differences," says Michael Brauer of the University of British Columbia. "But it's been hard to consistently demonstrate what they are."

A further complication comes from the probable overlap between the effects of PM2.5 and NO_2 – the two most harmful pollutants. "My personal belief is that the NO_2 epidemiology is largely a signal due to ultra-fine particles, which would already have been largely counted within PM2.5," says Jon Ayres at the University of Birmingham, UK, who was chair of COMEAP until 2011.

International estimates of mortality due to air pollution have also varied dramatically, but for Brauer, quibbling over numbers misses the point. "The actual number makes for a nice headline, but it's probably not that important," he says. "What is important is that we can see how air pollution compares to other major risks such as smoking, and so can prioritise policy and funding."

Air pollution ranks in the top 10 mortality risks in Europe



SOURCE: IHME/GBD

NICOLAS ASFOURI/AFP/GETTY IMAGES

HOW MANY UK DEATHS?

Estimates of the annual deaths attributable to air pollution vary wildly

In 2010, the Committee on the Medical Effects of Air Pollutants estimated 29,000 annual UK deaths were attributable to particulates less than 2.5 μm across (PM2.5), assuming no safe limit

Uncertainties meant there was a 75 per cent chance the number could be anything between 5000 and 55,000 deaths

In 2012, the World Health Organization put the figure at 16,400, assuming a safe limit of 7 μg/m^3

44,750-52,500 is the UK government estimate of deaths from PM2.5 and NO_2

40,000 is often quoted in the press as the number of deaths from air pollution. It comes from the UK Royal College of Physicians and Royal College of Paediatrics and Child Health, which estimate that PM2.5 and NO_2 cause between 30,000 and 50,000 deaths a year

Q. HOW DOES POLLUTION AFFECT MY HEALTH?

We still have a lot to learn about how outdoor air pollution causes ill health, not least because its effects on our bodies are likely to be multiple, complex and interdependent. Studies suggest PM2.5, NO_2 and ozone mess with oxidation reactions in the lungs and elsewhere in the body. This triggers inflammation and can cause tissue damage.

Most studies look for correlations between increased exposure to pollution and the prevalence of diseases. For instance, a 2014 study that followed some 100,000 people in five European countries for more than 11 years found that a 5 μg/m^3 increase in annual average PM2.5 exposure was associated with a 13 per cent increase in either heart attacks or unstable angina. Another study found the same increases in PM2.5 were associated with an 18 per cent increase in the risk of developing lung cancer.

There is also a well-established association between pollution and respiratory and pulmonary diseases, and stroke. A study published in January 2017 found that people living within 50 metres of a major road were 7 per cent more likely to develop dementia than those who lived 300 metres or more ›

WHAT'S IN THE AIR

Particulate matter (PM)

Comes from: power plants, factories, gas cookers, industry, volcanoes, dust storms, forest fires, car exhaust.

Health effects: Any dust or droplet less than 10 micrometres (μm) across (called PM10) can penetrate deep into your lungs. Those smaller than 2.5 μm (PM2.5) are the most damaging air pollutants. They include ultra-fine particles, which are smaller than 0.1 μm. Long-term exposure to PM2.5 can impair lung and heart function and increase mortality, especially among those at higher risk of heart disease and stroke.

Nitrogen oxides (NO_x)

Comes from: road transport – especially diesel engines – as well as indoor heating and power stations.

Health effects: NO_2 is assumed to be the second most harmful pollutant after PM2.5. Exposure can trigger respiratory problems. In lab studies, volunteers given extra NO_2 exhibit inflammatory responses. However, long-term effects are unclear.

Ground-level ozone (O_3)

Comes from: reactions between other chemicals including NO_x and volatile organic compounds, especially on warm, sunny days.

Health effects: ozone can cause wheezing, shortness of breath, inflamed and damaged airways, a range of lung diseases, exacerbated asthma. Children, elderly people and those who are most active are at greatest risk. O_3 is a powerful oxidant, so damages cells and tissues.

Sulphur dioxide (SO_2)

Comes from: burning fossil fuels, especially coal. As a result, levels have dropped significantly in the West where natural gas has largely replaced coal.

Health effects: irritated airways and eyes, breathing problems, heart and circulation problems. Those with asthma and other respiratory conditions are at greatest risk. SO_2 can form sulphates that become PM2.5, and it causes acid rain.

Ammonia (NH_3)

Comes from: decomposing organic matter, livestock and fertilisers.

Health effects: At the concentrations that it is present in the air, NH_3 is unlikely to harm human health. Its main effect is to acidify the soil.

away. Other research has linked air pollution with diabetes, kidney diseases, Alzheimer's, premature births and mental illness.

There is also growing evidence of effects on child development. A 2004 study found that 18-year-old Californians who had been exposed to 28 μg/m[3] of PM2.5 per year for eight years, on average, were 4.9 times more likely to have reduced lung function than those exposed to an average of 5 μg/m[3]. Researchers found delayed cognitive development in children in Barcelona who went to school in polluted areas.

A new US-UK-Chinese collaboration, led by Frank Kelly of King's College London, should offer a more precise understanding of the links between pollution and ill health. The study will give 120 Beijing residents and 120 people living in an outlying village portable pollutant monitors. It will then compare exposures with health data taken from urine and blood samples to help understand what pollution does to our bodies.

"Moving from just estimating people's exposure to actually measuring it and linking that to biological response markers is a major step forward," says Kelly.

Q. WHAT CAN WE DO ABOUT IT?

"Unlike finding a cure for cancer, we know how to tackle this problem because we've done it before," says Michael Brauer of the University of British Columbia. "The new laws introduced in the UK in the wake of 1952 [the pea-souper smog that killed 12,000 Londoners] and the way California has set standards to force industry to innovate and become cleaner point the way."

Still, solutions evade us. In the West, transport is the main cause for concern. Per capita car ownership roughly doubled between 1970 and 2012 in most of North America, western Europe, Australia and New Zealand. Cities have tried to keep cars off the road in several ways. Paris only allows vehicles with odd or even licence plate numbers on certain days. Freiberg in Germany has focused on providing cheap, efficient public transport. London and Stockholm have introduced congestion charges.

Evaluating these schemes is a challenge, because you would need to isolate their impact on pollution from other factors that might have also changed pollution levels. The best available evidence suggests many major efforts to reduce pollution from traffic have either failed or had little measurable effect.

SAM EDWARDS/PLAINPICTURE

Inside or outside of pavements?
That's a tricky one. Some studies show you get a lot less pollution just by moving a couple metres away from cars, but how the wind interacts with buildings can change everything.

Some road space rationing schemes have had perverse knock-on effects. According to some reports, people in Mexico and Beijing have started buying second vehicles with different licence plates to get around restrictions. Often the second car will be cheap and more polluting.

Charges or fines for taking larger, more polluting vehicles inside low-emission zones may be no better. There are more than 200 of these zones in Europe. But a 2015 evaluation of the London LEZ, the largest in the world, found it had no impact on levels of pollutants or related respiratory and allergy problems in children. The study authors speculated this was because of a delay in the introduction of tougher European standards for light goods vehicles and an increase in the proportion of diesel cars, encouraged by tax incentives. The results mirrored those in five Dutch cities, where LEZs had little effect on traffic-related pollution levels. A 2011 assessment of London's congestion charge scheme also found no compelling evidence that pollution had fallen two years after it was introduced.

Copenhagen and Amsterdam lead the world in getting their inhabitants to ditch their cars by providing better cycling infrastructure, but other governments are failing to put their money where their mouths are. In 2016, the UK government announced ambitious plans to double cycling journeys by 2025, reverse the decline in walking, reduce cycling fatalities and increase the proportion of children walking to school. The £316 million it has dedicated to achieving these goals over five years in England outside London is dwarfed by the £15 billion budget for major road improvements over the same period.

In the end, outside highly centralised,

Car or bus?
Bus. Studies suggest you inhale less pollution in a bus than inside a car (see "Leave the car at home", below).

authoritarian states, there is only so much governments can do. Individuals also need to realise the pollutants they breathe are produced when they drive short distances, fail to insulate their homes which would reduce heating bills, or buy intensively farmed meat, which produces a lot of NO_x.

Figuring out what you can do to make the most impact will depend on where you live. A European Union funded project, ClairCity is customising pollution models to help individuals in various cities to identify behaviours that will have the maximum impact in these locations. "Ultimately, we control our own behaviour," says Gavin Shaddick of the University of Bath, UK. "The tipping point will come when communities, societies and groups of people decide to do something about it."

Q. FIVE WAYS TO PROTECT YOURSELF AND YOUR FAMILY

1. Leave the car at home
Many drivers think windows protect them from the pollution they are generating. They are wrong, mostly. A 2012 study of commuters in Barcelona found that car passengers in busy traffic were exposed to 30 per cent more carbon dioxide than cyclists or pedestrians who can move faster along the same route, 25 times more carbon monoxide, and between two and three times more PM2.5 and ultra-fine particles. One reason may be that the air intake on most cars is at the front, right behind the exhaust pipe of the vehicle in front, says Shaddick.

There's a catch. Although they were exposed to less pollution, the active commuters inhaled more of it. Overall, bus commuters inhaled the least.

Some studies come to different conclusions. Outcomes vary with traffic speed, vehicle and fuel type, and the weather. "There is no one answer," says Ben Barratt of King's College London. "In slow-moving congested traffic, you will be more exposed than in free-flowing traffic. But the exercise benefits of walking and cycling generally outweigh any negatives."

2. Escape to the country, or a side street
Sometimes the simplest solutions are most effective. In this case, that means getting away from the source. Concentrations of NO_x and PM2.5 from car engines fall exponentially over the first few metres from the exhaust. "Moving from a congested road to a quieter street makes a big difference because you are much farther away and there are buildings in between you and the source," says Barratt.

Levels of NO_2, carbon monoxide and particulate matter are higher in towns and cities because of traffic. Country dwellers are less likely to suffer associated health risks, but ozone levels are on average higher in rural settings. Other pollutants can be just as elevated in the countryside as they are in urban environments. But overall, you are better off living in country lanes.

3. Wear a face mask
If you're shopping for a face mask, look for one with an N95 rating. It has been certified by the US National Institute for Occupational Safety and Health as filtering out 95 per cent of airborne particles larger than 0.3 µm. Many dust masks will fit this bill. Lab tests show that cotton handkerchiefs block just 28 per cent of particles from a diesel exhaust engine, and the simple cloth masks that tie behind the head won't help much more. Cycling masks vary from 55 to 85 per cent effectiveness. One study suggests that those with exhalation valves may work better. Surgical masks are surprisingly good: they can filter 80 per cent of particles in the lab. The results might not be replicated on the street though.

Masks won't keep out gases like NO_2 and SO_2 unless they have special components like charcoal filters. Some masks may help keep your blood pressure low. "If you have symptoms that worsen in polluted environments, a mask could be a sensible precaution," says Barratt. The key point in all cases is fit: all bets are off if a mask doesn't perfectly follow the contours of your face.

4. Give pushchairs the push
Children are at greater risk from pollution because their lungs and brains are still developing. Most vehicles release their exhaust fumes somewhere between 30 and 60 centimetres from the ground.

"Being at exhaust level is bound to give you a higher exposure," says atmospheric scientist Rob MacKenzie at the University of Birmingham, UK. One 2009 study found that infants in prams were exposed to twice as much particulate matter from diesel exhaust as the adults accompanying them. Yet research published in February 2017, found no difference between PM2.5 concentrations for babies in buggies and for adults pushing them, but did identify busy intersections and bus stops as pollution hotspots. The lead author of that study, Prashant Kumar of the University of Surrey, UK, suggests parents use pram covers, especially near busy traffic and intersections.

5. Get some fresh air
Air filters in most modern vehicles are designed to block large particles like dust and pollen, but are unlikely to capture all fine soot. Gases such as NO_x and ultra-fine particles will get through. Tests have shown that switching to a vehicle's recirculation mode can reduce ultra-fine particles by around 90 per cent, but will push up carbon monoxide levels.

Although some of their claims have been questioned, car manufacturers are increasingly including higher performance filters. Tesla (see "Help for heroes" page 121) says its high efficiency particulate air filter can reduce exposure to PM2.5 from very high to negligible. Air purifiers for buildings are becoming commonplace in China. Such devices can work well on sealed commercial buildings and potentially flats, but are less effective in leaky houses. ■

PLAINPICTURE/ALEKSANDAR ZAAR

It's not just tabloid newspapers that misrepresent medical statistics for dramatic effect, warn **Marianne Freiberger** and **Rachel Thomas**

Careless pork costs lives...

...and other medical myths

TYPE the word "cancer" into the website search engine of the *Daily Mail*, a British tabloid newspaper, and a wealth of information is just a mouse click away. Some of the reports are calming, most alarming – and all come with figures to back them up. Women who use talcum powder are 40 per cent more likely to develop ovarian cancer, says research. Cancer survival rates in the UK are among the worst in Europe, according to a study. The incidence of bowel cancer among the under-30s has soared by 120 per cent in 10 years, astonishing figures show.

The figures might make us worry for our health, but somehow we feel the better for their existence. Numbers help us make sense of the world: if you can put a number on a problem, then its extent is known and its impact can be circumscribed.

Yet that sense of solid certainty is all too often illusory. Statistics can be slippery, easily misused or misinterpreted. Nowhere is that more true than in the field of human health.

That's because the benefits of a particular medical treatment are often not obvious. "There are very few miracle cures. Most treatments require careful science to determine if there is any benefit and how big the benefit is," says David Spiegelhalter, a biostatistician at the University of Cambridge. "Working out the effects of an environmental risk factor is even more tricky," he adds. Saying anything sensible about human health requires large, reproducible clinical trials, and the careful observation of diverse populations – all of which implies the use of statistical methods to extract workable conclusions from the data.

The British epidemiologist Austin Bradford Hill recognised this when, in 1946, he ran the first trials in which participants were randomly assigned to two groups, one of which received the treatment and one of which didn't. One of these trials tested the effectiveness of the antibiotic streptomycin to treat tuberculosis, a condition that Bradford Hill himself had developed while serving in the first world war. After just six months, the results were so convincing that they led to streptomycin being adopted as the standard treatment. In 1950, together with Richard Doll, Bradford Hill used statistical methods to provide the first convincing evidence that smoking causes lung cancer.

Used well, statistics are a powerful tool. But caution is required. Sample size, the design of a study and even the definition of terms or the way a number is presented can all affect the value of the headline statistics we are offered. Generally, we are not privy to these details.

What's more, the decisions we take concerning health are often made at times of intense emotional stress. "People are very much influenced by culture, emotions and values when making judgements, and that's fine, that is part of being human," says Spiegelhalter. But it makes us all the more susceptible to seemingly incontrovertible numerical truths distilled into media headlines – and to the enthusiastic but sometimes equally misplaced insistence by researchers, doctors or advocates of a new treatment that it will do us good.

So when confronted with medical statistics, how do we know whether they are the real deal, or distorted before they get to us? How do errors creep in? What are the questions we need to ask to avoid falling for them?

YOUR NUMBER'S UP

Ratio bias

What would worry you more: being told that cancer kills 25 people out of 100, or that it kills 250 people out of 1000? Dumb question, you might say; both statements mean that a quarter of people die of cancer.

Yet such differences do matter – not to the risk itself, but to our perception of it. Those wishing to play up or play down a risk, whether to sell newspapers or a medical treatment, can follow the simple rule of "ratio bias". The bigger the number, the riskier the risk appears.

In one study of this effect, people rated cancer as riskier when told that it "kills 1286 people out of 10,000" than when told it "kills 24.14 people out of 100", even though the second statement equates to almost double the risk. Similarly, another study showed that 100 people dying from a particular form of cancer every day can be perceived as a lesser risk than 36,500 dying from the same disease each year, although the two are equivalent statements.

So when confronted with questions of risk, look carefully at the way the numbers are presented (see "Get smarter", page 12). And if you are comparing risks, make sure they are divided by the same number.

MORE HARM THAN GOOD?

Relative versus absolute risk

Is there anything that has not been claimed to cause cancer? Over the years we have learned, among other things, that drinking very hot cups of tea leads to an eightfold increase in the risk of developing oesophageal cancer; that a quarter of a grapefruit a day increases breast cancer risk by 30 per cent in post-menopausal women; and that a daily bacon sandwich raises the likelihood of bowel cancer by 20 per cent. This last finding was encapsulated by the British tabloid *The Sun* in the headline "Careless pork costs lives".

These assertions may or may not be valid, but hidden within them is a more important and insidious source of confusion. The figures quoted measure relative risks: how much more likely you are to get ill when indulging in the supposedly dangerous substance or activity compared with not indulging. But they tell you nothing about what that increase in risk amounts to in absolute terms, so there is no way of telling whether it is something worth being concerned about.

"For an average person, the chance of getting bowel cancer at some point in their life is around 5 per cent," says Spiegelhalter. So a 20 per cent relative increase in bowel cancer risk translates to an absolute increase in risk from 5 per cent to 6 per cent – just 1 per cent. That's big enough not to ignore, but less of a deterrent to those who like their daily bacon sandwich.

Journalists are by no means the only ones who exploit the greater headline-grabbing potential of relative risk; health professionals do it too. "One of the most misleading, but rather common, tricks is to use relative risks when talking about the benefits of a treatment, while potential harms are given in absolute risks," says Spiegelhalter.

This technique is known as mismatched framing. In his book *Reckoning with Risk*, psychologist Gerd Gigerenzer of the Max Planck Institute for Human Development in Berlin, Germany, quotes the example of a patient information leaflet concerning hormone replacement therapy. It claimed that HRT cuts the risk of bowel cancer by 50 per cent (a relative risk), but leads to 6 extra cases of breast cancer per 1000 women (an absolute risk). At first glance, the benefit here seems to hugely outweigh the additional breast cancer risk of just 0.6 per cent.

But until we know the absolute rates of bowel cancer in the target population, we are none the wiser. Assuming that rate is 5 per cent, as it is in the general population, the reduction in risk is 2.5 per cent, putting the benefit to harm ratio in a very different light.

Once you are aware of this trick, it's relatively easy to spot, but this doesn't eradicate it even from peer-reviewed medical journals. According to a study published in 2007, one-third of papers reporting on the benefits and harms of medical interventions in the *BMJ*, *The Lancet* and *The Journal of the American Medical Association* presented them using a mixture of different measures.

EVGENIY IVANOV/GETTY

Some like it hot – but is tea that's too hot a significant cancer risk?

Scary or not?

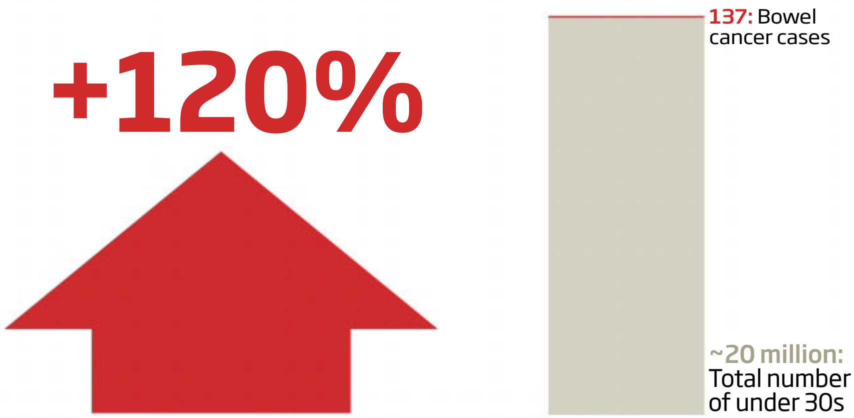
Different ways of presenting the same data can greatly influence our perception of risk

"Bowel cancer soars by 120% among the under 30s"

Daily Mail, 31 March 2009

137 people under 30 were diagnosed with bowel cancer in 2006 in England & Wales

up from 63 cases in 1997



"One-third of academic papers reporting medical benefits and harms used a mixture of statistical measures"

"Two things moving in tandem does not necessarily mean that one thing is causing the other to move"

TV KILLS

Correlation vs causation

It isn't surprising that a study with the title "Television viewing time and mortality" grabbed the headlines. It asked 8800 people about their health, lifestyle and television watching behaviour, and then followed them over the next six years, during which time 284 of them died. Among people who spent more than 4 hours a day in front of the TV, it found, the risk of their dying within the period of the study was 46 per cent higher than among those who watched less than 2 hours a day.

The sort of headlines generated – "TV kills, claim scientists" – were also predictable. But this is one case where two variables moving in tandem (correlation, in other words) does not necessarily mean that the change in one is responsible for change in the other (causation). In fact, the researchers were not primarily interested in TV viewing. They wanted to measure the amount of time people spent sitting still, and used TV watching as a shorthand for this; they explicitly excluded time spent watching TV while doing other, active things, such as ironing.

"At best, this study shows that sedentary behaviour, for which hours of TV watching is a proxy, is associated with modest elevations in death from heart disease and from all causes," says Nigel Hawkes, a health journalist and formerly the director of Straight Statistics (straightstatistics.org), a campaign to improve the use of statistics in the public arena. "There is nothing intrinsic in television that makes people more likely to die."

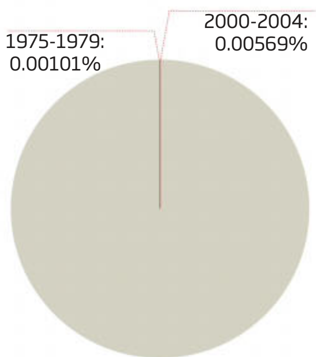
You don't have to look far to find confounding variables that might have been at work. People with certain underlying health problems sit or lie still for long periods, possibly in front of the TV, and these problems might also be associated with a raised risk of early death. Despite the study's apparent conclusions, it's probably still safe to switch on and zone out.

Before assigning cause and effect, it is essential to read between the lines. Bradford Hill identified the crucial question: Is there any other way of explaining the set of facts before us; and is any such explanation equally, or more, likely than cause and effect? The answer needs to be a resounding no.

Fox News
27 August 2008

"Throat cancer among white men up 400% in 30 years"

Percentage of white male population in the US diagnosed annually with adenocarcinoma of the oesophagus
Journal of National Cancer Institute, vol 100, p 1184



SIZE MATTERS

Clinical trial design

"Over 80 per cent of women say that this shampoo leaves their hair healthier and shinier." Such claims are common in advertising for all manner of consumer products. What they might not tell you is that only five women tested the shampoo. And of the four who certified its miraculous effect, one or two probably ended up with nicer hair purely by chance, or simply imagined the results.

Similar caveats apply to the effectiveness of medical treatments. Curing six out of 10 patients is promising. Curing 300 out of 500 is the same success rate, but far more convincing. "The sample size in a test is absolutely crucial in deciding whether any apparent improvement could have happened by chance alone," says Spiegelhalter.

The standard procedure for such trials is the one established by Bradford Hill more than 60 years ago: new medical treatments are tested in randomised controlled trials (RCTs), in which volunteers are randomly allocated to a study group that receives the new treatment or a control group that receives a placebo or existing treatment. "You can think of an RCT almost as a measuring instrument to measure a treatment's effectiveness," says Sheila Bird of the UK Medical Research Council Biostatistics Unit in Cambridge. To make sure any instrument is sensitive enough for its job, you need to assess how big an effect it is expected to measure.

Working out the size of the expected effect requires an analysis of past studies or the results of tests on animals. In the case of an RCT, the smaller the expected effect, the more people you need to enrol in your trial, and vice versa.

Another important consideration is the level of significance the trial is expected to achieve – that is, the likelihood that a useless treatment will register the effect you are after as a result of chance alone. RCTs are usually designed to achieve a 5 per cent significance level. This means that even if the drug is useless, it will register a positive result by chance in 1 out of 20 trials. For that reason, says Spiegelhalter, drug licensing authorities do not usually consider a single study sufficient evidence to approve a new drug. Repeat trials are needed.

So next time you hear of public acclaim for a miracle cure or wonder shampoo, ask three questions. How many people was it tested on? Was it tested in an RCT? And was the result confirmed by a second, independent test?

"Rudy Giuliani claimed you were only half as likely to survive prostate cancer in the UK as in the US. He was right – but also wrong"

Is the secret to a long life laying off the grapefruit?

EDWARD POND/GETTY

DIE ANOTHER DAY

Survival vs mortality

There can be few things in US politics more poisonous than discussions about healthcare. Over the years, the arguments have been accompanied by all sorts of dodgy claims and counterclaims, often with statistical evidence to back them up.

Take the statement by former New York City mayor Rudy Giuliani in his campaign to win the 2008 Republican presidential nomination. He quoted the chance of a man surviving prostate cancer – a disease he had himself experienced – as 82 per cent in the US, and compared this with a chance of just 44 per cent under the UK's taxpayer-funded National Health Service.

Survival rates a factor of two apart in two comparably developed countries? If right, surely that would be a damning indictment of the deadly inadequacy of socialised medicine. And there's no doubting Giuliani's figures were right.

Right – but also misleading. "Giuliani's numbers are meaningless for making comparisons across groups that differ dramatically in how the diagnosis is made," observed Gigerenzer and colleagues in a 2008 paper on risk communication.

That is because Giuliani was quoting five-year survival rates – the number of people diagnosed with a disease in a given year who are still alive five years later. But while prostate cancer in the US is generally diagnosed through screening, in the UK it is diagnosed on the basis of symptoms. Screening tends to pick up the disease earlier, leading to one source of bias in the comparison.

Suppose that of a group of men with prostate cancer all die at the age of 70. If the men do not develop symptoms until they are 67 or later, the five-year survival rate based on a symptoms approach is 0 per cent. Suppose, instead, that screening had picked up the cancer in all of these men at age 64. The five-year survival rate in this case is 100 per cent, despite the fact that mortality is the same. Better survival rates don't necessarily indicate a better outcome.

That is obviously an oversimplification, as earlier diagnosis through screening presumably increases the chance that corrective measures can be taken. But screening is not 100 per cent accurate. First there are false positives, in which the test incorrectly flags a healthy person as having cancer. Prostate screening also picks up non-progressive cancers, which will never lead to symptoms, let alone death. The exact extent of this overdiagnosis is unclear, but a rough estimate is that 48 per cent of men diagnosed in this way don't have a progressive form of the cancer.

Tricky comparison

False diagnosis and overdiagnosis both result in unnecessary treatment, and, potentially, significant harm – in the case of prostate cancer, men left impotent and incontinent. But overdiagnosis also inflates the five-year survival rate by including men who would not have died of prostate cancer anyway. "In the context of screening, survival is a biased metric," says Gigerenzer. "The bottom line is that to learn which country is doing better, you need to compare mortality rates."

The annual mortality from a disease is the proportion of people in the whole population who die from it in a given year. So which comes out better, the US or the UK? Figures from the period 2003 to 2007 published by the US National Cancer Institute indicate an age-adjusted mortality from prostate cancer of 24.7 per 100,000. Similar figures from Cancer Research UK for 2008 point to a mortality of 23.9 per 100,000. In statistical terms, that is a dead heat. Higher survival does not necessarily mean fewer deaths.

This kind of bias makes it tricky to compare survival rates in different countries, a difficulty often explicitly acknowledged by the authors of academic studies that use the metric. Equally often, that subtlety is overlooked by politicians and journalists in search of a shocking sound bite or headline.

So next time you are told that one country outperforms or underperforms another on some vital metric of health, take a close look at whether it is survival or mortality that is being quoted. If it's the former, take the figure with a pinch of salt. Be aware, though, that this may increase your risk of heart disease. ■

DAVID LEMM

Automated systems are running the show – often in secret. **Hal Hodson** lifts the curtain

NO ONE IN CONTROL

"AMAZON is all kinds of broken." If you caught that tweet on 12 December 2014, and were quick, you might have grabbed some exceptional bargains. For an hour only, Amazon was selling an odd mix of items – cellphones, video games, fancy-dress costumes, mattresses – for one penny.

The surprise price drop cost sellers dearly. Goods usually marked £100 went for a 99.99 per cent discount. Hundreds of customers leapt at the chance, often buying in bulk. Even though Amazon reacted quickly and cancelled many orders, they were unable to recall those that their automated system had already dispatched from warehouses. Once set in motion, the process was hard to stop. Thanks to a software glitch, a handful of independent traders using Amazon's Marketplace lost stock worth tens of thousands of dollars. Some faced bankruptcy.

We only notice when algorithms go wrong. Most of the time they get on with business out of sight and out of mind. And business is booming. Automated processes are no longer simply tools at our disposal: they often make the decisions themselves. Much of the news we read, the music we listen to and the products we buy are served up automatically, based on statistical guesswork about what we want. Invisible chaperones shape our online experiences. Systems we can't examine and don't understand determine the route we take to work, the rates we get for mortgages, and the price we see for airfares.

Many are proprietary and all are complex, pushing them beyond public scrutiny. How can we be sure they're playing fair? A new wave of algorithm auditors are on the case, intent on pulling back the curtain on the hidden workings and hunting for undue bias or discrimination. But is this the fix? Do algorithms need better policing, or must we accept their nature as a price we pay for our automated world?

There's nothing inherently mysterious about them: an algorithm is simply a set of instructions for getting something done. The trouble is that algorithms get nested inside or bolted on to others, interacting in ever more complex ways. It can also be hard to predict how algorithms will behave with real-world data once released into the wild.

The scope of their influence is often unclear. Some people swear blind that they've seen the price of flights on one website jump after checking out a rival site, for example. Others think that's bunk, an urban myth for our times. Such debates highlight the shadowy nature of today's systems.

Not only are most algorithms secret recipes, sometimes even the developers who wrote them are in the dark. When Aniko Hannak at the Central European University in Budapest, looked closely at how many of us have our search results skewed by factors like location and browsing history, she noted things even Google didn't know: for example, that around 12 per cent of searches get personalised. Google engineers thanked her. They'd never made such measurements and hadn't known the exact impact of their personalisation algorithms.

> "ALGORITHMS ARE NO LONGER SIMPLY TOOLS: THEY OFTEN MAKE THE DECISIONS"

Exposing hidden algorithms can cause outrage. That's what Christian Sandvig and his colleagues at the University of Michigan, Ann Arbor, found when they lifted the lid on Facebook's newsfeed algorithms, which decide which posts from friends and family we actually see. The team compared filtered and unfiltered feeds and found that Facebook's algorithms hid posts deemed uninteresting, according to unspecified criteria.

Around two-thirds of the participants in Sandvig's study had no idea that algorithms were deciding what they saw. Many were shocked and upset when posts from close friends or family were excluded. Some had been blaming themselves or their friends for the algorithms' work. "If you post something and it doesn't get any comments or likes, people assume that either their friends don't like the topic, or their friends don't like them," says Sandvig.

Even for news, it's a popularity contest. During the Ferguson riots in Missouri in 2014, for example, Facebook's newsfeeds were filled with posts about the Ice Bucket Challenge because these had hundreds of thousands of likes.

What Sandvig's team did for Facebook, Hannak and her colleagues are doing for other online activity. Hannak is interested in how algorithms can tailor prices to different shoppers. In one study, the researchers looked at how online retailers such as Walmart, Office Depot and Expedia varied prices according

"EVERYONE WAS HIGH ON FAST WEB-BASED LOANS. NO ONE ANTICIPATED THE PROBLEM"

to factors including a user's choice of browser, operating system and purchase history.

They found many instances of what they consider price discrimination though they are not sure of the rationale. Often the difference was small. Android users, for example, saw higher prices on about 6 per cent of items, though only by a few cents. In other cases, price quotes varied by up to $100. The greatest differences were typically seen between users who were logged in to a site and those who were not.

CRASH DAMAGE

Hannak and her colleagues now want to understand exactly how location influences search results. They have been simulating hundreds of Android phones and spreading them across Ohio using faked GPS coordinates. They have also been looking to see whether people from rich and poor neighbourhoods get different search results when hunting for financial services.

Evidence of that may already have come to light. Some think hidden algorithms played a part in the 2008 sub-prime mortgage crash. Between 2000 and 2007, US lenders like Countrywide Home Loans and DeepGreen doled out home loans at an unprecedented rate via automated online applications. "Everyone was saying what a great innovation it was," says Dan Power at the University of Northern Iowa in Cedar Falls. "Everyone was very high on these fast web-based loans. No one anticipated the problem."

The problem was granting so many high-risk loans without human oversight. Americans from minority groups suffered most in the resulting crash. Automated processes crunched through vast amounts of data to identify high-risk borrowers – who are charged higher interest rates – and targeted them to sell mortgages. "Those borrowers turned out to be disproportionately African American and Latino," says Seeta Gangadharan of the London School of Economics. "Algorithms played a role in that process."

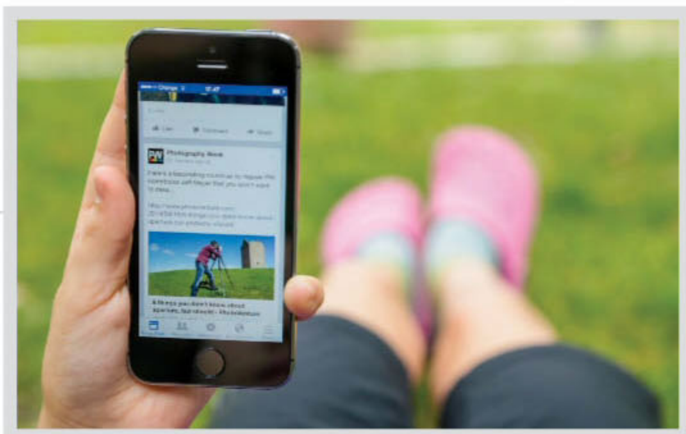
The exact degree to which algorithms were to blame remains unclear. But banks like Wells Fargo and Bank of America settled with several cities, including Baltimore, Chicago, Los Angeles and Philadelphia, for hundreds of millions of dollars over claims that their sub-prime lending had disproportionately affected minorities. Although the decision-making process big banks used to target and sell sub-prime loans may not have been new in itself, the reach and speed of those decisions when algorithms were the driving force was new. "It's the scale factor," says Gangadharan. "This was a problem that affected many people in the US and we have seen the effects fall along race and class lines in devastating ways."

Automated systems are replacing human discretion in ever more important decisions. In 2012, the US State Department started using an algorithm to randomly select the winners of the green card lottery. The system was buggy, however: it awarded visas only to people who applied on the first day, says Josh Kroll, a computer scientist at the University of California, Berkeley, who has investigated the event. Those visas were rescinded, but it's a good example of how hidden algorithms can have a life-changing effect.

In a similar example, the documents leaked by Edward Snowden revealed that the National Security Agency uses algorithms to decide whether a person is a US citizen. According to US law, only non-citizens can have their communications monitored without a warrant. In the absence of information about an individual's birthplace or parents' citizenship, the NSA algorithms use other criteria. Is this person in contact with foreigners? Do they appear to have accessed the internet from a foreign country? Depending on what you do online, your citizenship might change overnight. "One day you might be a citizen, another you might be a foreigner," says John Cheney-Lippold, at the University of Michigan in Ann Arbor. "It's a categorical assessment based on an interpretation of your data, not your passport or your birth certificate."

Algorithms are also used to police voter fraud. Several US states use software called Crosscheck to remove duplicate entries from electoral registers. But people have been deleted simply for having the same name. As with the sub-prime algorithms, minorities are again hit hardest. The names it scrubs are disproportionately those of black, Asian and Hispanic voters, who are more likely to share names – such as Jackson, Kim or Garcia.

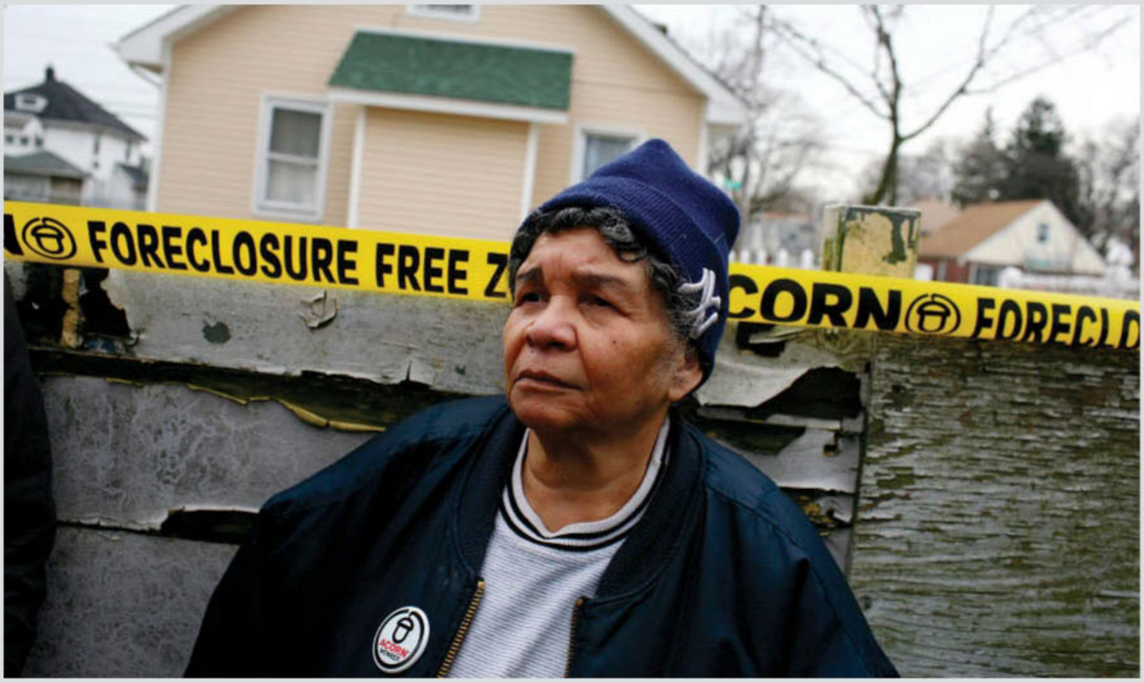
The next scandal may be prison sentencing. Judges and lawyers in Missouri can use a website to make an "Automated Sentencing Application". The system calculates incarceration costs for defendants, and weighs that against the likelihood the defendant will reoffend, based on prior criminal history and behavioural and demographic factors. Some think this will lead to minorities being given harsher sentences. Proxies like address, income and education level make it almost impossible to avoid racial bias. Similar systems are appearing across the US. "I think it's terrifying," says Sorelle Friedler,



ARTUR DEBAT/GETTY

Are systems that pick what you see keeping you in the dark?

Automated decisions may have culminated in sub-prime misery



ANDREW LICHTENSTEIN/ POLARIS / EYEVINE

a computer scientist at Haverford College in Pennsylvania.

The scales are falling from our eyes as the impact of algorithms is felt in almost every area of our lives. What should we do about it? In many of these examples, the problem is not the algorithms themselves, but the fact that they over-amplify an existing bias in the data.

HIGHER STANDARDS

"People who work with algorithms are comfortable with the idea that they might produce these unintended results," says Sandvig. But for a growing number of people, that's not good enough. Christo Wilson at Northeastern University, who has worked with Hannak, thinks that large technology companies like Google and Facebook ought to be considered as public services that huge numbers of people rely on. "Given that they have a billion eyeballs, I think they have a responsibility to hold themselves to a higher standard," he says.

Wilson thinks that automated systems might be made more trustworthy if users can control exactly how their results are personalised – such as leaving gender out of the equation or ignoring income bracket and address. It would also help us learn how these systems work, he says.

Others are calling for a new regulatory framework governing algorithms, much like we have for the financial industry, for example. A recent report commissioned by the White House recommends that policy-makers pay more attention to what the algorithms do with the data they collect and analyse. To ensure accountability, however, there would need to be independent auditors who inspect algorithms and monitor their impact. We cannot leave it to governments or industry alone to respond to the problems, says Gangadharan.

"The big question now for me is who are the watchdogs," says Sandvig. For now, he suggests it should be the researchers who are beginning to reveal algorithms' broader effects. Wilson, for example, has looked into setting up dummy credit profiles to better understand price-fixing systems. But independent auditors face tough obstacles. For a start, digging around inside many automated services violates their terms of use agreement, which prohibits attempts to analyse how they work. Under the US Computer Fraud and Abuse Act, such snooping may even be illegal. And while public scrutiny is important, the details of proprietary algorithms need to be kept safe from competitors or hackers, for example.

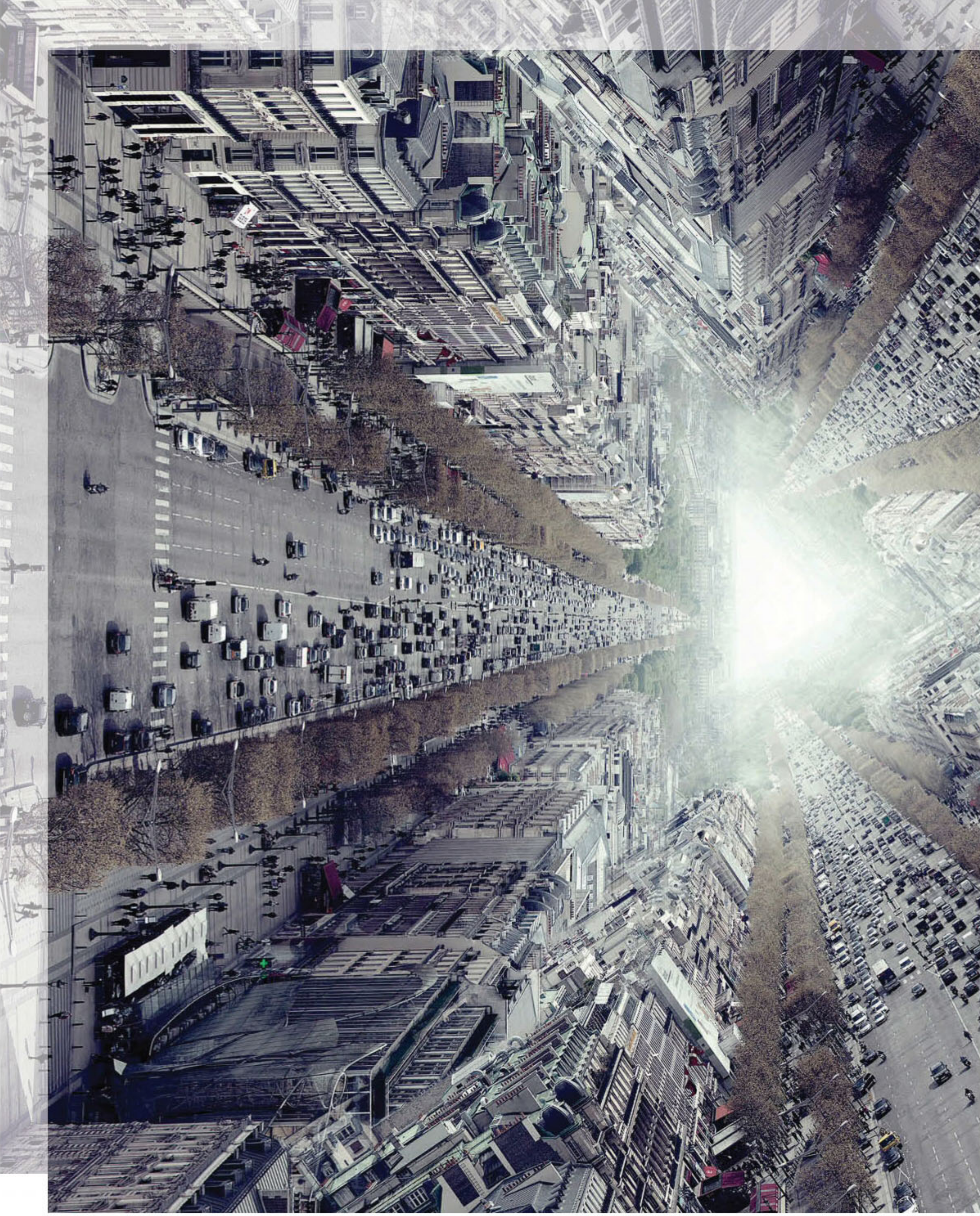
"HOW WE WANT ALGORITHMS TO BEHAVE MAY BE A HARDER QUESTION TO ANSWER THAN WE THINK"

What's more, most automated systems are too complex for humans to inspect by hand. So some researchers have developed algorithms that check other algorithms. Kroll is working on a system that would let an auditor verify that an algorithm did what it was supposed to with what it was given. In other words, it would provide a foolproof way of checking that the outcome of the green card lottery, for example, was in fact random. Or that a driverless car's algorithm for avoiding pedestrians treats both people walking and people in wheelchairs with the same caution.

Friedler has a different approach. By understanding the biases inherent in the underlying data, she hopes to eliminate bias in the algorithm. Her system looks for correlations between arbitrary properties – like height or address – and demographic groupings like race or gender. If the correlation is expected to lead to unwanted bias, then it would make sense to normalise the data. It is essentially affirmative action for algorithms, she says.

That's fine for cases where discrimination is clear, where a system is found to be unfair or illegal. But what if there is disagreement about how an algorithm ought to behave? Many would say Facebook's filtering of its newsfeed keeps it readable. Some would argue that highly personalised price adjustment can benefit both customers and retailers. What's acceptable to some won't be for others.

As Sandvig notes, unlike for financial systems, there are no standards of practice governing algorithms. But how we want them to behave may turn out to be a harder question for society to answer than we think. Maybe we'll need an algorithm for that. ■



SIMON GARDINER

The world maker

Is time running out for the clever piece of maths that runs modern life, asks Richard Elwes

YOU MIGHT not have heard of the algorithm that runs the world. Few people have, though it can determine much that goes on in our day-to-day lives: the food we have to eat, our schedule at work, when the train will come to take us there. Somewhere, in some server basement right now, it is probably working on some aspect of your life tomorrow, next week, in a year's time.

Perhaps ignorance of the algorithm's workings is bliss. The door to Plato's Academy in ancient Athens is said to have borne the legend "let no one ignorant of geometry enter". That was easy enough to say back then, when geometry was firmly grounded in the three dimensions of space our brains were built to cope with. But the algorithm operates in altogether higher planes. Four, five, thousands or even many millions of dimensions: these are the unimaginable spaces the algorithm's series of mathematical instructions was devised to probe.

Perhaps, though, we should try a little harder to get our heads round it. Because powerful though it undoubtedly is, the algorithm is running into a spot of bother. Its mathematical underpinnings, though not yet structurally unsound, are beginning to crumble at the edges. With so much resting on it, the algorithm may not be quite as dependable as it once seemed.

To understand what all this is about, we must first delve into the deep and surprising ways in which the abstractions of geometry describe the world around us. Ideas about such connections stretch at least as far back as Plato, who picked out five 3D geometric shapes, or polyhedra, whose perfect regularity he thought represented the essence of the cosmos. The tetrahedron, cube, octahedron and 20-sided icosahedron embodied the "elements" of fire, earth, air and water, and the 12-faced dodecahedron the shape of the universe itself.

Things have moved on a little since then. Theories of physics today regularly invoke strangely warped geometries unknown to Plato, or propose the existence of spatial dimensions beyond the immediately obvious three. Mathematicians, too, have reached for ever higher dimensions, extending ideas about polyhedra to mind-blowing "polytopes" with four, five or any number of dimensions.

A case in point is a law of polyhedra proposed in 1957 by the US mathematician Warren Hirsch. It stated that the maximum number of edges you have to traverse to get between two corners on any polyhedron is never greater than the number of its faces minus the number of dimensions in the problem, in this case three. The two opposite corners on a six-sided cube, for example, are separated by exactly three edges, and no pair of corners is four or more apart.

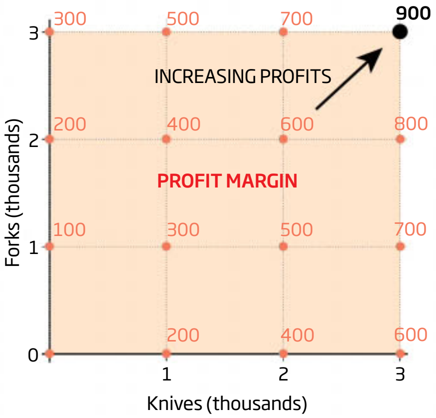
Hirsch's rule holds true for all 3D polyhedra. But it has never been proved generally for shapes in higher dimensions. The expectation that it should translate has come largely through analogy with other geometrical rules that have proved similarly elastic (see "Edges, corners and faces", page 67). When it comes to guaranteeing short routes between points on the surface of a 4D, 5D or 1000D shape, Hirsch's rule has remained one of those niggling unsolved problems of mathematics – a mere conjecture.

How is this relevant? Because, for today's mathematicians, dimensions are not just ›

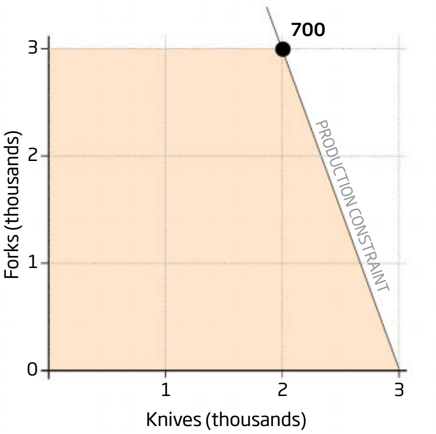
Room for improvement

Many business problems can be reduced to patterns in geometry – as this simple 2D example shows

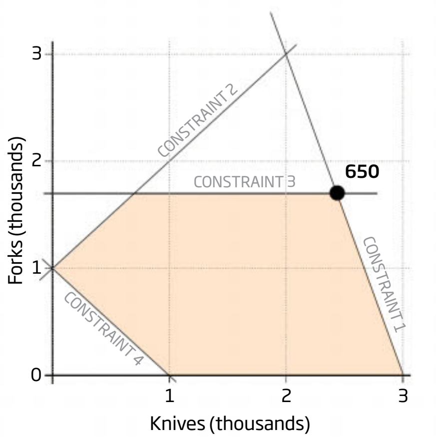
A cutlery factory makes $200 profit for every 1000 knives and $100 for every 1000 forks. With no constraints on production, **more profit** is made by making more of both



In the real world, finite staff or machine resources will mean the more forks you make, the fewer knives you can make. That constrains your operating space and the **maximum profit** you can make



Further constraints, such as demand for cutlery, restrict your operating space to a 2D geometric shape – and the **maximum achievable profit** always lies at a corner of that shape



about space. True, the concept arose because we have three coordinates of location that can vary independently: up-down, left-right and forwards-backwards. Throw in time, and you have a fourth "dimension" that works very similarly, apart from the inexplicable fact that we can move through it in only one direction.

But beyond motion, we often encounter real-world situations where we can vary many more than four things independently. Suppose, for instance, you are making a sandwich for lunch. Your fridge contains 10 ingredients that can be used in varying quantities: cheese, chutney, tuna, tomatoes, eggs, butter, mustard, mayonnaise, lettuce, hummus. These ingredients are nothing other than the dimensions of a sandwich-making problem. This can be treated geometrically: combine your choice of ingredients in any particular way, and your completed snack is represented by a single point in a 10-dimensional space.

Brutish problems

In this multidimensional space, we are unlikely to have unlimited freedom of movement. There might be only two mouldering hunks of cheese in the fridge, for instance, or the merest of scrapings at the bottom of the mayonnaise jar. Our personal preferences might supply other, more subtle constraints to our sandwich-making problem: an eye on the calories, perhaps, or a desire not to mix tuna and hummus. Each of these constraints represents a boundary to our multidimensional space beyond which we cannot move. Our resources and preferences in effect construct a multidimensional polytope through which we must navigate towards our perfect sandwich.

In reality, the decision-making processes in our sandwich-making are liable to be a little haphazard; with just a few variables to consider, and mere gastric satisfaction riding on the outcome, that's not such a problem. But in business, government and science, similar optimisation problems crop up everywhere and quickly morph into brutes with many thousands or even millions of variables and constraints. A fruit importer might have a 1000-dimensional problem to deal with, for instance, shipping bananas from five distribution centres storing varying numbers of fruit to 200 shops with different numbers in demand. How many items of fruit should be sent from which centres to which shops while minimising total transport costs?

A fund manager might similarly want to

PLAINPICTURE/GOZOOMA

arrange a portfolio optimally to balance risk and expected return over a range of stocks; a railway timetabler to decide how best to roster staff and trains; or a factory or hospital manager to work out how to juggle finite machine resources or ward space. Each such problem can be depicted as a geometrical shape whose number of dimensions is the number of variables in the problem, and whose boundaries are delineated by whatever constraints there are (see diagram, left). In each case, we need to box our way through this polytope towards its optimal point.

This is the job of the algorithm.

Its full name is the simplex algorithm, and it emerged in the late 1940s from the work of the US mathematician George Dantzig, who had spent the second world war investigating ways to increase the logistical efficiency of the US air force. Dantzig was a pioneer in the field of what he called linear programming, which uses the mathematics of multidimensional polytopes to solve optimisation problems.

The simplex algorithm directs wares to their destinations the world over

One of the first insights he arrived at was that the optimum value of the "target function" – the thing we want to maximise or minimise, be that profit, travelling time or whatever – is guaranteed to lie at one of the corners of the polytope. This instantly makes things much more tractable: there are infinitely many points within any polytope, but only ever a finite number of corners.

If we have just a few dimensions and constraints to play with, this fact is all we need. We can feel our way along the edges of the polytope, testing the value of the target function at every corner until we find its sweet spot. But things rapidly escalate. Even just a 10-dimensional problem with 50 constraints – perhaps trying to assign a schedule of work to 10 people with different expertise and time constraints – may already land us with several billion corners to try out.

The simplex algorithm finds a quicker way through. Rather than randomly wandering along a polytope's edges, it implements a "pivot rule" at each corner. Subtly different variations of this pivot rule exist in different implementations of the algorithm, but often it involves picking the edge along which the target function descends most steeply, thus ensuring each step takes us nearer the optimal value. When a corner is found where no further descent is possible, we know we have arrived at the optimal point.

Practical experience shows that the simplex method is generally a very slick problem-solver indeed, typically reaching an optimum solution after a number of pivots comparable to the number of dimensions in the problem. That means a likely maximum of a few hundred steps to solve a 50-dimensional problem, rather than billions with a suck-it-and-see approach. Such a running time is said to be "polynomial" or simply "P", the benchmark for practical algorithms that have to run on finite processors in the real world.

"Probably tens or hundreds of thousands of calls are made of the simplex algorithm every minute"

Dantzig's algorithm saw its first commercial application in 1952, when Abraham Charnes and William Cooper at what is now Carnegie Mellon University in Pittsburgh, Pennsylvania, teamed up with Robert Mellon at the Gulf Oil Company to discover how best to blend available stocks of four different petroleum products into an aviation fuel with an optimal octane level.

Since then the simplex algorithm has steadily conquered the world, embedded both in commercial optimisation packages and bespoke software products. Wherever anyone is trying to solve a large-scale optimisation problem, the chances are that some computer chip is humming away to its tune. "Probably tens or hundreds of thousands of calls of the simplex method are made every minute," says Jacek Gondzio, an optimisation specialist at the University of Edinburgh, UK.

Yet even as its popularity grew in the 1950s and 1960s, the algorithm's underpinnings were beginning to show signs of strain. To start with, its running time is polynomial only on average. In 1972, US mathematicians Victor Klee and George Minty reinforced this point by running the algorithm around some ingeniously deformed multidimensional "hypercubes". Just as a square has four corners, and a cube eight, a hypercube in n dimensions has 2^n corners. The wonky way Klee and Minty put their hypercubes together meant that the simplex algorithm had to run through all of these corners before landing on the optimal one. In just 41 dimensions, that leaves the algorithm with over a trillion edges to traverse.

A similar story holds for every variation of the algorithm's pivot rule tried since Dantzig's original design: however well it does in

EDGES, CORNERS AND FACES

Since Plato laid down his stylus, a lot of work has gone into understanding the properties of 3D shapes, or polyhedra. Perhaps the most celebrated result came from the 18th-century mathematician Leonhard Euler. He noted that every polyhedron has a number of edges that is two fewer than the total of its faces and corners. The cube, for example, has six faces and eight corners, a total of 14, while its edges number 12. The truncated icosahedron, meanwhile, is the familiar pattern of a standard soccer ball. It has 32 faces (12 pentagonal and 20 hexagonal), 60 corners – and 90 edges.

The French mathematician Adrien-Marie Legendre proved this rule in 1794 for every 3D shape that contains no holes and does not cut through itself in any strange way.

As geometry started to grow more sophisticated and extend into higher dimensions in the 19th century, it became clear that Euler's relationship didn't stop there: a simple extension to the rule applies to shapes, or polytopes, in any number of dimensions. For a 4D "hypercube", for example, a variant of the formula guarantees that the total number of corners (16) and faces (24) will be equal to number of edges (32) added to the number of 3D "facets" the shape possesses (8).

The rule derived by Warren Hirsch in 1957 about the maximum distance between two corners of a polyhedron was thought to be similarly cast-iron. Whether it truly is turns out to have surprising relevance to the smooth workings of the modern world (see main story).

2000 YEARS OF ALGORITHMS

George Dantzig's simplex algorithm has a claim to be the world's most significant (see main story). But algorithms go back far further.

c. 300 BC THE EUCLIDEAN ALGORITHM

From Euclid's mathematical primer *Elements*, this is the grandaddy of all algorithms, showing how, given two numbers, you can find the largest number that divides into both. It has still not been bettered.

820 THE QUADRATIC ALGORITHM

The word algorithm is derived from the name of the Persian mathematician Al-Khwarizmi. Experienced practitioners today perform his algorithm for solving quadratic equations (those containing an x^2 term) in their heads. For everyone else, modern algebra provides the formula familiar from school.

1936 THE UNIVERSAL TURING MACHINE

The British mathematician Alan Turing equated algorithms with mechanical processes – and found one to mimic all the others, the theoretical template for the programmable computer.

1946 THE MONTE CARLO METHOD

When your problem is just too hard to solve directly, enter the casino of chance. John von Neumann, Stanislaw Ulam and Nicholas Metropolis's Monte Carlo algorithm taught us how to play and win.

1957 THE FORTRAN COMPILER

Programming was a fiddly, laborious job until an IBM team led by John Backus invented the first high-level programming language, Fortran. At the centre is the compiler: the algorithm which converts the programmer's instructions into machine code.

1962 QUICKSORT

Extracting a word from the right place in a dictionary is an easy task; putting all the words in the right order in the first place is not. The British mathematician Tony Hoare provided the recipe, now an essential tool in managing databases of all kinds.

1965 THE FAST FOURIER TRANSFORM

Much digital technology depends on breaking down irregular signals into their pure sine-wave components – making James Cooley and John Tukey's algorithm one of the world's most widely used.

1994 SHOR'S ALGORITHM

Bell Labs's Peter Shor found a new, fast algorithm for splitting a whole number into its constituent primes – but it could only be performed by a quantum computer. If ever implemented on a large scale, it would nullify almost all modern internet security.

1998 PAGERANK

The internet's vast repository of information would be of little use without a way to search it. Stanford University's Sergey Brin and Larry Page found a way to assign a rank to every web page – and the founders of Google have been living off it ever since.

general, it always seems possible to concoct some awkward optimisation problems in which it performs poorly. The good news is that these pathological cases tend not to show up in practical applications – though exactly why this should be so remains unclear. "This behaviour eludes any rigorous mathematical explanation, but it certainly pleases practitioners," says Gondzio.

Flashy pretenders

The fault was still enough to spur on researchers to find an alternative to the simplex method. The principal pretender came along in the 1970s and 1980s with the discovery of "interior point methods", flashy algorithms which rather than feeling their way around a polytope's surface drill a path through its core. They came with a genuine mathematical seal of approval – a guarantee always to run in polynomial time – and typically took fewer steps to reach the optimum point than the simplex method, rarely needing over 100 moves regardless of how many dimensions the problem had.

The discovery generated a lot of excitement, and for a while it seemed that the demise of Dantzig's algorithm was on the cards. Yet it survived and even prospered. The trouble with interior point methods is that each step entails far more computation than a simplex pivot: instead of comparing a target function along a small number of edges, you must analyse all the possible directions within the polytope's interior, a gigantic undertaking. For some huge industrial problems, this trade-off is worth it, but for by no means all. Gondzio estimates that between 80 and 90 per cent of today's linear optimisation problems are still solved by some variant of the simplex algorithm. The same goes for a good few of the even more complex non-linear problems (see "Straight down the line", right). "As a devoted interior-point researcher I have a huge respect for the simplex method," says Gondzio. "I'm doing my best trying to compete."

We would still dearly love to find something better: some new variant of the simplex algorithm that preserves all its advantages, but also invariably runs in polynomial time. For US mathematician and Fields medallist Steve Smale, writing in 1998, discovering such a "strongly polynomial" algorithm was one of 18 outstanding mathematical questions to be dealt with in the 21st century.

Yet finding such an algorithm may not now even be possible.

That is because the existence of such an

SIMON GARDINER

"Cases where the algorithm fails have tended not to show up in practice – a pleasing behaviour that eludes explanation"

improved, infallible algorithm depends on a more fundamental geometrical assumption – that a short enough path around the surface of a polytope between two corners actually exists. Yes, you've got it: the Hirsch conjecture.

The fates of the conjecture and the algorithm have always been intertwined. Hirsch was himself a pioneer in operational research and an early collaborator of Dantzig's, and it was in a letter to Dantzig in 1957 musing

STRAIGHT DOWN THE LINE

When a young and nervous George Dantzig spoke about his new simplex algorithm at a conference of eminent economists and statisticians in Wisconsin in 1948, a rather large hand was raised in objection at the back of the room. It was that of the renowned mathematician Harold Hotelling. "But we all know the world is non-linear," he said.

It was a devastating put-down. The simplex algorithm's success in solving optimisation problems (see main story) depends on assuming that variables vary in response to other variables along nice straight lines. A cutlery company increasing its expenditure on metal, for example, will produce proportionately more finished knives, forks and profit the next month.

In fact, as Hotelling pointed out, the real world is jam-packed with non-linearity. As the cutlery company expands, economies of scale may mean the marginal cost of each knife or fork drops, making for a non-linear profit boost. In geometrical terms, such problems are represented by multidimensional shapes just as linear problems are, but ones bounded by curved faces that the simplex algorithm should have difficulty crawling round.

Surprisingly, though, linear approximations to non-linear processes turn out to be good enough for most practical purposes. "I would guess that 90 or 95 per cent of all optimisation problems solved in the world are linear programs," says Jacek Gondzio of the University of Edinburgh, UK. For those few remaining problems that do not submit to linear wiles, there is a related field of non-linear programming – and here too, specially adapted versions of the simplex algorithm have come to play an important part.

about the efficiency of the algorithm that Hirsch first formulated his conjecture.

Until recently, little had happened to cast doubt on it. Klee proved it true for all 3D polyhedra in 1966, but had a hunch the same did not hold for higher-dimensional polytopes. In his later years, he made a habit of suggesting it as a problem to every freshly scrubbed researcher he ran across. In 2001 one of them, a young Spaniard called Francisco Santos, now at the University of Cantabria in Santander, took on the challenge.

As is the way of such puzzles, it took time. After almost a decade working on the problem, Santos was ready to announce his findings at a conference in Seattle in 2010. Two years later, the resulting paper was published in the *Annals of Mathematics*. In it, Santos describes a 43-dimensional polytope with 86 faces. According to Hirsch's conjecture, the longest path across this shape would have (86 - 43) steps, that is, 43 steps. But Santos was able to establish conclusively that it contains a pair of corners at least 44 steps apart.

If only for a single special case, Hirsch's conjecture had been proved false. "It settled a problem that we did not know how to approach for many decades," says Gil Kalai of the Hebrew University of Jerusalem. "The entire proof is deep, complicated and very elegant. It is a great result."

A great result, true, but decidedly bad news for the simplex algorithm. Since Santos's first disproof, further Hirsch-defying polytopes have been found in dimensions as low as 20. The only known limit on the shortest distance between two points on a polytope's surface is now contained in a mathematical expression derived by Kalai and Daniel Kleitman of the Massachusetts Institute of Technology in 1992. This bound is much larger than the one the Hirsch conjecture would have provided, had it proved to be true. It is far too big, in fact, to guarantee a reasonable running time for the simplex method, whatever fancy new pivot rule we might dream up. If this is the best we can do, it may be that Smale's goal of an idealised algorithm will remain forever out of reach – with potentially serious consequences for the future of optimisation.

All is not lost, however. A highly efficient variant of the simplex algorithm may still be possible if the so-called polynomial Hirsch conjecture is true. This would considerably tighten Kalai and Kleitman's bound, guaranteeing that no polytopes have paths disproportionately long compared with their dimension and number of faces. A topic of interest before the plain-vanilla Hirsch conjecture melted away, the polynomial version has been attracting intense attention since Santos's announcement, both as a deep geometrical conundrum and a promising place to sniff around for an optimally efficient optimisation procedure.

As yet, there is no conclusive sign that the polynomial conjecture can be proved either. "I am not confident at all," says Kalai. Not that this puts him off. "What is exciting about this problem is that we do not know the answer."

A lot could be riding on that answer. As the algorithm continues to hum away in those basements it is still, for the most part, telling us what we want to know in the time we want to know it. But its own fate is now more than ever in the hands of the mathematicians. ■

INTELLIGENCE REINVENTED

The arrival of artificial intelligence in our lives is a story of hype, disappointment and unexpected triumph, explains ***Nello Cristianini***

IN THE summer of 1956, a remarkable collection of scientists and engineers gathered at Dartmouth College in Hanover, New Hampshire. Among them were computer scientist Marvin Minsky, information theorist Claude Shannon and two future Nobel prizewinners, Herbert Simon and John Nash. Their task: to spend the summer months inventing a new field of science called "artificial intelligence" (AI).

They did not lack in ambition, writing in their funding application: "every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it." Their wish list was "to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves". They thought that "a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer."

It took rather longer than a summer, but 60 years later, the field of AI seems to have finally found its way. These days, we can ask a computer questions, sit back while semi-autonomous cars negotiate traffic, and use smartphones to translate speech or printed text across most languages. We trust computers to check passports, screen our correspondence and fix our spelling. Even more remarkably, we have become so used to these tools working that we complain when they fail.

LEANDRO CASTELAO

As we rapidly get used to this convenience, it is easy to forget that AI hasn't always been this way.

At the Dartmouth conference, and at various meetings that followed it, the defining goals for the field were already clear: machine translation, computer vision, text understanding, speech recognition, control of robots and machine learning. For the following three decades, significant resources were ploughed into research, but none of the goals were achieved. It was not until the late 1990s that many of the advances predicted in 1956 started to happen. But before this wave of success, the field had to learn an important and humbling lesson.

While its goals have remained essentially the same, the methods of creating AI have changed dramatically. The instinct of those early engineers was to program machines from the top down. They expected to generate intelligent behaviour by first creating a mathematical model of how we might process speech, text or images, and then by implementing that model in the form of a computer program, perhaps one that would reason logically about those tasks. They were proven wrong.

They also expected that any breakthrough in AI would provide us with further understanding about our own intelligence. Wrong again.

Over the years, it became increasingly clear that those systems weren't suited to dealing with the messiness of the real world. By the early 1990s, with little to show for decades of work, most engineers started abandoning the dream of a general-purpose top-down reasoning machine. They started looking at humbler projects, focusing on specific tasks that were more likely to be solved.

Some early success came in systems to recommend products. While it can be difficult to know why a customer might want to buy an item, it can be easy to know which item they might like on the basis of previous transactions by themselves or similar customers. If you liked the first and second Harry Potter films, you might like the third. A full understanding of the problem was not required: you could detect useful correlations just by combing through a lot of data.

DATA BEATS THEORY

Could similar bottom-up shortcuts emulate other forms of intelligent behaviour? After all, there were many other problems in AI where no theory existed, but there was plenty of data to analyse. This pragmatic attitude produced success in speech recognition, machine translation and simple computer vision tasks such as recognising handwritten digits.

By the mid-2000s, with success stories piling up, the field had learned a powerful lesson: data can be stronger than theoretical models. A new generation of intelligent machines had emerged, powered by a small set of statistical learning algorithms and large amounts of data.

Researchers also ditched the assumption that AI would help us better understand our own intelligence. Try to learn from algorithms how humans perform those tasks, and you are wasting your time: the intelligence is more in the data than in the algorithm.

THE WINTERS OF AI DISCONTENT

Emergent technologies are often subjected to hype cycles, sometimes due to speculative bubbles inflated by excessive investor expectations. Some examples are railway mania in the UK in the 1840s and the dot-com bubble in the 1990s.

Artificial intelligence is perhaps unique in having undergone several hype cycles in a relatively short time. Its slumps of optimism even have a specific name: AI winters (see timeline, right). The two major winters occurred in the early 1970s and late 1980s. Both were caused largely by the withdrawal of public funding as progress stalled.

AI is now in a renewed phase of heightened optimism and investment. Unlike in previous cycles, however, AI today has a strong – and increasingly diversified – commercial revenue stream. Only time will tell whether this turns out to be a bubble.

The field had undergone a paradigm shift and had entered the age of data-driven AI. Its new core technology was machine learning, and its language was no longer that of logic, but statistics.

How, then, can a machine learn? It is worth clarifying here what we normally mean by learning in AI: a machine learns when it changes its behaviour (hopefully for the better) based on experience. It sounds almost magical, but in reality the process is quite mechanical.

Consider how the spam filter in your mailbox decides to quarantine some emails on the basis of their content. Every time you drag an email into the spam folder, you enable it to estimate the probability that messages from a given recipient or containing a given word are unwanted. Combining this information for all the words in a message allows it to make an educated guess about new emails. No deep understanding is required.

But when these ideas are applied on a very large scale, something surprising seems to happen: machines start doing things that would be difficult to program directly, like being able to complete sentences, predict our next click, or recommend a product. Taken to its extreme conclusion, this approach has delivered language translation, handwriting recognition, face recognition and more. Contrary to the assumptions of 60 years ago, we don't need to precisely describe a feature of intelligence for a machine to simulate it.

While each of these mechanisms is simple enough ›

1950 *Alan Turing publishes the seminal paper "Computing machinery and intelligence". Its opening sentence is "I propose to consider the question, 'Can machines think?'"*

1956 *The term "artificial intelligence" is coined at a workshop at Dartmouth College*

1959 *Computer scientists at Carnegie Mellon University create the General Problem Solver (GPS), a program that can solve logic puzzles*

1973 *The first AI winter sets in as funding and interest dry up*

1975 *A system called MYCIN diagnoses bacterial infections and recommends antibiotics using deduction based on a series of yes/no questions. It was never used in practice*

1987 *Second AI winter begins*

1989 *NASA's AutoClass computer program discovers several previously unknown classes of stars*

1994 *First web search engines launched*

1997 *IBM's Deep Blue beats world champion Garry Kasparov at chess*

1998 *NASA's Remote Agent is first fully autonomous program to control a spacecraft in flight*

2002 *Amazon replaces human product recommendation editors with an automated system*

2007 *Google launches Translate, a statistical machine translation service*

2009 *Google researchers publish an influential paper called "The unreasonable effectiveness of data". It declares that "simple models and a lot of data trump more elaborate models based on less data"*

2011 *Apple releases Siri, a voice-operated personal assistant that can answer questions, make recommendations and carry out instructions such as "call home"*

2011 *IBM's supercomputer Watson beats two human champions at TV quiz game Jeopardy!*

2012 *Google's driverless cars navigate autonomously through traffic*

2016 *Google's AlphaGo defeats Lee Sedol, one of the world's leading Go players*

"Modern artificial intelligence is a brilliant and powerful technology, but also a fundamentally disruptive one"

that we might call it a statistical hack, when we deploy many of them simultaneously in complex software, and feed them with millions of examples, the result might look like highly adaptive behaviour that feels intelligent to us. Yet, remarkably, the agent has no internal representation of why it does what it does.

This experimental finding is sometimes called "the unreasonable effectiveness of data". It has been a very humbling and important lesson for AI researchers: that simple statistical tricks, combined with vast amounts of data, have delivered the kind of behaviour that had eluded its best theoreticians for decades.

Thanks to machine learning and the availability of vast data sets, AI has finally been able to produce usable vision, speech, translation and question-answering systems. Integrated into larger systems, those can power products and services ranging from Siri and Amazon to the Google car.

Researchers' attention is now focused what fuels the engine of our intelligent machines: data. Where can they find data, and how can they make the most of this resource?

One important step has been to recognise that valuable data can be found freely "in the wild", generated as a byproduct of various activities – some as mundane as sharing a tweet or adding a smiley under a blog post.

Engineers and entrepreneurs have also invented a variety of ways to elicit and collect additional data, such as asking users to accept a cookie, tag friends in images, rate a product or play a location-based game centred on finding monsters in the street. Data became "the new oil".

CHALLENGES AHEAD

At the same time as AI was finding its way, we developed an unprecedented global data infrastructure. Every time you access the internet to read the news, do a search, buy something, play a game, or check your email, bank balance or social media feed, you interact with this infrastructure. It isn't just a physical one of computers and wires, but also one of software, including social networks and microblogging sites.

Data-driven AI both feeds on this infrastructure and powers it – it is hard to imagine one without the other. And it is hard to imagine life without either of them. This is what makes modern AI a brilliant and powerful technology, but also a fundamentally disruptive one.

The unified data infrastructure is not like any medium invented before. Unlike the copper cables that used to connect people in the telegraph or telephone age, it takes a keen interest in our actions. The medium looks back at us, anticipating our moves, guessing our intents, often trying to serve us better and sometimes to influence us. This gives a whole new meaning to the claim, made by the 1970s communications theorist Marshall McLuhan, that a medium can never be neutral.

The challenges AI might present us with include surveillance, discrimination, persuasion, unemployment and possibly even addiction. Are we prepared?

Intelligent machines need to collect data – often personal data – in order to work. This simple fact potentially turns them into surveillance devices: they know our location, our browsing history and our social networks. Can we decide who has access, what use can be made of the data, or whether the data gets deleted for ever? If the answer is no, then we don't have control.

AI's capability to make predictions is useful for insurance, loans and policing. But the quality of those predictions will depend on subtle design choices and on the way the information used to train it is collected, which creates a very real risk of implicit and unintended discrimination. A recent investigation by ProPublica, for example, claims to have uncovered a bias that would disadvantage African Americans in the software used in many US courts to make parole decisions. Another case has been reported where different job ads were targeted at different ethnic groups. Both illustrate the unintended effects of the complex interaction between algorithms and data.

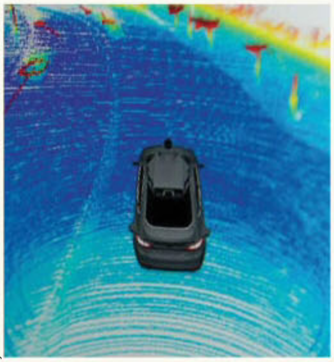
Another concern is persuasion. The business model of many AI companies is advertising, which means getting people to click on specific links. Research on how to steer users is well under way. The more the machines know about us, the better the job they can do of nudging us. Predictive interfaces might even induce addiction in vulnerable users, by actively rewarding them with the juiciest content that the web has to offer. This is something that needs to be carefully studied.

Employment will be affected too, as AIs learn from us (quite literally) how to do certain jobs, either because they watch how we do them, or because we are paid to generate their training data.

The emergence of internet crowdsourcing allows businesses to automatically outsource

RALPH FRESO/REUTERS

AI will increase the automation of warehouses like Amazon's



JEFF SWENSEN/THE NEW YORK TIMES/REDUX/EYEVINE

Driverless cars "see" better than ever, but mistakes can't be ruled out

micro-tasks that require human intelligence, by posting them on websites or apps where workers can choose the tasks they want to accept. In a way this works just like Uber, but for tasks other than driving, and is mediated by a computer system. Typical tasks would include transcribing handwriting or labelling images.

This also creates a workforce directly managed via computers, and defines a set of tasks that are the ideal candidate for automation. Indeed, many of those task-workers are actually generating or annotating the data being used to train their AI replacements. At the same time, many call centres and warehouses will be increasingly automated within a decade.

I do not believe that we yet have the legal and cultural tools to handle these and many other challenges. Who do we turn to if an intelligent algorithm denies us parole, medical treatment or a diploma? Are we prepared for our character and trustworthiness to be ranked just like our credit history, as some countries are proposing? Do we want the state to have access to our online activities and knowledge of our preferences? Do we want our children to spend their online time in the company of persuasive machines, designed to steer their behaviour in a given direction? What happens to society if large numbers of people are put out of work?

Artificial intelligence has come a long way. It is now being integrated into our lives, and promises to improve them. We might not call it AI once it is deployed, but we can expect benefits in fields ranging from healthcare to transportation, from communications to schooling.

And research is not slowing down. The machine-learning paradigm has been effective in addressing many areas like vision and speech processing, and it is likely that future AI will also find a way to integrate some top-down reasoning methods descended from earlier approaches. What will come after that may surprise us again.

As our AI efforts continue to open up new possibilities, we can imagine seamless conversations with machines, fluent real-time translation of speech, and many useful ways to automate our houses and cars.

But we might want to resist the temptation to introduce AI into as many domains as possible, at least before the cultural and legal framework evolves. Widespread adoption of AI brings remarkable opportunities, but also potential risks. Contrary to popular belief these are not existential risks to our species, but rather a possible erosion of our privacy and autonomy.

So as we finally enjoy the benefits of six decades of research in AI, with machines joining us in our everyday lives, we should celebrate – but also tread carefully. ■

YOU WIN SOME...

BEN HIDER/GETTY

One of the most celebrated successes of machine learning (see main story) came in 2016 when an algorithm called AlphaGo defeated South Korean master Lee Sedol at the game Go – something none of its programmers could come close to doing themselves. AlphaGo combined various machine-learning methodologies to analyse databases of more than 30 million Go moves, as well as playing thousands of games against itself. A similar strategy earlier allowed IBM's Watson supercomputer to win at the TV quiz game Jeopardy! (pictured above).

Given the right data, it seems that machines can improve their intelligence a great deal. But we should remember that machine learning is a statistical exercise, and therefore it can always fail.

We have also seen some blunders caused by machine learning. In 2015, Google apologised after one of its products automatically labelled photos of two black people "gorillas"; this year Microsoft had to withdraw a conversational bot called Tay because it had learned offensive language. In both cases it was not a failure of the algorithm, but of the training data that had been fed to it.

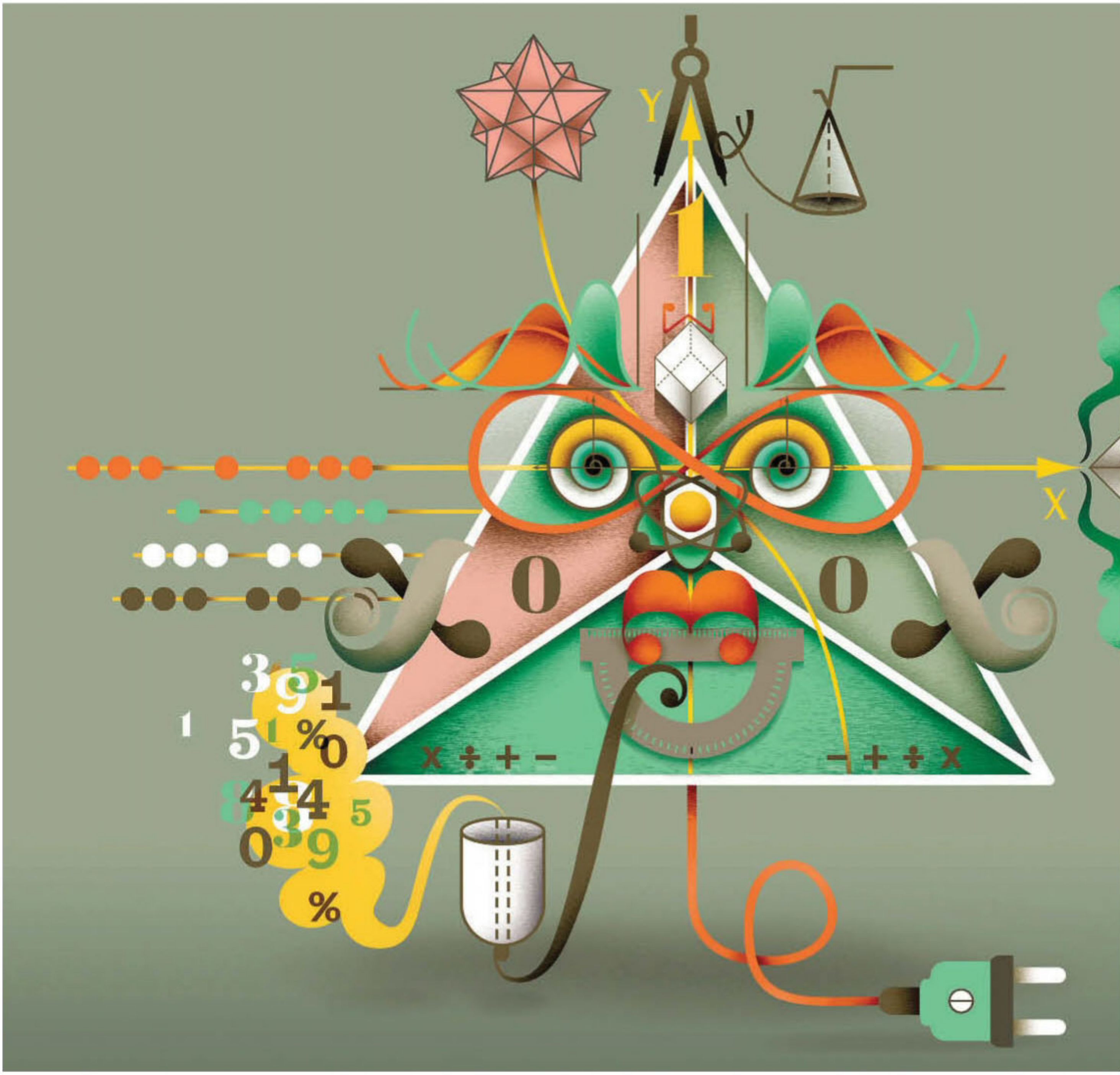
2016 also saw the first fatality linked to a "driverless" car, when a driver put a Tesla on autopilot and it failed to detect a trailer on the road. The conditions were unusual, with a white obstacle against a light sky, and the computer vision system simply made a mistake. I do not expect it to be the last one as many companies move into that market.

On the other hand, there are countless stories that do not end up in the news, simply because the AI systems are doing their work as expected. They include search engines, online shops and semi-autonomous cars.

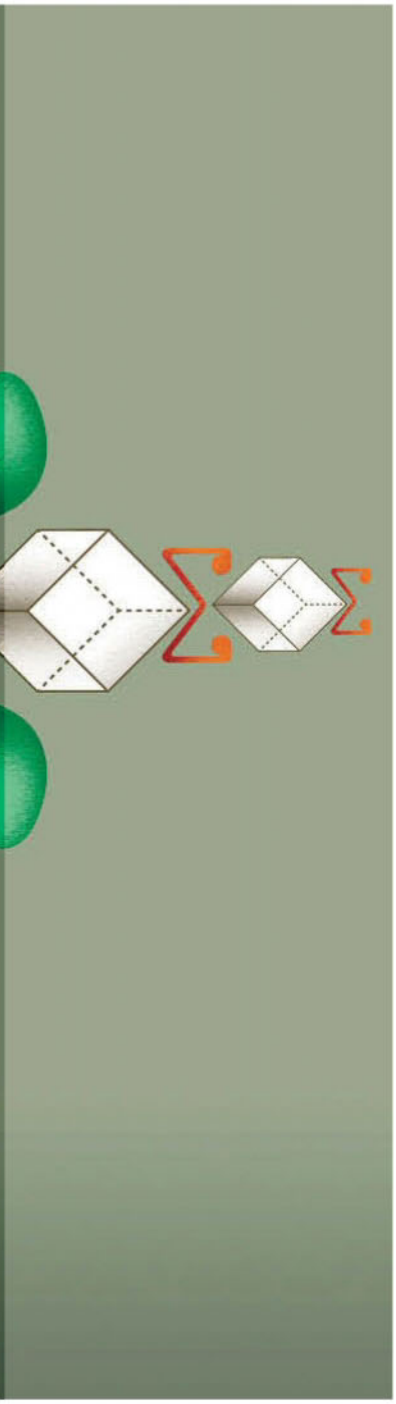
As we entrust machines with increasingly sensitive decisions, we need to pay careful attention to the kind of data we feed to them. It is not just the technology, but its deployment in our everyday life, that needs better understanding.

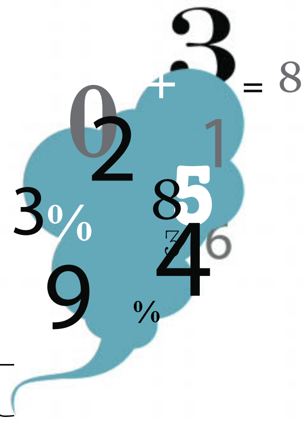
BEYOND KNOWLEDGE

We have always invented tools to take us beyond our natural abilities. But computers changed our ability to think, and they are now on the verge of transcending the limits of our ingenuity



ROYA HAMBURGER





Proof of concept

Computers can crack theorems humans struggle with. Could they extend mathematics itself, asks Jacob Aron

SHINICHI MOCHIZUKI is still waiting. In 2012, the highly respected mathematician at Kyoto University in Japan published more than 500 pages of dense maths on his website. It was the culmination of years of work. Mochizuki's inter-universal Teichmüller theory described previously uncharted areas of the mathematical realm and let him prove a long-standing conundrum about the true nature of numbers, known as the ABC conjecture. Other mathematicians hailed the result, but warned it would take a lot of effort to check. Months passed, then years, with no conclusion.

Ask a mathematician what a proof is and they're likely to tell you it must be absolute – an exhaustive sequence of logical steps leading from an established starting point to an undeniable conclusion. But that's not the whole story. You can't just publish something you believe is true and move on; you have to convince others that you haven't made any mistakes. For a truly groundbreaking proof, this can be a frustrating experience.

It turns out that very few mathematicians are willing to put aside their own work and dedicate the months or even years it would take to understand a proof like Mochizuki's. And as maths becomes increasingly fractured into subfields within subfields, the problem is set to get worse. Some think maths is reaching a limit. Real breakthroughs can be too complicated for others to check, so many mathematicians occupy themselves with more attainable but arguably less significant problems. What's to be done?

For some, the solution lies in employing digital help. A lot of mathematicians already work alongside computers – they can help check proofs and free up time for more creative work. But it might mean changing how maths is done. What's more, computers may one day make genuine breakthroughs on their own. Will we be able to keep up? And what does it mean for maths if we can't?

The first major computer-assisted proof was published 40 years ago and it immediately sparked a row. It was a solution to the four-colour theorem, a puzzle dating back to the mid-19th century. The theorem states that all maps need only four colours to make sure no adjacent regions are coloured the same. You can try it as many times as you like and find it to be true. But to prove it, you need to rule out the very possibility of there being a bizarre map that bucks the trend.

In 1976, Kenneth Appel and Wolfgang Haken did just that. They showed you could narrow the problem down to 1936 sub-arrangements that might require five colours. They then used a computer to check each of these potential counterexamples, and found that all could indeed be coloured with just four colours.

Job done, or so you'd think. "Mathematicians were reluctant to accept this as a proof," says Xavier Leroy at the Institute for Research in Computer Science and Automation (INRIA) in Paris, France. What if there was an error in the code? "They said: 'We're not going to recheck your thousand particular cases by hand, we don't trust your program, and that's not a real proof'."

They had a point. Checking software that tests a mathematical conjecture can be harder than proving it the traditional way, and a coding mistake can make the results totally unreliable. "It's very difficult to check whether a given program does the proper calculation just by inspection," says Georges Gonthier, also at INRIA. "The computer goes over the code many times, so it can amplify even the smallest error." ›

The trick is to use software to check software. Working with a type of program known as a proof assistant, mathematicians can verify that each step of a proof is valid. "It's a fairly interactive process, you type commands into the tool and then the tool will spellcheck it, if you like," says Leroy. And what if the proof assistant has a bug? It's always possible, but these programs tend to be small and relatively easy to check by hand. "More importantly, this is code that is run over and over again," says Gonthier. "You have massive experimental data to show that it is computing properly."

However, using proof assistants means embracing a different way of working. When mathematicians write out proofs, they skip a lot of the boring details. There is no point in laying out the foundations of calculus every time, for example. But such shortcuts don't fly with computers. To work with a proof, they must account for every logical step, even apparent no-brainers such as why 2 + 2 = 4.

"To make mathematical proof easier for computers, we must redefine maths itself"

Translating human-written proofs into computer-speak is still an active area of research. A single proof can take years. One early breakthrough came in 2005, when Gonthier and his colleagues updated the proof of the four-colour theorem, making every part of it computer-readable. Previous versions, ever since Appel and Haken's work in 1976, relied on an area of maths called graph theory, which draws on our spatial intuition. Thinking about regions on a map comes naturally to humans, but not computers. The whole thing needed reworking.

"You have to turn everything into algebra, and that forces you to be more precise," says Gonthier. "That precision ends up paying off." Gonthier discovered that a part of the proof – widely assumed to be true because it seemed so obvious – had in fact never been proved at all because it was deemed not worth the effort. The assumption turned out to be correct, but it illustrates an added benefit of extra precision.

Tackling the four-colour theorem was just a warm-up, however. "It has relatively few uses in the rest of mathematics," says Gonthier. "It was a brain-teaser." So he turned to the Feit-Thompson theorem, a large and foundational proof in group theory from the 1960s. For many years the proof had been built upon and rewritten and it was eventually published in two books. By formalising it, Gonthier hoped to demonstrate the computer's capacity to digest a meatier proof that touched many different branches of mathematics. "The perfect test case," he says.

It was a success. "In the process they found a couple of minor mistakes in the books," says Leroy. "They were easily fixable, but still things that every human mathematician missed." People took notice, says Gonthier. "I got letters saying how wonderful it was."

In both cases, the result was never in doubt. Gonthier was taking well-established maths and translating it for computers. But others have been forced to redo their work in this way just to get their proofs accepted.

In 1998, Thomas Hales at the University of Pittsburgh, Pennsylvania, found himself in a similar position to Mochizuki's today. He had just published a 300-page proof of the Kepler conjecture, a 400-year-old problem that concerns the most efficient ways to stack a collection of spheres. As with the four-colour theorem, the possibilities boiled down to variations on a few thousand arrangements. Hales and his then student Samuel Ferguson used a computer to check them all.

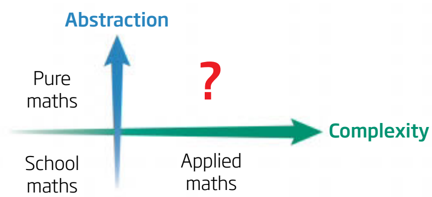
Hales submitted his result to the journal *Annals of Mathematics*. Five years later, reviewers for the journal announced they were 99 per cent certain that the proof was correct. "Referees in mathematics generally do not want to check computer code. They don't see that as part of their job," says Hales.

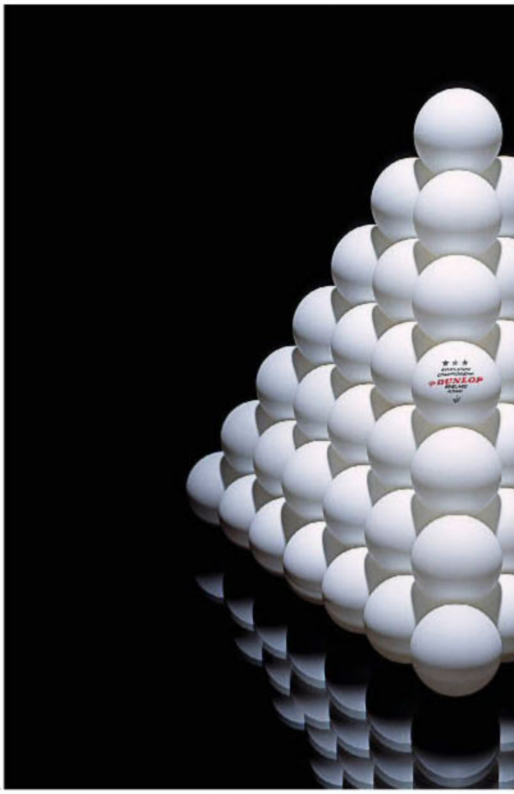
Convinced he was right, Hales started to rework his proof in 2003, so that it could be checked with a proof assistant. It essentially meant starting all over again, he says. He finally completed the project in 2014.

Gonthier's and Hales's research has shown that the approach can be applied to important mathematics. "The big theorems in maths that we're proving now seemed a distant dream 10 years ago," says Hales. But despite advances like the proof assistant, proving things with a computer is still a laborious process. Most mathematicians don't bother.

Uncharted terrain

Maths that is both highly abstract and highly complex may be beyond human ability. Some think computers could open up this new territory for us




JAMES DAY/GALLERY STOCK

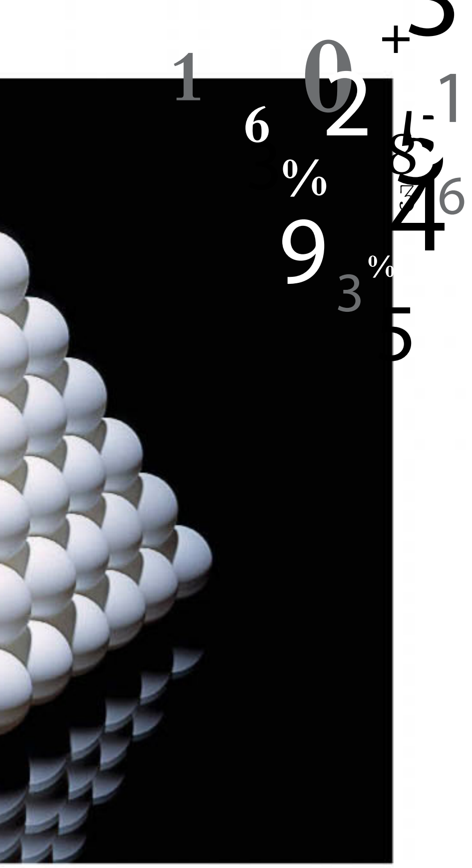
That's why some are working in the opposite direction. Rather than making proof assistants easier to use, Vladimir Voevodsky at the Institute for Advanced Study in Princeton, New Jersey, wants to make mathematics more amenable to computers. To do this, he is redefining its very foundations.

True to type

This is deep stuff. Maths is currently defined in terms of set theory, essentially the study of collections of objects. For example, the number zero is defined as the empty set, the collection of no objects. One is defined as the set containing one empty set. From there you can build an infinity of numbers. Most mathematicians don't worry about this on a day-to-day basis. "People are expected to understand each other without going down to that much detail," says Voevodsky.

Not so for computers, and that's a problem. There are multiple ways to define certain mathematical objects in terms of sets. For us, that doesn't matter, but if two computer proofs use different definitions for the same thing, they will be incompatible. "We cannot compare the results, because at the core they are based on two different things," says Voevodsky. "The existing foundations of maths don't work very well if you want to get everything in a very precise form."

Voevodsky's alternative approach swaps sets for types – a stricter way of defining mathematical objects in which every concept has exactly one definition. Proofs built with



Software confirmed the best way to stack spheres

types can also form types themselves, which isn't the case with sets. This lets mathematicians formulate their ideas with a proof assistant directly, rather than having to translate them later. In 2013 Voevodsky and colleagues published a book explaining the principles behind the new foundations. In a reversal of the norm, they wrote the book with a proof assistant and then "unformalised" it to produce something more human-friendly.

This backwards working changes the way mathematicians think, says Gonthier. "The book is entirely written in non-formalised prose, but if you have any kind of experience with using the computer system, you quickly realise that the prose closely reflects what is going on in the formal system."

It also allows much closer collaboration between large groups of mathematicians, because they don't have to constantly check each other's work. "They've really started to popularise the idea that proof assistants can be good for the working mathematician," says Leroy. "That's a really exciting development."

And it may be just the beginning. By making maths easer for computers to understand, Voevodsky's redefinition might take us into new territory. As he sees it, mathematics is split into four quadrants (see chart, left). Applied maths – modelling the airflow over a wing, for example – involves high complexity but low abstraction. Pure maths, the kind of pen and paper maths that is far removed from our everyday lives, involves low complexity but high abstraction. And school-level maths is neither complex nor abstract. But what lies in that fourth quadrant?

"It is very difficult at the present to go into the high levels of complexity and abstraction, because it just doesn't fit into our heads very well," says Voevodsky. "It somehow requires abilities that we don't posses." By working with computers, perhaps humans could access this fourth mathematical realm. We could prove bigger, bolder and more abstract problems than ever before, pushing our mastery of maths to ultimate heights.

Or perhaps we'll be left behind. In 2014, Alexei Lisitsa and Boris Konev at the University of Liverpool, UK, published a computer-assisted proof so long that it totalled 13 gigabytes, roughly the size of Wikipedia. Each line of the proof is readable, but for anyone to go through the entire result would take several tedious lifetimes.

The pair have since optimised their code and reduced the proof to 800 megabytes – a big improvement, but still impossible to digest. "From a human viewpoint, there's not much difference," says Lisitsa. Even if you did devote your life to reading something like this, it would be like studying a photograph pixel-by-pixel, never seeing the larger picture. "You cannot grasp the idea behind it."

Although it is on a far grander scale, the situation is similar to the original proof of the four-colour theorem, where mathematicians could not be sure an exhaustive computer search was correct. "We still don't know why the result holds true," says Lisitsa. "It could be a limit of human understanding, because the objects are so huge."

Doron Zeilberger of Rutgers University in Newark, New Jersey, thinks there will even come a time when human mathematicians will no longer be able to contribute. "For the next hundred years humans will still be needed as coaches to guide computers," he says. But after that? "They could still do it as an intellectual sport, and play each other like human chess players still do today, even though they are much inferior to machines."

Zeilberger is an extreme case. He has listed his computer, nicknamed Shalosh B. Ekhad, as a co-author for decades and thinks humans should put pen and paper aside to focus on educating our machines. "The most optimal use of a mathematician's time is knowledge transfer," he says. "Teach computers all their tricks and let computers take it from there."

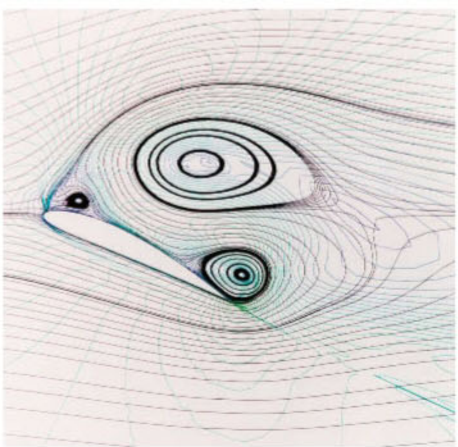
Spiritual discipline

But most mathematicians bristle at the idea of software that churns out proofs beyond human comprehension. "The idea that computers are going to replace mathematicians is misplaced," says Gonthier.

Besides, computer mathematicians would risk churning out an accelerating stream of unread papers. As it stands, scientific results often fail to garner the recognition they deserve, but the problem is particularly marked for maths. In 2014 there were more than 2000 maths papers posted to the online repository arXiv.org each month, more than in any other discipline, and the rate is increasing. "If you have too many new results that keep appearing, many just go unnoticed," says Leroy. Maybe we could at least create software to read everything and help humans keep up with the important bits, he says.

Gonthier feels this is missing the point: "Mathematics is not as much about finding proofs as it is about finding concepts." The nature of maths itself is under scrutiny. If humans do not understand a proof, then it doesn't count as maths, says Voevodsky. "The future of mathematics is more a spiritual discipline than an applied art. One of the important functions of mathematics is the development of the human mind."

All of this may be too late for Shinichi Mochizuki, however. A handful of mathematicians now understand his proof. Even so, his work is so advanced, so far removed from mainstream maths, that having a computer check it would be far more difficult than coming up with the original proof. "I don't even know if it would be possible to formalise what he's done," says Hales. For now, humans remain the ultimate judge – even if we don't always trust ourselves. ■



M.D. SANETRIK/CORBIS

Modelling and visualising airflow is a task computers handle well

Eureka machines

Inventors rely a lot on happenstance. Can computers help us make our own luck? Paul Marks investigates

ON A summer's day in 1899, a bicycle mechanic in Dayton, Ohio, slid a new inner tube out of its box and handed it to a customer. The pair chatted and the mechanic toyed idly with the empty box, twisting it back and forth. As he did so, he noticed the way the top of the box distorted in a smooth, spiral curve. It was a trivial observation – but one that would change the world.

The shape of the box just happened to remind the mechanic of a pigeon's wing in flight. Watching that box flex in his hands, Wilbur Wright saw how simply twisting the frame supporting a biplane's wings would give him a way to control an aircraft in the air.

Serendipity and invention go hand in hand. The Wright brothers' plane is just one of many examples. Take velcro: George de Mestral invented the material after he noticed the hook-covered seeds of the burdock plant sticking to his dog. And Harry Coover's liquid plastic concoction failed miserably as a material for cockpit canopies, as it stuck to everything. But it had a better use: superglue.

It may be romantic, but it is an achingly slow way to advance technology. Relying on happenstance means inventions that could be made today might not appear for years. "The way inventions are created is hugely archaic and inefficient," says Julian Nolan, CEO of Iprova, a company based in Lausanne, Switzerland, which specialises in generating inventions. Nothing has changed for hundreds of years, he says. "That's totally out of sync with most other industries."

But we are starting to make our own luck. Those eureka moments could soon be dialled up on demand as leaps of imagination are replaced by the steady steps of software. From algorithms that mimic nature's way of producing the best designs to systems that look for gaps between existing patented technologies that new designs might fill, computer-assisted invention is here.

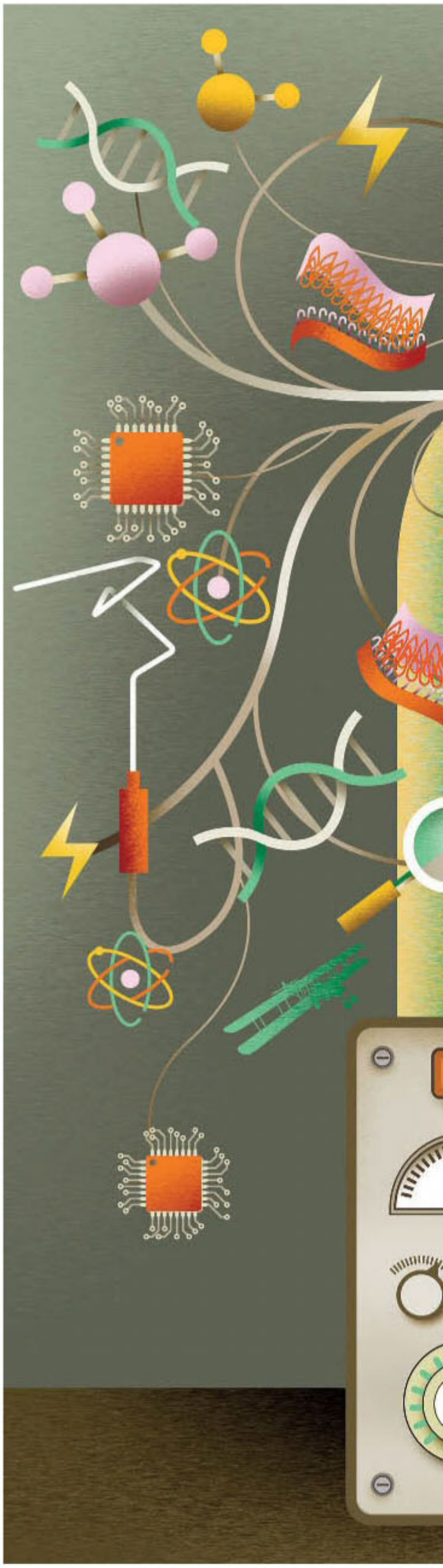
The impact could be huge. Some claim automated invention will speed up technological progress. It could also level the playing field, making inventors of us all. But what happens if the currency of ideas is devalued? To qualify for a patent, for example, an idea can't be "obvious". How does that apply when ideas are found by brute force?

The first group to mimic evolution in patent design – pioneering the use of so-called genetic algorithms (see "As nature intended", page 80) – was led by John Koza at Stanford University in California in the 1990s. The team tested their algorithms by seeing if they could reinvent some of the staples of electronic design: the early filters, amplifiers and feedback control systems developed at Bell Labs in the 1920s and 1930s. They succeeded. "We were able to reinvent all the classic Bell Labs circuits," says Koza. "Had these techniques been around at the time, the circuits could have been created by genetic algorithms."

In case that was a fluke, the team tried the same trick with six patented eyepiece lens arrangements used in various optical devices. The algorithm not only reproduced all the optical systems, but in some cases improved on the originals in ways that could be patented.

The versatility of this type of algorithm is clear from the showcase of evolved inventions at the annual Genetic and Evolutionary Computation Conference (GECCO). Innovations at the 2015 event included efficient swimming gaits for a four-tentacled, octopus-like underwater drone – evolved by a team at the BioRobotics Institute in Pisa, Italy – and the most fuel-efficient route for a future space probe to clean up low Earth orbits. Engineers at the European Space Agency's advanced concepts lab in Noordwijk, the Netherlands, treated the task like a cosmic version of the famous travelling salesman problem – but instead of cities, their probe visits derelict satellites and dead rocket bodies to nudge them out of orbit.

ROYA HAMBURGER



However, the big prize at GECCO is the human competitiveness award, or "Humie", for inventions deemed to compete with human ingenuity. The first Humie, in 2004, was awarded for an odd-shaped antenna, evolved for a NASA-funded project. It worked brilliantly even though it looked like a weedy sapling, with a handful of awkwardly angled branches, rather than a regular stick-like antenna. It certainly wasn't something a human designer would produce.

That is often the point. "When computers are used to automate the process of inventing, they aren't blinded by the preconceived notions of human inventors," says Robert Plotkin, a patent lawyer in Burlington, Massachusetts. "So they can produce designs that a human would never dream of."

The Humie winner in 2015 was a way to improve the accuracy of super-low-power computers. So-called approximate computers are built from simple logic circuitry that consumes very little power but can make a lot of mistakes. By evolving smart software routines for such computers, Zdenek Vasicek at Brno University of Technology in the Czech Republic was able to correct many of the errors introduced by the simple design. The result is a greener chip for use in applications where computational exactness doesn't matter, like streaming music or video.

There's just one problem with using genetic algorithms: you need to know in advance what you want to invent so that your algorithm can modify it in fruitful ways. "Genetic algorithms work well when you already know all the relevant features and can vary them until you get a solution that satisfies all your fitness constraints," says Tony McCaffrey, chief technology officer of Innovation Accelerator based in Natick, Massachusetts. Nolan agrees: "Genetic algorithms tend to be good at optimising pre-existing inventions but typically not ones of great commercial value." That's because they don't take big, inventive steps, he says, and so have less chance of making a commercially valuable hit.

Innovation Accelerator's approach is to use software to help inventors notice easily missed features of a problem that, if addressed, could lead to a novel invention. "An invention is something new that was not invented before because people overlooked at least one thing that the inventor noticed," says McCaffrey. "If we can get people to notice more obscure features of a problem, we raise the chances that they will notice the key features needed to solve the problem."

To do that, the firm has written software ›

One algorithm found a way to make helmets safer using magnets

INTERVIEW

The art of programming

Computers can paint and make discoveries – the challenge is to teach them to code, says AI researcher **Simon Colton**

Simon Colton is professor of digital games technology at Falmouth University, UK. He works on software that behaves in ways that would be deemed creative if seen in humans – such as painting the image of him above

You designed software called HR to make its own discoveries. Has it had much success?
One thing HR came up with was a classification of mathematical structures known as Latin squares. Like a Suduko puzzle, these are grids of symbols where each row and column contains every symbol. HR produced some of the first algebraic classifications of these structures.

A version of HR also independently came up with Goldblach's conjecture – that every even integer greater than 2 can be expressed as the sum of two primes.

Are mathematicians interested in using the system?
We found that mathematicians like software to do the boring grunt work – the massive calculations and trivial proofs they know are true. But creative things like inventing concepts and spotting conjectures they like to do themselves. I once sent Herbert Simon, the Nobel prizewinning economist and computer scientist, an email about a conjecture that HR had proved. He later told me that he hadn't read to the end because he wanted to solve the puzzle himself. His wife said she had to call him to bed.

How do you make software discover things?
You give it data that you want to find something out about, but rather than looking for known unknowns – as with machine learning, where you know what you're looking for but not what it looks like – it tries to find unknown unknowns.

We want software to surprise us, to do things we don't expect. So we teach it how to do general things rather than specifics. That contradicts most of what we do in computer science, which is to make sure software does exactly what you want. It takes a lot of effort for people to get their heads round it.

Can computers make breakthroughs?
I think we will only see computers making true discoveries when software can program itself. The latest version of HR is specifically designed to write its own code. But it's a challenge; it turns out that writing software is one of the most difficult things that people do. And, ultimately, there are mathematical concepts that you can't turn into code, especially ones dealing with infinity.

Another program of yours, The Painting Fool, creates portraits. How do people respond to this type of creativity?
Mathematicians will accept a computer as being creative if it produces great results over and over again. But in the art world, people take more convincing.

When you buy a painting you buy it for many reasons, only one of which is that it will look good with your sofa. When you like a painting, you're celebrating the humanity that went into it. How can we get software to fit in with that?

I don't want to do Turing tests where we try and confuse people about who or what is doing something. We want people to relate to what the software does on its own terms. But computers won't replace people in the creative industries because we will always pay for humanity – for blood, sweat and tears.
Interview by Douglas Heaven

BRIAN KERSEY/PRESS ASSOCIATION

AS NATURE INTENDED

Genetic algorithms tackle the problem of design by mimicking natural selection. Desired characteristics are described as if they were a genome, where genes represent parameters such as voltages, focal lengths, or material densities, say.

The process starts with a more or less random sample of such genomes, each a possible, albeit suboptimal, design. By combining parent genomes from this initial gene pool – and introducing "mutations" – offspring are created with features of each parent plus potentially beneficial new traits. The fitness of the offspring for a given task is tested in a simulation. The best are selected and become the gene pool for the next round of breeding. This process is repeated again and again until, as with natural selection, the fittest design survives (see diagram, opposite).

As well as evolving new designs, evolutionary algorithms can be used to evolve "parasites" that inflict maximal damage to test safety or security features. "Nature has been very good and very creative at finding loopholes in every possible complex system," says Eric Bonabeau of Icosystem of Cambridge, Massachusetts, who has used this technique to improve the design of ships for the US navy.

that lets you describe a problem in human language. It then "explodes" the problem into a large number of related phrases and uses these to search the US Patent and Trademark Office database for inventions that solve similar problems. But similar is the operative word, says McCaffrey. The system is designed to look for analogues to the problem in other domains. In other words, the software does your lateral thinking for you.

In one example, McCaffrey asked the system to come up with a way to reduce concussion among American football players. The software exploded the description of the problem and searched for ways to reduce energy, absorb energy, exchange forces, lessen momentum, oppose force, alter direction and repel energy. Results for how to repel energy led the firm to invent a helmet that contained strong magnets to repel other players' helmets, lessening the impact of head clashes. Unfortunately, someone else beat them to the patent office by a few weeks. But it proved the principle, says McCaffrey.

In another case, the software duplicated a ski-maker's recent innovation. The problem was to find a way to stop skis vibrating so skiers could go faster and turn more safely. The manufacturer eventually stumbled upon an answer, but Innovation Accelerator's software was able to find it quickly. "A violin builder had a method to produce purer music by reducing vibrations in the instrument," says McCaffrey. "The method was applied to the skis and made them vibrate less."

"Ninety per cent of problems have already been solved in some other field," says McCaffrey. "You just have to find them." He now plans to use IBM's supercomputer Watson, which draws inferences from millions of documents, to help his system understand patents and technical papers far more deeply.

The technology at Nolan's firm, Iprova, also helps inventors to think laterally – but with ideas derived from sources far beyond patent documents. The company is unwilling to reveal exactly how its Computer Accelerated Invention technique works, but in a 2013 patent, Iprova says it provides clients with "suggested innovation opportunities" by interrogating not only patent databases and technical journals, but also blogs, online news sites and social networks.

Of particular interest is the fact that it alters its suggestions as tech trends on the internet change. The result seems to be extremely productive. "We use our technology to create hundreds of high-quality inventions per month, which we then communicate to our customers," says Nolan. "They can then choose to patent them." If their wide range of customers in the healthcare, automative and telecommunicaitons industries is anything to go by, Iprova appears to have hit pay dirt. One of its clients is Philips, a major technology multinational. Such firms don't add outside expertise to their R&D teams lightly.

All this means that algorithm-led discovery is likely to be the most productive inventing process of the future. "Human inventors who learn to leverage computer-automated innovation will leapfrog peers who continue to invent the old-fashioned way," says Plotkin

But where do we draw the line between the two? "I don't think there is a clear separation between human and algorithm," says Eric Bonabeau, founder of Icosystem, a company based in Cambridge, Massachusetts. "The key is to find the right division of labour."

"Human inventors who use automated innovation will leapfrog those who don't"

Icosystem uses genetic algorithms to optimise everything from inventions to business processes – an approach Bonabeau calls "enhanced serendipity".

However, if the division of labour is too much on the computer's side, it could undermine the patent system itself. Currently a "person having ordinary skill in the art" must believe that an invention isn't obvious if it is to be granted a patent. But if inventors are only tending a computer, the inventions that arise could be deemed an obvious output of that computer, like hot water from a kettle.

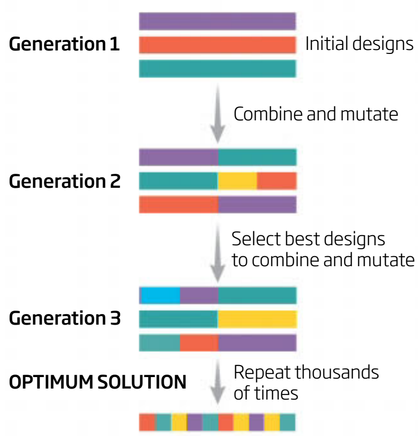
These concerns have already been raised with drug discovery, says Gregory Aharonian, a consultant based in San Francisco, who specialises in patents. "If drug discovery tools become so powerful that a researcher is just overseeing the tools' activity, does that make the whole process obvious and so not patentable? Industry could be shooting itself in the foot by developing such technology."

Another concern is that broad access to smart invention tools could speed up human technological development. Making the resulting gadgets may consume Earth's resources all the quicker. McCaffrey is more optimistic. "I am really impressed with engineers who are creating ways to improve housing, food storage, crop growth, water purification and transportation in the developing world," he says. "I sincerely hope we use this emerging invention assistance technology to address the really important problems faced by humanity."

Chance favours the prepared mind. If Wilbur Wright hadn't been thinking about his problem, he may never have had his eureka moment. "Automating that amounts to making accidental encounters orders of magnitude more efficient," says Bonabeau. "In other words, outsource serendipity to the algorithm." ■

It's evolutionary

Genetic algorithms try to find the optimum solution to a problem by repeatedly combining and mutating the best in each generation of potential solutions



I WORK THEREFORE I AM

In the Western world work defines us as never before. That might be a good thing, argue **Michael Bond** and **Joshua Howgego**

"WHAT do you do?" It's simultaneously the most common and least elegant way to begin a conversation with a stranger. But it sure gets to the rub. Our work permeates our sense of self. Often that begins with our names: if you are an English Smith, a German Schmidt or an Italian Ferraro, you are just one of many with a brand identity determined by the employment your ancestors took.

In the rich countries of the world, where switching jobs is often routine and opportunities for self-expression and development outside of work are greater than ever before, you might expect this identification with work to be diminishing. Not so. In a 2014 Gallup survey, 55 per cent of US workers said they got a sense of identity from their job, a figure that rises to 70 per cent for college graduates. In an increasingly automated world where the nature of work is in flux, that could present a problem. But it is also an opportunity: start to unpick the reasons why what we do means so much to us, and the effects it has on us, and we can begin to make work work better for all of us.

In doing so, it's important to first realise how that work has meant very different things at different times. We have evidence for employer-employee relationships stretching back thousands of years (see "The past of the pay cheque", right), but the concept of working in a distinct profession for a set number of hours each week is a relatively recent one. Even in medieval Europe, when the rise of differentiated professions led to the invention of surnames, our sense of belonging was more likely to be determined by our family, religion or the place we lived, says Benjamin Hunnicutt, who studies the history of work and leisure at the University of Iowa. It was only with the rise of paid employment in the 19th century that the notion of work as an end in itself – and a source of identity – begins to crop up, he says.

Wind forward to today and one thing is for sure: work does fill a lot of our lives. Although in rich countries the average amount of time people work each year has declined over the past half-century – from around 2100 hours in 1960 to below 1600 hours in 2005, according to a 2011 OECD report – factors such as the rise in paid leave account for a lot of that. For white-collar workers not on vacation, work dominates. In 2005, the proportion of high-skilled people in the UK working at least 50 hours a week hit 20 per cent. That has since gone down a bit, but an analysis published in 2015 shows that such extreme working hours have been on an overall upward trajectory in the US, Canada and Europe since 1970.

Even when we aren't at work, it can feel like we are. Smartphones mean

3000 BC

THE WEALTHY BEGIN EMPLOYING OTHER PEOPLE IN SPECIALISED TRADES

TESSA BUNNEY/MILLENNIUM IMAGES, UK

white-collar workers are connected to their jobs at all times. "Modernist distinctions like home-office, work-leisure, public-private and even self-other no longer hold fast," wrote Princeton University sociologist Dalton Conley in his 2009 book *Elsewhere, USA*. Since then, the proliferation of mobile technologies means this always-on culture has spread enormously, he says.

It's easy to see that as a bad thing for ourselves and our relationships with others. And sure, work can be long, stressful, boring and just plain hard. But it's not all bad.

"The miserable effects of unemployment are pretty well documented by social scientists," says David Frayne at Cardiff University, UK. That goes beyond simply the poverty that usually accompanies unemployment. In 2005, Brian Faragher, then at the University of Manchester in the UK, and his colleagues looked at 485 previous studies of the relationship between job satisfaction and health. They showed that people who were happy in their jobs were more likely to be healthy, and in particular were less likely to experience depression, anxiety or low self esteem compared with those less satisfied with their jobs. A review carried out for the UK government in 2006 showed that any stress work creates is, on balance, likely to be outweighed by the problems of not having a job.

Live to work?

That tallies with an idea first articulated by the Austrian psychiatrist Viktor Frankl in the 1940s. While held in concentration camps during the second world war, he helped fellow prisoners endure the horror around them by getting them to focus on the lives they might later lead. In his 1946 book *Man's Search for Meaning*, he argued that these future lives could hold meaning, and that one way of finding it was through work. "Everyone has his own specific vocation or mission in life to carry out a concrete assignment which demands fulfillment," he wrote.

In 2014, Patrick Hill at Carleton University in Ottawa, Canada, and Nicholas Turiano at West Virginia University set out to test this relationship between purpose and well-being. They used data on over 6000 people, captured in the late 1990s as part of a US longitudinal survey. They had been asked how strongly they agreed or disagreed with three statements: "Some people wander aimlessly through life, but I am not one of them"; "I live life one day at a time and don't really think about the future"; "I sometimes feel as if I've done all there is to do in life". Following up on the participants' subsequent life histories, the researchers found a strong link between mortality and the way the volunteers had answered the questionnaire. Those who had more ›

THE PAST OF THE PAY CHEQUE

Perhaps it's no surprise that one of the earliest known examples of writing features two basic human concerns: alcohol and work. About 5000 years ago, the people living in the city of Uruk, in modern day Iraq, wrote in a picture language called cuneiform. On one tablet excavated from the area we can see a human head eating from a bowl, meaning "ration", and a conical vessel, meaning "beer". Scattered around are scratches recording the amount of beer for a particular worker. It's the world's oldest known payslip, implying that the concept of worker and employer was familiar five millennia ago.

It was not ever thus. Çatalhöyük in what is now Turkey was one of the first towns. Houses and human remains dating from its foundation some 9000 years ago are all very similar, suggesting equality. "Everyone was involved in small-scale farming or hunting," says Ian Hodder, an anthropologist at Stanford University in California who has excavated at Çatalhöyük since 1993. No one owned the land, and produce was shared. The residents of this city are unlikely to have considered their daily chores "work", says Hodder. "My view is that they would see it as just part of their daily activities, along with cooking, rituals and feasts that were such an important part of their lives."

Change came a few thousand years later. The trigger seems to have been an agricultural revolution in which new methods of cultivation and animal domestication increased food production and allowed some individuals to build up wealth. The surplus food freed some from toiling in the fields to focus on specialised tasks such as carpentry or pottery. By the time our beer ration card was written, a professional class had been born, a transition that seems to mark the beginning of what we know as work.

These same conditions set the scene for other working practices that still endure. "There was a change from social sharing to hierarchies," says Hodder. "This was an inevitable consequence of living in large communities and intensifying agricultural production." Throughout the Middle East and China between 6000 and 4000 BC, as towns became bigger, a powerful elite commandeered not only resources, but labour.

Back then you didn't work for money, but rather food, shelter and protection. Since then work has slowly morphed into the employment we have today. Our human needs, for food and shelter, haven't changed – but the way we get them has. Alison George ■

of a sense of purpose lived longer, even when other psychological and health conditions were accounted for – and the trend held for people of all ages.

Why that should be remains unclear. Hill thinks having a direction, something work provides, may give people a reason to take better care of themselves and thus lead them to adopt healthier lifestyles (see "Why am I here", page 28). It may also help us cope with stress.

"People who have a strong sense of meaning in life view stressors differently," says Neal Krause, of the University of Michigan in Ann Arbor. "They are more likely to see them as challenges rather than unwanted and painful setbacks." So although work may be a source of stress, it might also help us face whatever else life throws at us.

Work to live

Meaningful work seems to stave off cognitive decline, too. A study led by Carole Dufouil at INSERM, the French national health research institute, showed that for every extra year someone works before retiring, their risk of developing dementia decreases by 3 per cent. This follows a previous study by researchers at King's College London and elsewhere that found that people with dementia who had worked beyond age 65 delayed the onset of their symptoms by more than six weeks for every extra year worked.

This is not just about staying cognitively active. Patricia Boyle, a neuropsychologist who works with Alzheimer's patients at the Rush University Medical Center in Chicago, says a sense of purpose appears to make older people more resilient on many levels, perhaps because it improves immune function and decreases the risk of vascular diseases – though the biological mechanism is unknown.

All this presumes we can find meaningful employment that allows us to enjoy work's positive psychological benefits. Brent Rosso at Montana State University has come up with a list of six attributes that make work meaningful, based on years of academic surveys. Almost any job can exhibit at least one of these attributes (see "Find your meaning", right), although Rosso reckons that an individual's culture and personality will influence which ones they find meaning in.

1400s

SURNAMES REFLECTING A PERSON'S TRADE BECOME WIDELY USED IN EUROPE

LOUIS QUAIL/PICTURETANK

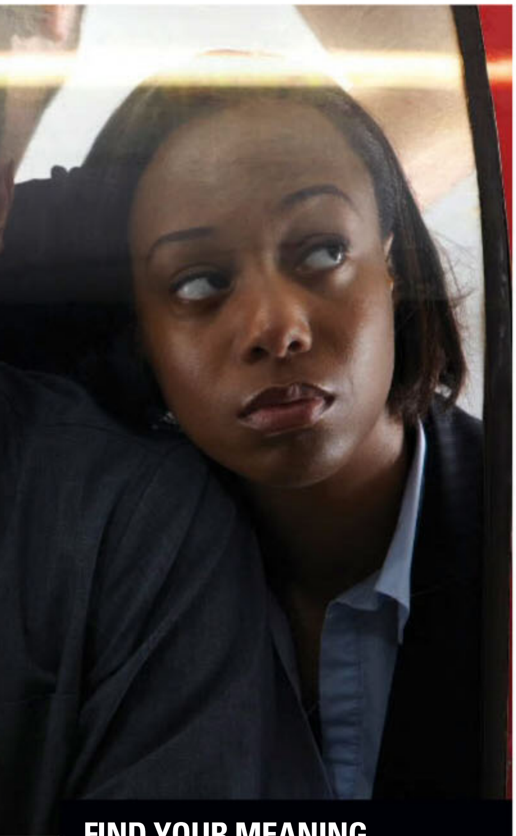
One important point is a sense of belonging: identifying with the people you work with has been shown to increase not only job satisfaction, but productivity, too. Alex Pentland of the Massachussets Institute of Technology, for example, has shown that the more cohesive and communicative a team is – the more they chat and gossip – the more they get done.

As the nature of work changes over the coming years, its effects on our psyches will no doubt continue to evolve. The key will be to change work so that we can continue to find meaning in it – because not all work is made equal. As the American poet and philosopher Henry David Thoreau once said: "It is not enough to be industrious. So are the ants." ■

NICE WORK

IF YOU CAN GET IT

Technology has always been seen as a threat to human jobs. **Jon White** asks if this time it's serious



CHRIS TURNER/GETTY

FIND YOUR MEANING

People find meaning in work in six main ways; which aspects someone finds most important depends on them and their society

AUTHENTICITY Going to work makes you feel you are accessing your "true self" – maybe that you are following a calling or can be yourself.
AGENCY You are able to make significant decisions and feel as if you "make a difference". This taps into our desire to believe that we have free will.
SELF-WORTH Your job make you feel valuable; you are able to see milestones of achievement, no matter how small.
PURPOSE You see your work as moving you closer to a strongly held goal. The downside is that you are more likely to sacrifice pay and personal time too.
BELONGING It's not what you do, it's who you do it with. You belong to a special group of colleagues, even if your job seems mundane or poorly rewarded.
TRANSCENDENCE Your job is about sacrifice for a greater cause. Your meaning comes from following this, or perhaps a truly inspirational boss.

JOHN MAYNARD KEYNES always assumed that robots would take our jobs. According to the British economist, writing in 1930, it was all down to "our means of economising the use of labour outrunning the pace at which we can find new uses for labour". And that was no bad thing. Our working week would shrink to 15 hours by 2030, he reckoned, with the rest of our time spent trying to live "wisely, agreeably and well".

It hasn't happened like that – indeed, if anything many of us are working more than we used to (see "I work therefore I am, page 82). Advanced economies that have seen large numbers of manual workers displaced by automation have generally found employment for them elsewhere, for example in service jobs. The question is whether that can continue, now that artificial intelligence is turning its hand to all manner of tasks beyond the mundane and repetitive.

Fear of machines taking jobs dates back at least as far as the Luddites, a group of British weavers who went on a mill-burning rampage in 1811 when power looms made them redundant. Two centuries on, many of us could face the same predicament. In 2013 Carl Frey and Michael Osborne of the Oxford Martin Programme on Technology and Employment at the University of Oxford looked at 702 types of work and ranked them according to how easy it would be to automate them. They found that just under half of all jobs in the US could feasibly be done by machines within two decades.

The list included jobs such as telemarketers and library technicians. Not far behind were less obviously susceptible jobs, including models, cooks and construction workers, threatened respectively by digital avatars, robochefs and prefabricated buildings made in robot factories. The least vulnerable included mental health workers, teachers of young children, clergy and choreographers. In general, jobs that fared better required strong social interaction, original thinking and creative ability, or very specific fine motor skills of the sort demonstrated by dentists and surgeons.

Others find that list overblown. A recent working paper for the rich- ›

1860s

THE FIRST COMPANY PENSION SCHEMES START TO APPEAR

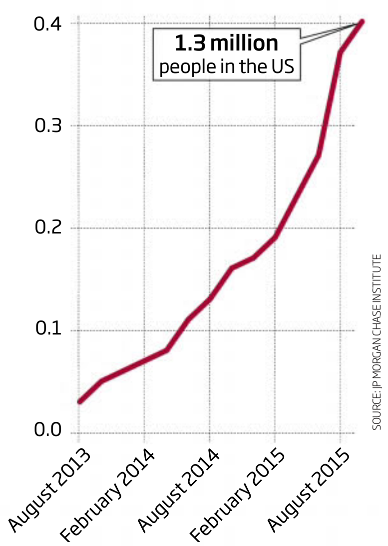
world OECD club suggests that AI will not be able to do all the tasks associated with all these jobs – particularly the parts that require human interaction – and only about 9 per cent of jobs are fully automatable. What's more, past experience shows that jobs tend to evolve around automation.

According to this more Keynesian view, technological progress will continue to improve our lives. The most successful innovations are those that complement rather than usurp us, says Ben Shneiderman, who founded the human-computer interaction lab at the University of Maryland. Witness for instance the prominence of "cobots" at last month's annual automation expo in Chicago. Such robots are designed to work alongside people, making their work safer and easier, not replacing them. "Technologies are most effective when their designs amplify human abilities," says Shneiderman. They could help us solve problems, communicate widely, or create art, music and literature, he believes.

The weight of expert opinion is behind him. In 2014 the Pew Research Center, a US think tank, asked 1896 experts whether they thought that by 2025, technology would have destroyed more jobs than it creates. The optimists outnumbered the pessimists.

Get the gig

The **percentage** of the US population earning some income from online labour platforms such as Uber is rising fast



1880s

GERMANY INTRODUCES THE FIRST STATE WELFARE SYSTEM

That's not to deny that AI is spreading into some surprising settings – whether it be organising nightly maintenance on Hong Kong's subway system, or helping out with subtle legal research, as does ROSS, an AI assistant built on IBM's Watson computer (see "Intelligence reinvented, page 70). This suggests that AI could still cause short-term turbulence in the labour market.

One unfolding example is the gig economy. Here AI systems serve up a platter of casual labour to a convenient app for consumers. Examples include the taxi firm Uber and outfits like TaskRabbit, which helps people find casual labourers to complete all sorts of chores. Although the gig economy is still small in absolute terms, a study of 1 million people who bank with JP Morgan Chase suggested that the number of people getting some of their income from the gig economy has increased tenfold in two years (see graph, left).

In such set-ups, workers are typically considered self-employed contractors, so the company has no obligation to keep supplying work or provide benefits like holiday pay or pensions. That has already led to strikes.

How can we adapt? The answer might simply be to update our social frameworks to reflect the new reality of work. Many countries are considering new regulatory frameworks for the gig economy. In the US Uber and Lyft, another taxi service, face ongoing lawsuits about the classification of drivers as contractors rather than employees. Drivers may vote with their wheels, too: Transunion Car Service, established in New Jersey in 2015, is an Uber-like taxi business owned by its drivers that promises health and retirement benefits.

Others are thinking more radically about how to reconfigure our whole relationship with work (see "All play and no work", right). That speaks to an important point: ultimately we, not AI, are in charge of our own destiny. Given the benefits of work for our health and well-being (see "I work therefore I am," page 82), maybe we'll opt not to abolish fulfilling, rewarding work. "There will be inequities and disruptions, but that's been going on for hundreds of years," says Shneiderman. "The question is: is the future human-centred? I say it is." ■

STEPHEN WILKES/GETTY

ALL PLAY AND NO WORK

The rise of automation could require a radical rethink of how we distribute the fruits of labour, says **Hal Hodson**

EACH month, Nathalie Kuskoff repeats the process that ensures her family's security. Her two young children both have chronic illnesses, so their apartment in southern Finland is mostly paid for by the government, which also helps with childcare, medical bills and education. "I get a lot of different social benefits because of my situation – I mean a lot," she says. They come at a price: relentless form-filling.

Most developed economies have some form of welfare state to redistribute wealth from the economically active to those who are unemployed or can't work. People differ about who they think should get what, but few dispute the principle of a basic safety net.

But as Kuskoff and many others find,welfare on the basis of need is a cumbersome, bureaucratic affair. And as automation continues its march, many more of us may find ourselves caught in its net (see "Nice work if you can get it", page 85). This is the background to a radical idea to rejig the way we distribute welfare that has recently been in the headlines: universal basic income.

At its simplest, the premise is to replace welfare with a contract promising everyone the same money unconditionally, covering basic human needs – food, shelter, clothing – which people can add to by working. Its proponents cite an array of advantages including higher employment, better community cohesion and improved health. Others see it as an excuse to shirk. Now, as the debate rages, several huge social experiments could settle these differences.

Universal basic income has a long history. Thomas Paine, a US founding father, believed that natural resources were a common heritage and that landowners sitting on them should be taxed and the income redistributed. While the idea has never fully materialised, neither has it entirely gone away. In a few corners of the world variants are discreetly part of the furniture. In Alaska, for example, an annual dividend from state oil revenues is paid to citizens each year – a windfall of $2072 per person in 2015.

The idea has been gaining adherents across the political spectrum. In the UK, for example, proponents include the left-wing Green party and a right-wing think tank, the Adam Smith Institute. The main opposition Labour party has also toyed with the idea. In Canada, testing the approach forms part of the policy platform of the Liberal party, elected to government in 2015.

Licence to laze

The perceived threat of automation is a timely spur to revisiting universal basic income, but it isn't the only one. First, conventional welfare systems are not just bureaucratic but also costly. Even though basic income pays out more money, it cuts the costs of red tape. Various schemes have been proposed that look affordable, including one last year from the RSA, a UK think tank (see chart, page 88). Basic income also promises to eliminate financial disincentives to work that bedevil many welfare systems – under a basic income system, you always earn more if you work.

The most entrenched criticism is that too many would exploit a guaranteed income to sit on their hands, grinding the economy to a halt. There are signs, though, that this is too gloomy a view.

For four years beginning in 1975, the 10,000 citizens of Dauphin in Manitoba, Canada, were guaranteed a basic level of financial security: if their monthly income dropped below a certain level, the government would top it up. Support for this experiment soon dried up, and it was never properly analysed.

Evelyn Forget at the University of Manitoba in Winnipeg recently revisited the experiment, comparing public records from Dauphin with those from similar small towns. Forget found the only groups that spent less time in work during the trial were teenage boys and new mothers. The boys were staying in school rather than bowing to pressure to take agricultural jobs, and the mothers were nursing. What's more, Dauphin had noticeably lower hospitalisation rates and fewer depression-related illnesses.

That was just one small-town trial. But in Alaska, experience suggests ›

Basic income

15 million

Number of UK jobs threatened by automation

SOURCE: BANK OF ENGLAND

One way to ease the impact of job losses is for the state to pay everyone a basic income. Here's how much people would get under a scheme devised by the RSA, a UK think tank

Age	
0-4	£3387 (£4290 for the first child)
5-24	£2925
25-64	£3692
65+	£7420

SOURCE: RSA

The total cost of this scheme is only slightly more than the cost of the present welfare system, which it would replace

Package would cost: **£282.7 billion**

Estimated 2016-17 UK welfare bill: **£270 billion**

SOURCE: RSA

SOURCE: OFFICE FOR BUDGET RESPONSIBILITY

The RSA points out that tax breaks handed out by the UK government last year were worth almost double the **extra money needed** to introduce their version of basic income

£12.7 billion shortfall

£19.5 billion cuts in corporation tax and fuel duty in 2015-16

SOURCE: INSTITUTE FOR FISCAL STUDIES

that a basic income could help reduce the rising inequality that has been hobbling world economies. Economist Scott Goldsmith at the University of Alaska Anchorage points out that the state is the only one in the US in which the income of the poorest 20 per cent grew faster than that of the top 20 per cent between the 1980s and 2000.

Now experiments are afoot to test such effects more exactingly. One, in Finland, is one of the grandest social experiments ever conceived, says social scientist Jurgen De Wispelaere at the University of Bath, UK. "There's nothing like it happening anywhere." Starting in 2017, as many as 10,000 Finns receive a no-strings-attached monthly income of €600 for two years. That sum is designed to guarantee subsistence, says Helsinki University's Ville-Veikko Pulkka, who has worked for Finland's social insurance department Kansaneläkelaitos (Kela), covering housing, food and services like water and electricity.

Kela is planning to publish the full trial design, but the point is to test whether a basic income gets more people working. "Removing disincentives to joining the labour force is the key task given to us by government," says Pulkka. The ideal is to give people a platform to enter the labour market on their own terms.

In Finland, that taps into a well-anchored social principle called universalism: that the same services and education should be available to everyone. "At some level, people want to believe in this system," says Pulkka. Kuskoff would certainly be interested in participating. "Getting the money without all the paperwork sounds like heaven," she says.

Reducing bureaucracy is the driver of a similar large-scale experiment in the Netherlands. It started when the Dutch government passed a law giving municipalities the responsibility for administering welfare. Their staff baulked at taking on the job of continually vetting welfare applicants as the central government had been doing. "People realised this was going to do their heads in and they needed to change it," says De Wispelaere.

Nineteen municipalities are now changing how they administer welfare payments, says Sjir Hoeijmakers, who is coordinating the experiments. Each tests different supposed benefits of a basic income like those Forget flagged in Dauphin. In Eindhoven, for example, the focus is on whether the changes help build strong neighbourhoods, while other municipalities are concentrating on randomised controlled trials to determine how individuals fare. A certain amount of freeloading is expected, says De Wispelaere: "In any policy you have good and bad. We want to know how many people move to the couch, and then compare the positive effects."

Private companies are also getting in on the act. Y Combinator, a venture

1929

THE WALL STREET CRASH CUTS THE WORKING WEEK TO FIVE DAYS

THOMAS DWORZAK/MAGNUM PHOTOS

capital firm with stakes in the taxi app Uber, is running a basic income experiment, with a pilot phase in Oakland, California, that began in January 2017.

The most important arguments in favour of basic income are about improved health and well-being, says Louise Haagh, a social economist at the University of York, UK. These too are now coming under more scrutiny. For example, a study of 1000 children by Kimberly Noble of Columbia University in New York found a strong positive correlation between family income and brain development. One theory is that families with a secure income can focus extra resources on their children. "But with purely correlational data we can't say which way the arrow is pointing," says Noble.

To find out, she is running an experiment in which 1000 low-income mothers across the US will receive a basic income for three years. One group will receive a nominal $20 a month, the other $333. Noble's focus is on brain development, not economics, but a pilot study in New York in which money was handed out on trackable, prepaid debit cards suggested freeloading wasn't a problem: of 1100 transactions, most of the money went on groceries. Just three happened at a liquor store.

So is a basic income a panacea?

FINE ART IMAGES/HERITAGE IMAGES/GETTY IMAGES

2003

THE BLACKBERRY 6210 MOBILE PHONE MEANS PEOPLE CAN CHECK WORK EMAILS ANYWHERE

Some, like Kuskoff, who have special care needs, worry that such a system might push them to a harsher edge of the welfare state in the name of homogenisation and efficiency. And Haagh thinks that a half-hearted implementation might entrench, not dissolve, social inequalities by offering rich and poor the same. Governments could end up subsidising companies that give few or no benefits to their workforces, while the lucky few with more conventional employment receive far more. The problem is that a poorly designed basic income "might not end up changing society that much", says Haagh.

Other variants do exist. Negative income tax is a means-tested version of universal basic income: poor people receive a guaranteed income from the government, middle earners aren't taxed, while the rich are.

It sounds fairer, but could have a significant disadvantage, as the work of Silvia Avram at the University of Essex, UK, hints. She asked people to perform a tedious task to earn money under different taxation models. The participants were divided into two groups. One group started with a lump sum that was reduced as they earned – much as would happen under a negative income tax – while the others were taxed as they earned. Both groups ended up with the same money for a given amount of work, but the first group was far quicker to quit the task, suggesting that a well-documented human tendency to loss aversion was kicking in: we are wired to place more importance on minimising losses on what we already have than realising gains of the same value.

For Anthony Painter, director of the Action and Research Centre at the RSA, and author of its report on how basic income could work in the UK, it is an indication that negative income tax wouldn't be as effective at getting people into work as a basic income. Painter and others also think a basic income could benefit society in other ways, freeing up people to look after older relatives and children, or to pursue creative and innovative work that traditionally pays less, like music, arts and invention (see "Eureka machines", page 78).

Such supposed whole-society benefits aren't easy to test objectively, and that might be the most crucial point. If the referendum on basic income that took place in Switzerland in June 2016 is any indication, basic income has a long way to go to gain public acceptance. During the debate, triggered when a group of citizens collected more than the necessary 100,000 signatures for a vote on such constitutional change, no political party endorsed the idea: it was widely seen as indulging shirkers. In the end, 77 per cent of voters rejected it.

However, basic-income campaigners were celebrating that evening, saying their objective was to get people talking. The conversation continues. Maybe the mark of ultimate success for the proponents of universal income, says Hoeijmakers, will be if at parties the unfashionable question "what do you do?" morphs into: "why do you do?". ■


NewScientist THE COLLECTION
THE BIG QUESTIONS
THE UNKNOWN UNIVERSE
THE SCIENTIFIC GUIDE TO A BETTER YOU
THE HUMAN BRAIN
THE WONDERS OF SPACE
BEING HUMAN
15 IDEAS
ORIGIN, EVOLUTION, EXTINCTION
THE QUANTUM WORLD
WILD PLANET
MIND-EXPANDING IDEAS
EINSTEIN'S MIND-BENDING UNIVERSE
ESSENTIAL KNOWLEDGE
EVERYTHING YOU NEED TO KNOW TO MAKE SENSE OF THE 21ST CENTURY

ON THE ROAD AGAIN

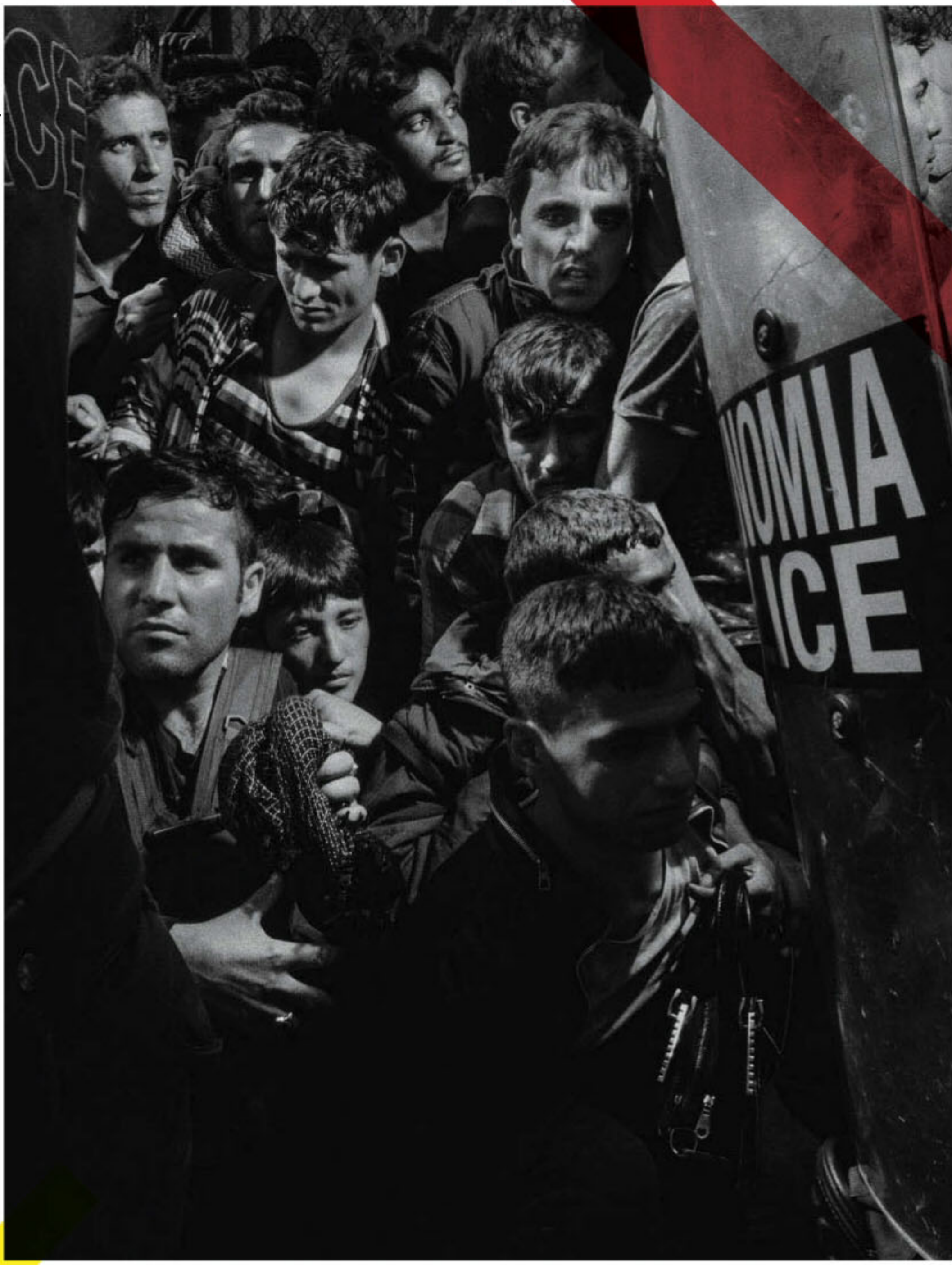
From our origins in Africa we've conquered the world by migrating. Can modern immigration really be a crisis, asks Debora MacKenzie

OPENING PAGE IMAGE: SAM FALCONER ALEX MAJOLI/MAGNUM PHOTOS

HUMANS migrate. It is a characteristic of our species. Yet these days a migration crisis is headline news. More than a million desperate people fled to Europe in 2015, and nearly 4000 died trying. The influx is increasing, starting every spring as the weather improves. The United Nations says Europe faces "an imminent humanitarian crisis, largely of its own making". And it is not alone. The UN has also censured Australia for sending boatloads of refugees to squalid camps in other countries. And President Trump has promised to build a wall while tens of thousands of lone children flee violence in Latin America across the US-Mexican border.

In 2016, the World Economic Forum ranked large-scale refugee flows as its global risk of highest concern. When the US Council on Foreign Relations drew up its top 10 priorities for conflict prevention, it included political instability in the EU caused by the influx of migrants. Concerns about refugees and economic migrants are grist to the mill for those who persuaded Britain to vote to leave the EU. And there's no doubt that migration will increase as the world's economy becomes more globalised, and as demographic and environmental pressures bite.

Should we be alarmed? What is the truth about migration? It is an emotive issue. But the scientific study of what happens when humans move is starting to supply some non-emotive answers. It's showing that many widespread beliefs don't hold up to scrutiny. "Concern about immigrants falls sharply when people are given even the most basic facts," says Peter Sutherland, the UN Special Representative for migration. One analyst even says that removing all barriers to migration would be like finding trillion dollar bills on the sidewalk.



A SPECIES ON THE MOVE

200,000 years ago

Homo sapiens appears in what is now East Africa

125,000

Humans leave Africa for western Asia, but settlements later replaced by Neanderthals

The millions fleeing Syria have shone a spotlight on refugees, but that tragedy is just a small part of a bigger picture. More than 240 million people worldwide are international migrants. Refugees account for fewer than 10 per cent of the total and, in theory, they are the least contentious group, because many countries have signed international commitments to admit them. The rest are moving to work, or to join family members who have jobs.

When such people travel with refugees, they are often derided as "just" economic migrants. This is unfair, says Alex Betts, head of the Refugee Studies Centre at the University of Oxford. Whether or not they meet the official definition of a refugee, many are escaping dire conditions that pose a threat to their survival. Although globalisation of the world's economy has lifted millions out of poverty, it has not been able to create enough jobs where there are people in need of work. Aid funds are starting to address this problem – but for the most part people must go where there are jobs.

That's why some see migration as a crisis. The 2008 financial crash spawned insecurity about jobs and concerns about economic migrants. Several populist parties took the opportunity to warn of a flood of freeloaders at the gates, increasing the issue's political visibility and hardening the policies of some mainstream parties, including in the UK, where the desire limit free movement of people contributed to the Brexit vote. The US government decided not to bail out firms that hired too many immigrants. Spain paid migrants to leave – even after they had stopped coming as jobs disappeared. And feelings of insecurity remain.

"The logic driving this is the idea that migrant workers present additional competition for scarce jobs," says Ian Goldin at the University of Oxford. Indeed, it is probably part of our evolved nature to think that more for you means less for me (see "The Origins of xenophobia", above). But that's not how modern economies work.

If economies really were zero-sum games in this way, wages would go down as labour supply increased and natives might well lose jobs to immigrants. But no modern economic system is that simple, says Jacques Poot at the University of Waikato, New Zealand. The knock-on of economic migration is that increased labour also brings an increase in profit, which business owners can invest in more production. They can also diversify, creating opportunities for a broader range of workers. In addition, migration means workers can be more efficiently matched to demand, and make the economy more resilient by doing jobs natives won't or can't do.

"More people expand the economy," says Goldin, because people are moving from where they cannot work productively to where they can. In a survey of 15 European countries, the UN's International Labour Organisation (ILO) found that for every 1 per cent increase in a country's population caused by immigration, its GDP grew between 1.25 and 1.5 per cent. The World Bank estimates that if immigrants increased the workforces of wealthy countries by 3 per cent, that would boost world GDP by $356 billion by 2025. And removing all barriers to migration could have a massive effect. A meta-analysis of several independent mathematical models suggests it

THE ORIGINS OF XENOPHOBIA

All the evidence suggests that migrants boost economic growth. So why don't we just fly people who want to work to countries where there are jobs and welcome them with open arms? Prejudices rooted in humanity's evolutionary past may be partly to blame.

"Perceptions of competition drive a lot of our thinking and are difficult to avoid," says Victoria Esses at the University of Western Ontario in London, Canada. Humans think of their support systems as a zero-sum game – so if one person gains, another must lose out. Such perceptions were accurate during our evolutionary history as hunter-gatherers when the appearance of others on our patch meant fewer mastodons or mushrooms for us. If they were close relatives they might share – or at least our common genes would benefit from their success. But anyone displaying different cultural markers was likely to be a competitor. A modern capitalist economy is not a zero-sum game – if you add more workers, it grows (see main story). Regardless of this, our evolutionary hang-ups make it difficult to accept the economic sense in welcoming immigrants.

That's not all. We are instinctively wary of close contact with strangers because in our evolutionary past this helped us guard against infectious disease, says Mark Schaller at the University of British Columbia in Canada. Separate groups of people often have different histories of exposure and acquired immunity to pathogens. A disease carried innocuously by one might devastate another, as happened to the Native Americans after Europeans arrived.

Steven Neuberg at Arizona State University in Tempe notes that groups also evolve different survival-enhancing practices. "Foreigners with different rules might interfere with the social coordination you need to do important tasks, or might get members of your group to follow their rules instead," he says. "Chaos could emerge if your group makes decisions by consensus but theirs is authoritarian."

Schaller and Neuberg believe that for both these reasons, human cultures evolved to be wary of close interaction with people who were different from their group.

This xenophobia persists, says Neuberg, who has found that people feel threatened by groups with different values of many kinds. Ethnic groups in modern cities often form enclaves rather than mixing randomly – which can foster strong local communities but also engenders wider mistrust. To live in multicultural societies, we will need to learn to get past such evolved tendencies.

100,000 – Humans head to east Asia, interbreeding with Neanderthals along the way

75,000 – Humans reach West Africa

60,000 – Humans leave Africa via the Middle East

46,000 – Humans reach Australia

44,000 – Humans reach Europe

BETTMANN/CORBIS

Millions migrated to the Americas in the 19th century, but far more stayed at home

would increase world GDP by between 50 and 150 per cent. "There appear to be trillion-dollar bills on the sidewalk" if we lift restrictions on emigration, says Michael Clemens at the Center for Global Development, a think tank in Washington DC, who did the research.

But who gets those billions? Most of the extra wealth goes to migrants and to their home countries. In 2015, migrants sent home $440 billion, two and a half times the amount those countries received in foreign aid – promoting development and jobs at home. But what do natives of countries that attract migrants get out of it?

In the EU it has been difficult to tease out the effect of free movement of workers from other economic results of membership. However, a study of non-EU member Switzerland is illuminating. Different parts of Switzerland allowed free access to EU workers at different times, enabling Giovanni Peri of the University of California, Davis, to isolate the effects. He found that while the workforce grew by 4 per cent, there was no change in wages and employment for natives overall. Wages increased a little for more educated Swiss people, who got jobs supervising newcomers, while some less educated Swiss people were displaced into different jobs.

Peri has also looked at the situation in the US. "Data show that immigrants expand the US economy's productive capacity, stimulate investment and promote specialisation, which in the long run boosts productivity," he says. "There is no evidence that immigrants crowd out US-born workers in either the short or the long run." Natives instead capitalise on language and other skills by moving from manual jobs to better-paid positions. Peri calculates that immigration to the US between 1990 and 2007 boosted the average wage by $5100 – a quarter of the total wage rise during that period.

THE RELUCTANT MIGRANT

Humans have always migrated. Our species started as African apes and now covers the planet. Tales of migration are central to our religions, our literature and our family histories. And migration is at the heart of modern life. I am a migrant. You may be too. In 2016, 38 per cent of scientists working in the US and 33 per cent in the UK were foreign-born. Yet they may be exceptions to an ancient rule: in fact, few people migrate. And when we do, often it's because we feel we have no other option.

Take our ancient ancestors who left Africa between 65,000 and 55,000 years ago. At the time, humans had evolved 35 different lineages of mitochondrial DNA, a collection of genes that changes very slowly. The migrants were carrying just two of these, which with other DNA data suggests that they could have numbered as few as 1000. The vast majority of human diversity outside Africa stems from this single migration, suggesting this small band of pioneers may not have gone far, occupying the first lands they came to in the Middle East and discouraging followers. Their descendants would then have expanded into further territories when those hunting grounds got crowded. In this way, over tens of thousands of years, humans occupied the world, moving first to Asia and Australia, then to Europe, and finally colonising the Americas.

The biggest emigration the world has ever seen is much more recent. A mass movement of people from Europe to the New World occurred between 1850 and 1910. At its peak, over 2 million people a year were relocating. Nevertheless, the vast majority chose to stay put. On average, only 5 per cent of the population of Britain – among the biggest sources of migrants – left each decade.

Today, just 3.3 per cent of the world's people are migrants, little more than in 1990. Even within the European Union, where citizens are free to live wherever they choose, only 2.8 per cent, 14 million people, now reside outside their native country. "The idea that, without controls, everyone moves is contradicted by the evidence," says Philippe Legrain at the London School of Economics. "Niger is next to Nigeria, Nigeria is six times richer and there are no border controls, but Niger is not depopulated. Sweden is six times richer than Romania, the EU permits free movement, but Romania is not depopulated." Even strong economic incentives are often not enough to tempt us to leave home.

16,000 years ago – Humans reach the Americas

11,000 – Mainland Europeans follow retreating ice into Britain

8000 – Farmers from Middle East invade Europe and interbreed with hunter-gatherers

6000 – In Sumeria, first cities attract migrants and spawn technological innovation

4500 – Yamnaya pastoralists from the steppe invade Europe

Further evidence comes from a meta-analysis Poot did in 2010, which collated all the research done up until that point. It reveals that rises in a country's workforce attributable to foreign-born workers have only a small effect on wages, which could be positive or negative. At worst, a 1 per cent rise caused wages to fall by 0.2 per cent, mostly for earlier generations of immigrants. The impact on the availability of jobs for natives is "basically zero", he says. Any tendency for wages to fall with an increase in immigration can be counteracted by enforcing a minimum wage.

The UK Migration Advisory Committee came to a similar conclusion in 2012. "EU and non-EU migrants who have been in the UK for over five years are not associated with the displacement of British-born workers," it reported. Very recent migrants do have a small impact, but mainly on previous migrants. What's more, the ILO notes that low-skilled migrants do "dirty, dangerous and difficult" jobs, which locals do not want – crop picking, care work, cleaning and the like. Meanwhile, highly skilled migrants plug chronic labour shortages in sectors such as healthcare, education and IT. Nearly a third of UK doctors and 13 per cent of nurses are foreign-born.

Another presumption made about migrants is that they put a strain on benefit systems. This is also not borne out by the evidence. "It is widely assumed that economic migrants are mainly poor people out to live off the tax money of the relatively rich," says human rights expert Ian Buruma. "Most of them are not spongers. They want to work." A lot go not to countries offering generous benefits, but to where there are jobs. Some 82 million people, 36 per cent of the world's current migrants, have moved from one developing country to another, especially from Haiti to the Dominican Republic, Egypt to Jordan, Indonesia to Malaysia and Burkina Faso to Ivory Coast.

Those who do end up in wealthier countries are not the burden people sometimes assume. The Organisation for Economic Co-operation and Development, which represents 34 of the world's wealthiest nations, calculates that its immigrants on average pay as much in taxes as they take in benefits. Research shows that EU workers in the UK take less from the benefits system than native Brits do, mostly because they are younger on average. Moreover, they bring in education paid for by their native countries, and many return to their homeland before they need social security. Based on recent numbers, Britain should conservatively expect 140,000 net immigrants a year for the next 50 years. The Office for Budget Responsibility, the UK's fiscal watchdog, calculates that if that number doubled, it would cut UK government debt by almost a third – while stopping immigration would up the debt by almost 50 per cent.

Illegal migrants make a surprising extra contribution, says Goldin. While many work "informally" without declaring income for taxes, those in formal work often have taxes automatically deducted from their pay cheques, but rarely claim benefits for fear of discovery. Social security paid by employers on behalf of such migrants, but never claimed by them, netted the US $20 billion between 1990 and 1998, says Goldin. That, plus social security contributions by young legal migrants who will not need benefits for decades, is now keeping US social security afloat, he says.

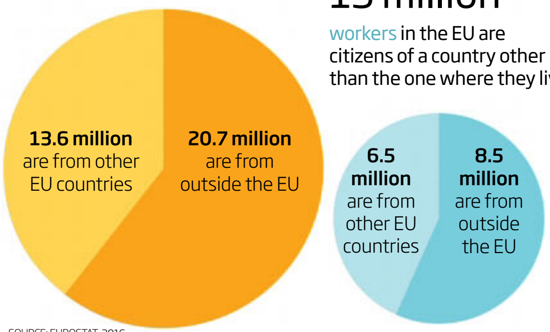
"One of the dominant, but empirically unjustified images is of masses of people flowing in… taking away jobs, pushing up housing prices and overloading social services," write Stephen Castles at the University of Sydney, Australia, and two colleagues in their book, *The Age of Migration*. They argue that an increase in migration is often the result rather than the cause of economic changes that harm natives – such as neoliberal economic policies. "The overwhelming majority of research finds small to no effects of migration on employment and wages," says Douglas Nelson of Tulane University in New Orleans. "On purely economic grounds, immigration is good for everyone." ›

34.3 million

people in the EU are citizens of a country other than the one where they live

15 million

workers in the EU are citizens of a country other than the one where they live



SOURCE: EUROSTAT, 2016

244,000,000

Number of immigrants globally in 2015

Source: UN Population Fund

3.3%

Percentage of people worldwide who are migrants

Source: World Bank

3000 – Western Eurasians move back into eastern then southern Africa

1500 – Humans populate Polynesia

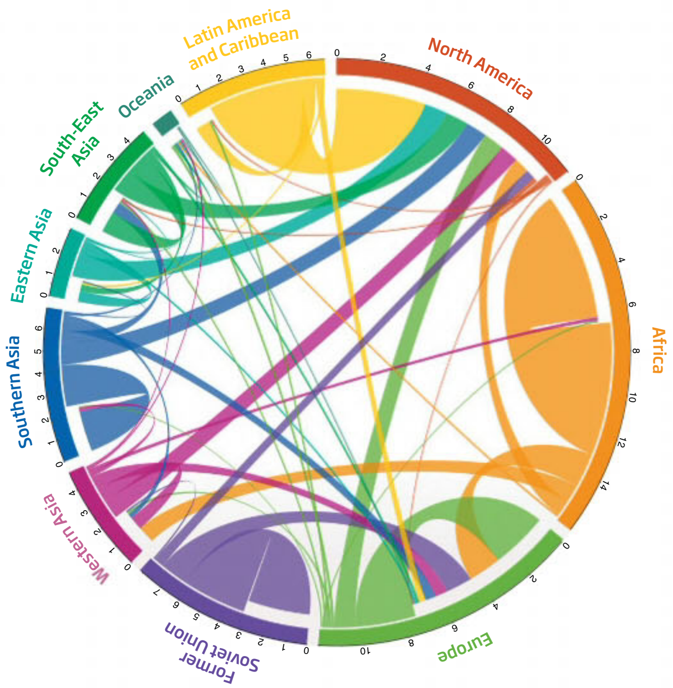
AD 370 – Eurasian Huns move west, driving Germanic tribes into Rome, leading to its fall

450 – Anglo-Saxons from Denmark colonise Britain

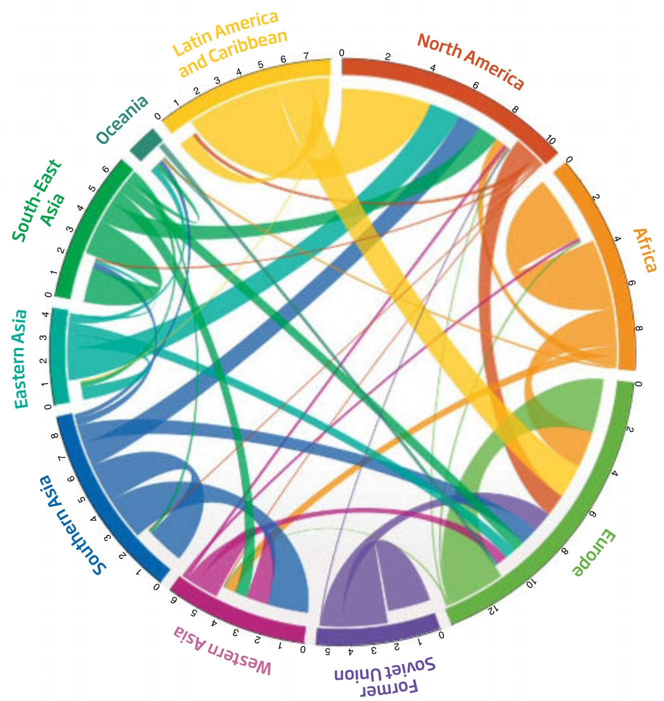
632 – Arabs spread across western Asia and North Africa creating Islamic caliphate

1995-2000

The movement of people from region to region (millions)



2000-2005



PASSPORT TO SUCCESS?

Governments only started to control who entered their country relatively recently. Other than in wartime, authorities worried more about people getting out. Roman and medieval laws kept peasants bound to their farms. In the 1600s, English labourers needed locally issued passes to travel for work, partly to stop them "benefits shopping" for parish poor relief. But controls were largely internal.

External passports were mere requests for safe conduct, rather than restrictive documents determining where you could go, says John Torpey at the City University of New York. This was partly because technology to identify individuals, such as photography, was not widely available until the late 19th century.

But the main reason was that an individual's nationality had little political meaning before the late 1700s. The passport as an instrument of state regulation was born of the French revolution of 1789. At first, ordinary people were issued passes to control internal movement, especially to Paris. But after the king tried to escape, and foreign aristocrats attacked the revolution, the authorities started requiring such papers for exit and entry to the country. The revolution created one of the world's first "nation-states", defined by the "national" identity of its people rather than its monarchs' claims. "This novel importance of the people and their nationality made identity papers integral to creating the modern state," says Torpey.

As the idea of the nation-state spread, so did passports. But as the industrial revolution snowballed in the 19th century, there was pressure to allow free movement of all the factors of production – money, trade and labour. Passport requirements were widely relaxed across Europe – in 1872, the British foreign secretary, Earl Granville, even wrote: "all foreigners have the unrestricted right of entrance into and residence in this country". The situation was similar in North America.

In the early 20th century, European legal experts were divided over whether states even had the right to control people's international movements. But the nationalism that was propelling Europe towards war changed that. Among other things, it meant foreigners might be spies. Passport controls were reapplied, and never lifted again.

AD 800 – Vikings colonise Britain

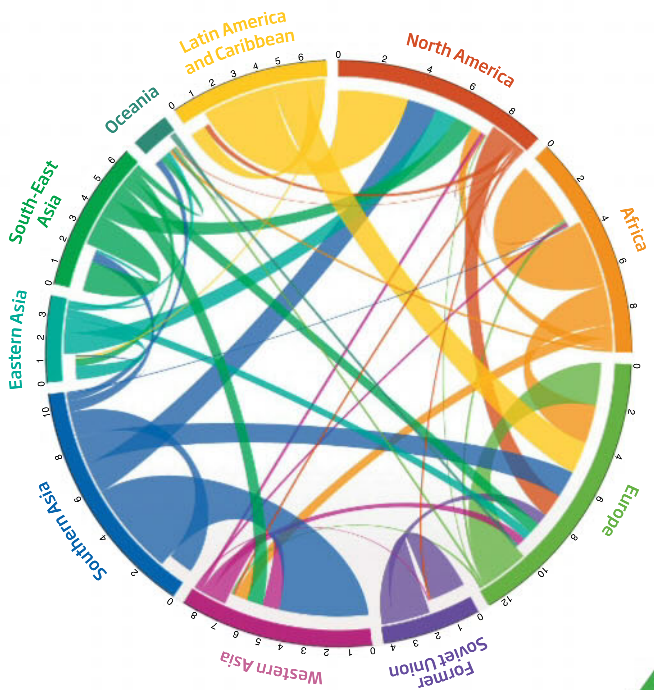
980 – Vikings migrate to Iceland, Greenland and Newfoundland

1200 – Migrants from northern Mexico establish Aztec empire

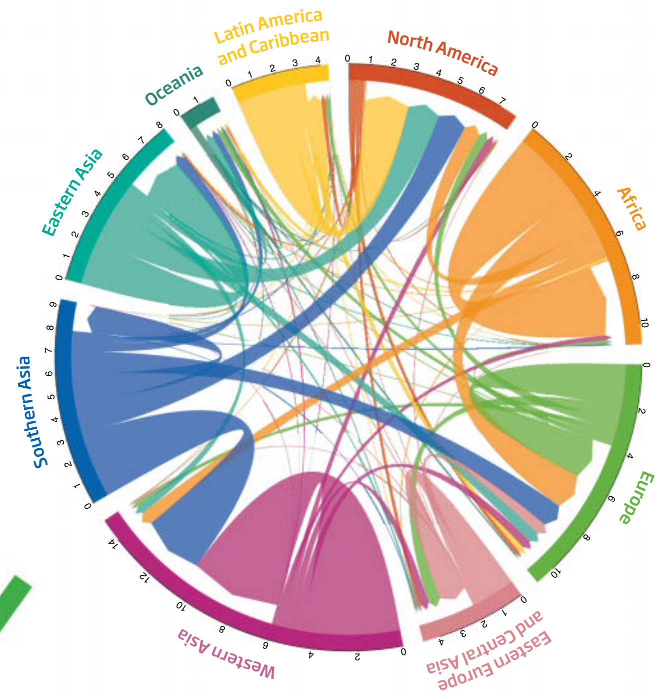
1220 – Mongols spread across Eurasia

1492 – Columbus reaches the Caribbean, sparking a wave of migration from Europe to the Americas

2005-2010



2010-2015



GUY ABEL, ASIAN DEMOGRAPHIC RESEARCH INSTITUTE, SHANGHAI UNIVERSITY, AND VIENNA INSTITUTE OF DEMOGRAPHY, AUSTRIAN ACADEMY OF SCIENCES, BIT.LY/NS_MIGRATIONDATA

That may come as a welcome surprise to many. But economics is not the whole story. If perceptions about jobs and wages were the only problem, you would expect anti-immigrant views to run high where jobs are scarce. Yet a 2013 study of 24 European countries found that people living in areas of high unemployment tended not to have negative views of migrants. So, what else are we worried about?

One major issue is a perceived threat to social cohesion. In particular, immigrants are often associated with crime. But here again the evidence doesn't stack up. In 2013, Brian Bell at the London School of Economics and his colleagues found no change in violent crime in Britain linked either to a wave of asylum seekers in the 1990s,

$356 billion

Boost to world GDP by 2025 if immigration increased workforces in high-income countries by 3 per cent

Source: World Bank

$39-$117 trillion

Estimated boost to world GDP if all barriers to migration fell

Source: Center for Global Development

or eastern EU migrants after 2004. The asylum seekers were associated with a small increases in property crime such as theft – boosting existing local crime rates some 2 per cent – perhaps because they were not allowed to work, suggest the authors. But areas where eastern Europeans settled had significantly less of any crime. Another study found that immigrants had no impact on crime in Italy. And immigrants in the US are much less likely to commit crimes and are imprisoned less often than native-born Americans. Tim Wadsworth of the University of Colorado has even suggested that a rise in immigration in the 1990s may have driven an overall drop in US crime rates since then.

Nevertheless, immigrants can put pressure on local communities. High ›

1520 – European ships start transporting slaves from West Africa to the New World

1820 – 2.6 million Europeans living in the New World, a quarter of them indentured servants

1840 – Start of the age of mass migration from Europe to North America and Australia

1847 – 1.5 million Irish people flee famine to Britain and America

1851 – 50,000 Chinese people migrate to join the California gold rush

AGE CONCERNS

As birth rates plummet in the developed world, migrants are keeping our economies afloat. They account for half of the increase in the US workforce since 2005, and 70 per cent in Europe. Even so, the number of people of working age supporting each retiree over 65 is falling. In 2000, this "dependency ratio" was 4:1 across the European Union. Today it is 3.5:1. And even with current levels of migration it is set to fall to 2 by 2050.

In 2000, the UN Department of Economic and Social Affairs ran a detailed simulation to see how many immigrants would be needed to support the population over 65 in developed countries. They found that with no migration, Europe's population is set to fall 17 per cent by 2050 – with a 30 per cent decrease in people of working age. To maintain overall numbers, the EU needs 850,000 immigrants per year – for comparison, the net migrant number from outside the EU in 2013 was 540,000. However, to keep the working-age population from falling, it needs nearly double that: 1.5 million a year. That would mean recent migrants and their children would account for 14 per cent of the UK population and over a third of Germany's and Japan's. Even then, the dependency ratio would be just over 2. The US fares better – current and expected migration kept its dependency ratio at 3.

"Migration might be the most relevant force to have an impact on the age distribution in Europe to 2050," says demographer Pablo Lattes, an author of the study. Germany, which has a shortfall of 1.8 million skilled workers, is keenly aware of this. Officials have been saying quietly at international meetings that this is why they have accepted so many of Europe's current wave of refugees. In 2000, the government tried to bring in 20,000 foreign high-tech workers, but this was met with strong opposition from the public. Germany may hope refugees will be harder for people to object to.

rates of arrival can temporarily strain schools, housing and other services. "That is what people tend to see," says Goldin. He says investment is required to mitigate these problems. "Governments need to manage the costs, which tend to be short-term and local," he says. That's a challenge, but it can be done. Bryan Caplan of George Mason University in Fairfax, Virginia, points out that since the 1990s, 155 million Chinese have moved from the countryside to cities for work. "This shows it's entirely possible to build new homes for hundreds of millions of migrants given a couple of decades."

China may be managing the biggest mass migration in history, but there's one problem it mostly doesn't face. Perceived threats to national identity often top natives' list of concerns about immigrants. It can even be an issue when such identities are relatively recent constructs. But countries with a clear ethnic identity and no recent history of significant immigration face the biggest problem, says Nelson. "It's tricky for Sweden, which went from essentially no immigrants to 16 per cent in half a generation," he says. And Denmark is another nation where anxiety over the loss of cultural homogeneity has been blamed for a backlash against immigrants.

Refugees account for fewer than 10 per cent of total migrant numbers

YURI KOZYREV/NOOR

Elsewhere, there has been a hardening of attitudes. Ellie Vasta of Macquarie University in Sydney, Australia, is trying to understand why Europe, which embraced multiculturalism in the 1970s, today calls for cohesion and nationalism, demanding that immigrants conform and testing them for "Britishness" or "Dutchness". She blames an increasing loss of cohesion in society due to "individualising" forces from mass media to the structure of work. As people rely more on their own resources, they have a longing for community. The presence of foreigners appears to disrupt this, creating a "desire to control differences", she says.

Research by Robert Putnam at Harvard University suggests this move away from multiculturalism could be problematic. He finds that increased diversity lowers "social capital" such as trust, cooperation and altruism. However, this can be overcome in societies that accommodate, rather than erase, diversity by creating "a new, broader sense of 'we'". In other words, success lies not in assimilation, but in adaptation on both sides. Canada has tried to achieve this by basing its national identity on immigration. Canadian prime minister Justin Trudeau told the World Economic Forum in Davos, Switzerland, in 2016 that "diversity is the engine of investment. It generates creativity that enriches the world."

This view is shared by complex systems analyst Scott Page at the University of Michigan, Ann Arbor. He argues that culturally diverse groups, from cities to research teams, consistently outperform less diverse groups due to "cognitive diversity" –

AD 1913 – European migration to the Americas peaks at 2.1 million per year

1917 – Bolshevik revolution in Russia displaces more than a million people into western Europe

1945 – Second world war displaces 30 million people

1947 – 18 million people move between India and Pakistan following partition exposure to disagreement and alternative ways of thinking. "Immigration provides a steady inflow of new ways of seeing and thinking – hence the great success of immigrants in business start-ups, science and the arts," he says. But more diversity means more complexity, and that requires more energy to maintain – investment in language skills, for example. The fact that immigrants have settled more successfully in some places than others suggests that specific efforts are required to get this right. Achieving broad agreement on core goals and principles is one, says Page.

We had better learn how to manage diversity soon because it's about to skyrocket in wealthy countries. As birth rates fall, there's a growing realisation that workers from abroad will be required to take up the slack (see "Age concerns", left). In addition, the fertility of incomers can stay higher than that of natives for several generations. In 2011, for the first time since mass European migration in the 19th century, more non-white than white babies were born in the US, mainly to recent Asian and Hispanic immigrants and their children. By 2050, white Americans will be a minority, says Bill Frey of the Brookings Institution in Washington DC. That's good news for the US, he adds, because it gives the country a younger workforce and outlook than its competitors in Europe and Japan.

Even if we finesse multiculturalism, there is a potential game changer looming on the horizon. Massive automation and use of robotics could make production less dependent on human labour (see "Nice work if you can get it", page 85). This "fourth industrial revolution" may see governments paying their citizens a guaranteed minimum wage independent of work. There has been little discussion of how this might affect a mobile global workforce.

SI BARBER/EYEVINE

However, some warn that cheap, automated production in wealthy countries could destroy export markets for poor countries. This would worsen unemployment and political instability – and also massively boost migration pressure.

One way to prepare for this would be to take a more coordinated and strategic approach to the global workforce. As it is, it's hard to track migration amidst a mess of non-standardised data and incompatible rules. Countries do not agree on who is a migrant. Even the EU has no common policy or information for matching people to jobs. Migrants are usually managed by foreign ministries, not labour ministries that understand the job market. "What could be of real value would be for governments, companies and trade unions to get together and look at where the labour shortages are, and how they could be filled, with natives or migrants," says Michelle Leighton, head of migration at the ILO.

Amazingly, says Goldin, there is no global body to oversee the movement of people. Governments belong to the International Organisation for Migration but it is not an official UN agency so cannot set common policy. Instead, each country jealously guards its borders while competing for workers. Goldin and others think there should be a UN agency managing migration in the global interest, rather than leaving it to nations with differing interests – and power. This, combined with real empirical understanding of the impacts of migration, might finally allow humanity to capitalise on the huge positive potential of its ancient penchant for moving. ■

Migrants make economies more resilient by doing jobs that natives won't or can't

4%

of the UK population were foreign citizens in 1993

8.5%

of the UK population were foreign citizens in 2014

Source: Migration Observatory, University of Oxford

1975
UN High Commissioner for Refugees counts 2.4 million refugees worldwide

2000
UNHCR counts 12.1 million refugees worldwide

2015
UNHCR counts 15.1 million refugees worldwide

Imagine there's no countries...

...it isn't hard to do, sang John Lennon. Actually, it is, argues **Debora MacKenzie**. Is there an alternative?

TRY, for a moment, to envisage a world without countries. Imagine a map not divided into neat, coloured patches, each with clear borders, governments, laws. Try to describe anything our society does – trade, travel, science, sport, maintaining peace and security – without mentioning countries. Try to describe yourself: you have a right to at least one nationality, and the right to change it, but not the right to have none.

Those coloured patches on the map may be democracies, dictatorships or too chaotic to be either, but virtually all claim to be one thing: a nation state, the sovereign territory of a "people" or nation who are entitled to self-determination within a self-governing state. So says the United Nations, which now numbers 193 of them.

And more and more peoples want their own state, from the Scots demanding another vote for independence to jihadis declaring a new state in the Middle East. Many of the big news stories of the day, from ongoing conflicts in Gaza to rows over immigration in Europe and the US, and the divisive Brexit vote, are linked to nation states in some way.

Even as our economies globalise, nation states remain the planet's premier political institution. Large votes for nationalist parties in elections across Europe, not to mention Britain's decision to go it alone, prove that nationalism remains alive – even as the EU tries to transcend it.

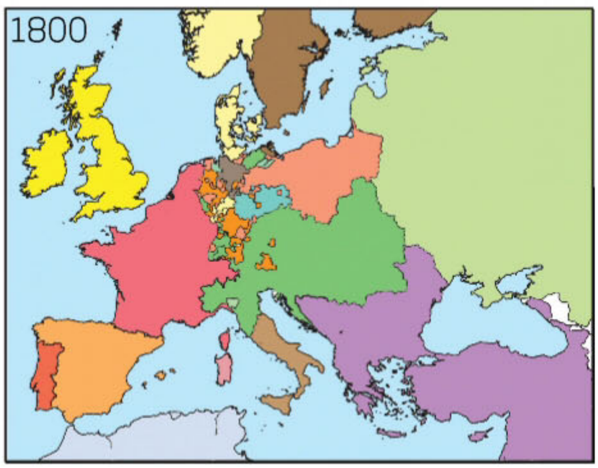
Yet there is a growing feeling among economists, political scientists and even national governments that the nation state is not necessarily the best scale on which to run our affairs. We must manage vital matters like food supply and climate on a global scale, yet national agendas repeatedly trump the global good. At a smaller scale, city and regional administrations often seem to serve people better than national governments.

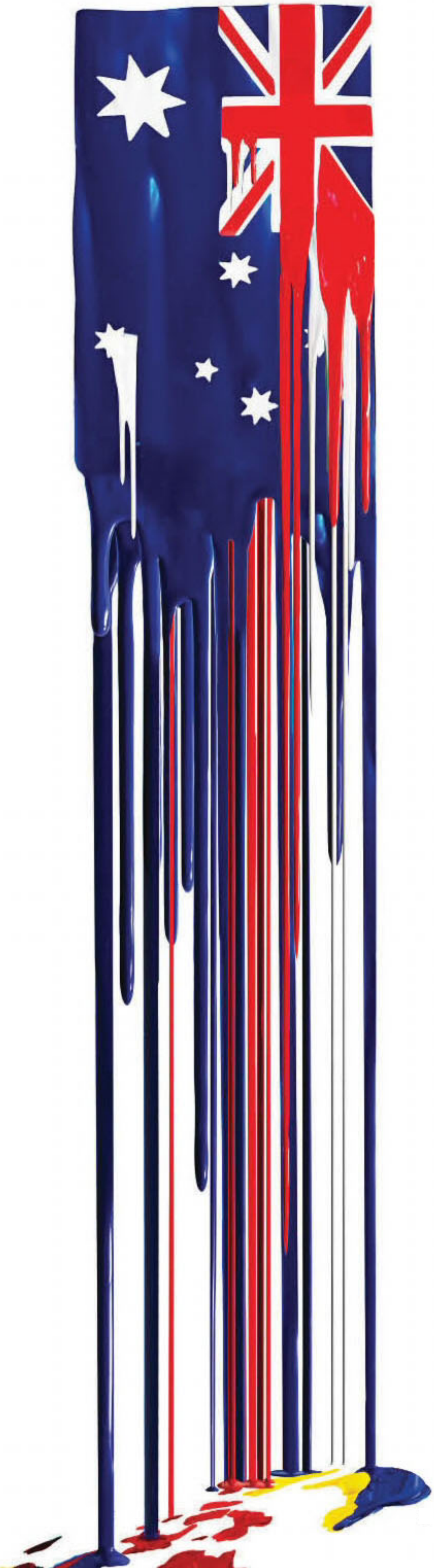
How, then, should we organise ourselves? Is the nation state a natural, inevitable institution? Or is it a dangerous ›

MAP: NORMAN KIRBY; PHOTOGRAPH: TATSURO NISHIMURA

Shifting sands

National borders can feel permanent and immutable – until you look at how they have changed over the past two centuries, especially in Europe, the cradle of the modern nation state





anachronism in a globalised world?

These are not normally scientific questions – but that is changing. Complexity theorists, social scientists and historians are addressing them using new techniques, and the answers are not always what you might expect. Far from timeless, the nation state is a recent phenomenon. And as complexity keeps rising, it is already mutating into novel political structures. Get set for neo-medievalism.

Before the late 18th century there were no real nation states, says John Breuilly of the London School of Economics. If you travelled across Europe, no one asked for your passport at borders; neither passports nor borders as we know them existed. People had ethnic and cultural identities, but these didn't really define the political entity they lived in.

That goes back to the anthropology, and psychology, of humanity's earliest politics. We started as wandering, extended families, then formed larger bands of hunter-gatherers, and then, around 10,000 years ago, settled in farming villages. Such alliances had adaptive advantages, as people cooperated to feed and defend themselves.

War and peace

But they also had limits. Robin Dunbar of the University of Oxford has shown that one individual can keep track of social interactions linking no more than around 150 people. Evidence for that includes studies of villages and army units through history, and the average tally of Facebook friends.

But there was one important reason to have more friends than that: war. "In small-scale societies, between 10 and 60 per cent of male deaths are attributable to warfare," says Peter Turchin of the University of Connecticut at Storrs. More allies meant a higher chance of survival.

Turchin has found that ancient Eurasian empires grew largest where fighting was fiercest, suggesting war was a major factor in political enlargement. Archaeologist Ian Morris of Stanford University in California reasons that as populations grew, people could no longer find empty lands where they could escape foes. The losers of battles were simply absorbed into the enemy's domain – so domains grew bigger.

How did they get past Dunbar's number? Humanity's universal answer was the invention of hierarchy. Several villages allied themselves under a chief; several chiefdoms banded together under a higher chief. To grow, these alliances added more villages, and

if necessary more layers of hierarchy.

Hierarchies meant leaders could coordinate large groups without anyone having to keep personal track of more than 150 people. In addition to their immediate circle, an individual interacted with one person from a higher level in the hierarchy, and typically eight people from lower levels, says Turchin.

These alliances continued to enlarge and increase in complexity in order to perform more kinds of collective actions, says Yaneer Bar-Yam of the New England Complex Systems Institute in Cambridge, Massachusetts. For a society to survive, its collective behaviour

"The view of the state as a necessary framework for politics does not stand up"

must be as complex as the challenges it faces – including competition from neighbours. If one group adopted a hierarchical society, its competitors also had to. Hierarchies spread and social complexity grew.

Larger hierarchies not only won more wars but also fed more people through economies of scale, which enabled technical and social innovations such as irrigation, food storage, record-keeping and a unifying religion. Cities, kingdoms and empires followed.

But these were not nation states. A conquered city or region could be subsumed into an empire regardless of its inhabitants' "national"

DOV MAKABAW/ALAMY

Emotional attachment to a nation state is a recent invention

identity. "The view of the state as a necessary framework for politics, as old as civilisation itself, does not stand up to scrutiny," says historian Andreas Osiander of the Humboldt University in Berlin.

One key point is that agrarian societies required little actual governing. Nine people in 10 were peasants who had to farm or starve, so were largely self-organising. Government intervened to take its cut, enforce basic criminal law and keep the peace within its undisputed territories. Otherwise its main role was to fight to keep those territories, or acquire more.

Even quite late on, rulers spent little time governing, says Osiander. In the 17th century Louis XIV of France had half a million troops fighting foreign wars but only 2000 keeping order at home. In the 18th century, the Dutch and Swiss needed no central government at all. Many eastern European immigrants arriving in the US in the 19th century could say what village they came from, but not what country: it didn't matter to them.

Before the modern era, says Breuilly, people defined themselves "vertically" by who their rulers were. There was little horizontal interaction between peasants beyond local markets. Whoever else the king ruled over, and whether those people were anything like oneself, was largely irrelevant.

Such systems are very different from today's states, which have well-defined boundaries filled with citizens. In a system of vertical loyalties, says Breuilly, power peaks where the overlord lives and peters out in frontier territories that shade into neighbouring regions. Ancient empires are coloured on modern maps as if they had firm borders, but they didn't. Moreover, people and territories often came under different jurisdictions for different purposes.

Simple societies

Such loose control, says Bar-Yam, meant pre-modern political units were only capable of scaling up a few simple actions such as growing food, fighting battles, collecting tribute and keeping order. Some, like the Roman Empire, did this on a very large scale. But complexity – the different actions society could collectively perform – was relatively low.

Complexity was limited by the energy a society could harness. For most of history that essentially meant human and animal labour. In the late Middle Ages, Europe harnessed more, especially water power. This boosted social complexity – trade increased, for example– requiring more government. A decentralised feudal system gave way to centralised monarchies with more power.

But these were still not nation states. Monarchies were defined by who ruled them, and rulers were defined by mutual recognition – or its converse, near-constant warfare. In Europe, however, as trade grew, monarchs discovered they could get more power from wealth than war.

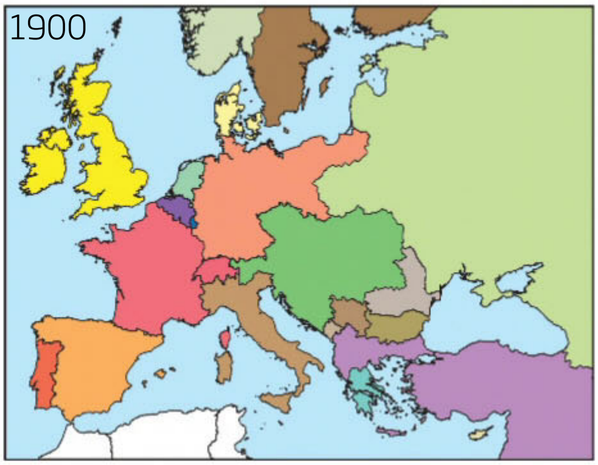
In 1648, Europe's Peace of Westphalia ended centuries of war by declaring existing kingdoms, empires and other polities "sovereign": none was to interfere in the internal affairs of others. This was a step towards modern states – but these sovereign entities were still not defined by their peoples' national identities. International law is said to date from the Westphalia treaty, yet the word "international" was not coined until 132 years later.

By then Europe had hit the tipping point of the industrial revolution. Harnessing vastly more energy from coal meant that complex behaviours performed by individuals, such as weaving, could be amplified, says Bar-Yam, producing much more complex collective behaviours.

This demanded a different kind of government. In 1776 and 1789, revolutions in the US and France created the first nation states, defined by the national identity of their citizens rather than the bloodlines of their rulers. According to one landmark history of the period, says Breuilly, "in 1800 almost nobody in France thought of themselves as French. By 1900 they all did." For various reasons, people in England had an earlier sense of "Englishness", he says, but it was not expressed as a nationalist ideology.

By 1918, with the dismemberment of Europe's last multinational empires such as the Habsburgs in the first world war, European state boundaries had been redrawn largely along cultural and linguistic lines. In Europe at least, the nation state was the new norm.

Part of the reason was a pragmatic adaptation of the scale of political control required to run an industrial economy. ›



Unlike farming, industry needs steel, coal and other resources which are not uniformly distributed, so many micro-states were no longer viable. Meanwhile, empires became unwieldy as they industrialised and needed more actual governing. So in 19th-century Europe, micro-states fused and empires split.

These new nation states were justified not merely as economically efficient, but as the fulfilment of their inhabitants' national destiny. A succession of historians has nonetheless concluded that it was the states that defined their respective nations, and not the other way around.

"At the revolution in 1789, half of France's residents did not speak French"

France, for example, was not the natural expression of a pre-existing French nation. At the revolution in 1789, half its residents did not speak French. In 1860, when Italy unified, only 2.5 per cent of residents regularly spoke standard Italian. Its leaders spoke French to each other. One famously said that, having created Italy, they now had to create Italians – a process many feel is still taking place.

Sociologist Siniša Maleševic of University College Dublin in Ireland believes that this "nation building" was a key step in the evolution of modern nation states. It required the creation of an ideology of nationalism that emotionally equated the nation with people's Dunbar circle of family and friends.

That in turn relied heavily on mass communication technologies. In an influential analysis, Benedict Anderson of Cornell University in New York described nations as "imagined" communities: they far outnumber our immediate circle and we will never meet them all, yet people will die for their nation as they would for their family.

Such nationalist feelings, he argued, arose after mass-market books standardised vernaculars and created linguistic communities. Newspapers allowed people to learn about events of common concern, creating a large "horizontal" community that was previously impossible. National identity was also deliberately fostered by state-funded mass education.

The key factor driving this ideological process, Maleševic says, was an underlying structural one: the development of far-reaching bureaucracies needed to run complex industrialised societies. For example, says Breuilly, in the 1880s Prussia became the first government to pay unemployment benefits. At first they were paid only in a worker's native village, where identification was not a problem. As people migrated for work, benefits were made available anywhere in Prussia. "It wasn't until then that they had to establish who a Prussian was," he says, and they needed bureaucracy to do it. Citizenship papers, censuses and policed borders followed.

That meant hierarchical control structures ballooned, with more layers of middle management. Such bureaucracy was what really brought people together in nation-sized units, argues Maleševic. But not by design: it emerged out of the behaviour of complex hierarchical systems. As people do more kinds of activities, says Bar-Yam, the control structure of their society inevitably becomes denser.

In the emerging nation state, that translates into more bureaucrats per head of population. Being tied into such close bureaucratic control also encouraged people to feel personal ties with the state, especially as ties to church and village declined. As governments exerted greater control, people got more rights, such as voting, in return. For the first time, people felt the state was theirs.

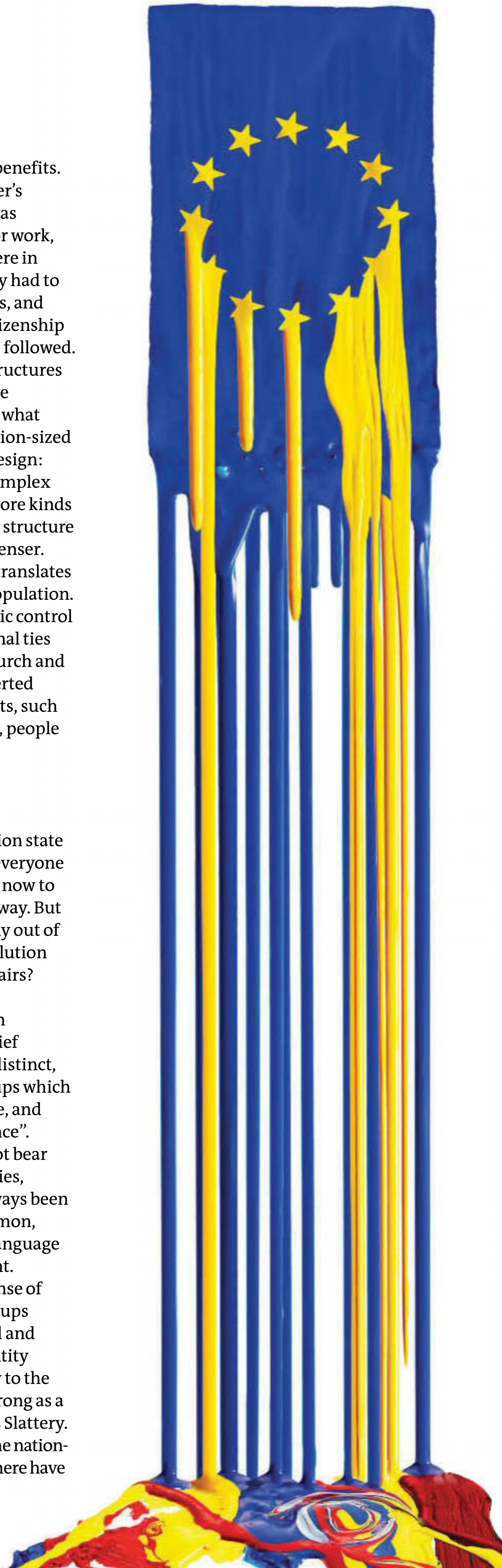
Natural state of affairs?

Once Europe had established the nation state model and prospered, says Breuilly, everyone wanted to follow suit. In fact it's hard now to imagine that there could be another way. But is a structure that grew spontaneously out of the complexity of the industrial revolution really the best way to manage our affairs?

According to Brian Slattery of York University in Toronto, Canada, nation states still thrive on a widely held belief that "the world is naturally made of distinct, homogeneous national or tribal groups which occupy separate portions of the globe, and claim most people's primary allegiance". But anthropological research does not bear that out, he says. Even in tribal societies, ethnic and cultural pluralism has always been widespread. Multilingualism is common, cultures shade into each other, and language and cultural groups are not congruent.

Moreover, people always have a sense of belonging to numerous different groups based on region, culture, background and more. "The claim that a person's identity and well-being is tied in a central way to the well-being of the national group is wrong as a simple matter of historical fact," says Slattery.

Perhaps it is no wonder, then, that the nation-state model fails so often: since 1960 there have



IVA ZIMOVA/PANOS PICTURES

Even today, conflicts usually revolve around issues of nationhood

been more than 180 civil wars worldwide.

Such conflicts are often blamed on ethnic or sectarian tensions. Failed states, such as Syria, are typically riven by violence along such lines. According to the idea that nation states should contain only one nation, such failures have often been blamed on the colonial legacy of bundling together many peoples within unnatural boundaries.

But for every Syria or Iraq there is a Singapore, Malaysia or Tanzania, getting along okay despite having several "national" groups. Immigrant states in Australia and the Americas, meanwhile, forged single nations out of massive initial diversity.

What makes the difference? It turns out that while ethnicity and language are important, what really matters is bureaucracy. This is clear in the varying fates of the independent states that emerged as Europe's overseas empires fell apart after the second world war.

According to the mythology of nationalism, all they needed was a territory, a flag, a national government and UN recognition. In fact what they really needed was complex bureaucracy.

Some former colonies that had one became stable democracies, notably India. Others did not, especially those such as the former Belgian Congo, whose colonial rulers had merely extracted resources. Many of these became dictatorships, which require a much simpler bureaucracy than democracies.

Dictatorships exacerbate ethnic strife because their institutions do not promote citizens' identification with the nation. In such situations, people fall back on trusted alliances based on kinship, which readily elicit Dunbar-like loyalties. Insecure governments allied to ethnic groups favour their own, while grievances among the disfavoured groups grow – and the resulting conflict can be fierce.

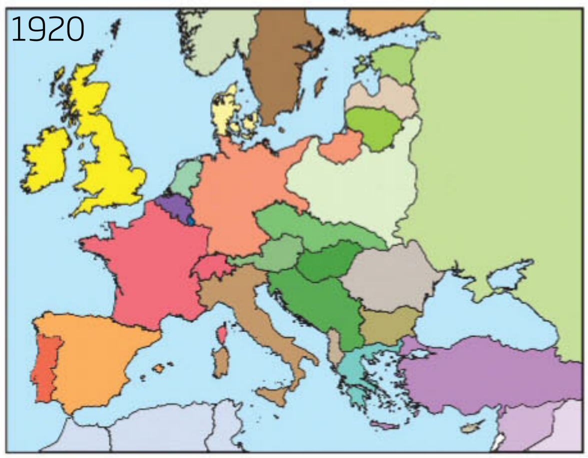
Recent research confirms that the problem is not ethnic diversity itself, but not enough official inclusiveness. Countries with little historic ethnic diversity are now having to learn that on the fly, as people migrate to find jobs within a globalised economy (see "On the road again", page 91).

How that pans out may depend on whether people self-segregate. Humans like being around people like themselves, and ethnic enclaves can be the result. Jennifer Neal of Michigan State University in East Lansing has used agent-based modelling to look at the effect of this in city neighbourhoods. Her work suggests that enclaves promote social cohesion, but at the cost of decreasing tolerance between groups. Small enclaves in close proximity may be the solution.

But at what scale? Bar-Yam says communities where people are well mixed – such as in peaceable Singapore, where enclaves are actively discouraged – tend not to have ethnic strife. Larger enclaves can also foster stability. Using mathematical models to correlate the size of enclaves with the incidences of ethnic strife in India, Switzerland and the former Yugoslavia, he found that enclaves 56 kilometres or more wide make for peaceful coexistence – especially if they are separated by natural geographical barriers,

Switzerland's 26 cantons, for example, which have different languages and religions, meet Bar-Yam's spatial stability test – except one. A French-speaking enclave in German-speaking Berne experienced the only major unrest in recent Swiss history. It was resolved by making it a separate canton, Jura, which meets the criteria.

Again, though, ethnicity and language are only part of the story. Lars-Erik Cederman of the Swiss Federal Institute of Technology in Zurich argues that Swiss cantons have achieved peace not ›



MARK HENLEY/PANOS PICTURES

Multi-ethnic states such as Malaysia can get along quite well

by geographical adjustment of frontiers, but by political arrangements giving cantons considerable autonomy and a part in collective decisions.

Similarly, having analysed civil wars since 1960, Cederman finds that strife is indeed more likely in countries that are more ethnically diverse. But careful analysis confirms that trouble arises not from diversity alone, but when certain groups are systematically excluded from power.

Governments with ethnicity-based politics were especially vulnerable. The US set up just such a government in Iraq after the 2003 invasion. Exclusion of Sunni by Shiites led to insurgents declaring a Sunni state in occupied territory in Iraq and Syria. True to nation-state mythology, it rejects the colonial boundaries of Iraq and Syria, as they force dissimilar "nations" together.

Ethnic cleansing

Yet the solution cannot be imposing ethnic uniformity. Historically, so-called ethnic cleansing has been uniquely bloody, and "national" uniformity is no guarantee of harmony. In any case, there is no good definition of an ethnic group. Many people's ethnicities are mixed and change with the political weather: the numbers who claimed to be German in the Czech Sudetenland territory annexed by Hitler changed dramatically before and after the war. Russian claims to Russian-speakers in eastern Ukraine may be equally flimsy.

Both Bar-Yam's and Cederman's research suggests one answer to diversity within nation states: devolve power to local communities, as multicultural states such as Belgium and Canada have done.

"We need a conception of the state as a place where multiple affiliations and languages and religions may be safe and flourish," says Slattery. "That is the ideal Tanzania has embraced and it seems to be working reasonably well." Tanzania has more than 120 ethnic groups and about 100 languages.

In the end, what may matter more than ethnicity, language or religion is economic scale. The scale needed to prosper may have changed with technology – tiny Estonia is a high-tech winner – but a small state may still not pack enough economic power to compete.

That is one reason why Estonia is such an enthusiastic member of the European Union. After the devastating wars in the 20th century, European countries tried to prevent further war by integrating their basic industries. That project, which became the European Union, now primarily offers member states profitable economies of scale, through manufacturing and selling in the world's largest single market.

What the EU fails to inspire is nationalist-style allegiance – which Maleševic thinks nowadays relies on the "banal" nationalism of sport, anthems, TV news programmes, even song contests. That means Europeans' allegiances are no longer identified with the political unit that handles much of their government.

Ironically, says Jan Zielonka of the University of Oxford, the EU has saved Europe's nation states, which are now too small to compete individually. He argues that the call by nationalist parties to "take back power from Brussels", which was successful during the Brexit campaign in the UK, would lead to weaker countries, not stronger ones.

He sees a different problem. Nation states grew out of the complex hierarchies of the industrial revolution. The EU adds another layer of hierarchy – but without enough underlying integration to wield decisive power. It lacks both of Maleševic's necessary conditions: nationalist ideology and pervasive integrating bureaucracy.

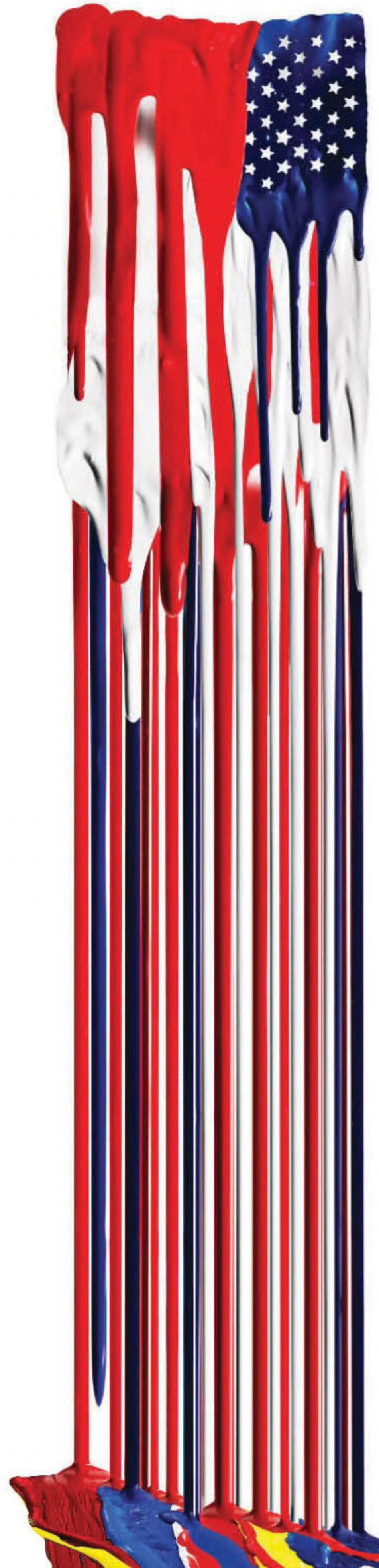
Even so, the EU may point the way to what a post-nation-state world will look like.

Zielonka agrees that further integration of Europe's governing systems is needed as economies become more interdependent. But he says Europe's often-paralysed hierarchy cannot achieve this. Instead he sees the replacement of hierarchy by networks of cities, regions and even non-governmental organisations. Sound familiar? Proponents call it neo-medievalism.

"The future structure and exercise of political power will resemble the medieval model more than the Westphalian one," Zielonka says. "The latter is about concentration of power, sovereignty and clear-cut identity." Neo-medievalism, on the other hand, means overlapping authorities, divided sovereignty, multiple identities and governing institutions, and fuzzy borders.

Anne-Marie Slaughter, head of the New American think tank and a former US assistant secretary of state, also sees hierarchies giving way to global networks primarily of experts and bureaucrats from nation states. For example, governments now work more through flexible networks such as the G7 (or 8, or 20) to manage global problems than through the UN hierarchy.

Ian Goldin of the Oxford Martin School at the University of Oxford, which analyses global problems, thinks such networks must emerge. He believes existing institutions such as UN agencies and the World Bank are structurally unable to deal with problems that emerge from global interrelatedness, such as economic instability, pandemics, climate



change and cybersecurity – partly because they are hierarchies of member states which themselves cannot deal with these global problems. He quotes Slaughter: "Networked problems require a networked response."

Again, the underlying behaviour of systems and the limits of the human brain explain why. Bar-Yam notes that in any hierarchy, the person at the top has to be able to get their head around the whole system. When systems are too complex for one human mind to grasp, he argues that they must evolve from hierarchies into networks where no one person is in charge.

"The future exercise of power will resemble the medieval model"

Where does this leave nation states? " They remain the main containers of power in the world," says Breuilly. And we need their power to maintain the personal security that has permitted human violence to decline to all-time lows.

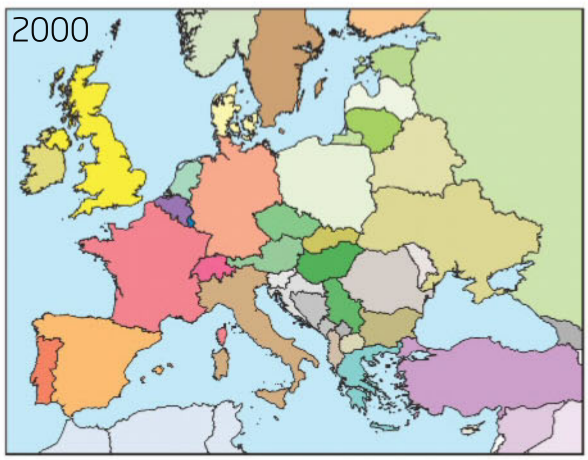
Moreover, says Dani Rodrik of the John F. Kennedy School of Government at Harvard University, the very globalised economy that is allowing these networks to emerge needs something or somebody to write and enforce the rules. Nation states are currently the only entities powerful enough to do this.

Yet their limitations are clear, both in solving global problems and resolving local conflicts. One solution may be to pay more attention to the scale of government. Known as subsidiarity, this is a basic principle of the EU: the idea that government should act at the level where it is most effective, with local government for local problems and higher powers at higher scales. There is empirical evidence that it works: social and ecological systems can be better governed when their users self-organise than when they are run by outside leaders.

However, it is hard to see how our political system can evolve coherently in that direction. Nation states could get in the way of both devolution to local control and networking to achieve global goals. With climate change, it is arguable that they already have (see "Living with climate change", page 108).

There is an alternative to evolving towards a globalised world of interlocking networks, neo-medieval or not, and that is collapse. "Most hierarchical systems tend to become top-heavy, expensive and incapable of responding to change," says Marten Scheffer of Wageningen University in the Netherlands. "The resulting tension may be released through partial collapse." For nation states, that could mean anything from the renewed pre-eminence of cities to Iraq-style anarchy. An uncertain prospect, but there is an upside. Collapse, say some, is the creative destruction that allows new structures to emerge.

Like it or not, our societies may already be undergoing this transition. We cannot yet imagine there are no countries. But recognising that they were temporary solutions to specific historical situations can only help us manage a transition to whatever we need next. Whether or not our nations endure, the structures through which we govern our affairs are due for a change. Time to start imagining. ■



LIVING WITH CLIMATE CHANGE

In December 2015, 195 nations gathered at the Paris climate talks and agreed to take action to limit global warming to 2°C.

The Paris Agreement was dealt a blow in June 2017 when Donald Trump announced his intent to take the US out of it. But the world can and will move on.

Over these eight pages, we look at the reality of climate change, and our response to it: what we're doing, what more we must do, and what the future holds in a warmer world

1

LIVING WITH CLIMATE CHANGE

THE NEW NORMAL

IT MAY not be immediately obvious, but the world outside your window is already a changed one. Since the industrial revolution, global temperatures have risen by about 1°C, which has had an impact at even the largest scales. For example, melting glaciers in Greenland are shifting the distribution of water on Earth, and nudging the planet's axis. As a result, the position of the North Pole has moved eastwards by more than 1 metre since 2005. An upshot of this is that Earth will spin faster and, by 2200, days could be 0.12 milliseconds shorter.

Earth's tilt is unlikely to affect your life or even that of your children, but other changes are happening closer to home. In the UK, for instance, spring is beginning about two weeks earlier on average than it did half a century ago, and autumn a week later. In the seas, many animals have shifted their range hundreds of kilometres polewards. On land, we are seeing similar shifts, but it can be much harder for terrestrial wildlife to move, not least because of roads and cities.

Another subtle change is that nights are warming faster than days. Night-time is a chance for heat to escape back out into space, but the extra greenhouse gases in the atmosphere are trapping ever more of it. This is particularly bad news during heatwaves: if our bodies don't get a chance to cool down at night, it is harder to cope with the heat of the day.

Not only are heatwaves more difficult to deal with in our changing world, they are also more frequent and more extreme. The 2003 European heatwave killed 70,000 people, many of them elderly or young children – groups who are less able to regulate their core temperature. A 2004 study showed that global warming has at least doubled the risk of such a weather event occurring.

Heatwaves are just one example of extreme weather affected by global warming. Of course, there have always been floods, storms and droughts, but now more rain can fall because a hotter atmosphere can hold more moisture, storms are more violent because there is more energy to power them and droughts can be more severe because water is evaporating faster.

Researchers are working on systems that would give an indication of how likely it is that extreme weather events are a result of climate change, in near-real time. But all weather events are affected to some degree.

Satellite studies show that rising carbon dioxide levels are also making the planet greener, particularly in dry areas. The results are complex, and not always good. In Australia, for instance, the extra vegetation is sucking up more water and reducing river flows by as much as a third.

All the changes due to climate change are overlaid on top of natural swings in the weather and climate, making it difficult to tell what is down to climate change. But long-term studies leave no room for doubt: as temperatures slowly climb, "normal" is a constantly shifting state. **Michael Le Page**

MARCAUX/GETTY

LIVING WITH CLIMATE CHANGE

CAN WE LIMIT WARMING TO 2 °C?

2

AT THE core of the Paris climate change agreement is the aspiration to "[hold] the increase in the global average temperature to well below 2 °C above pre-industrial levels". At current rates of greenhouse gas emissions, we have 20 years before such a rise is inevitable. To avoid it, we need emissions to peak as soon as possible – preferably by 2020 – before making their way to zero by about 2070.

There are some grounds for optimism: energy and industry emissions may already be peaking as the world moves away from the dirtiest of fossil fuels, coal (see "Have we reached peak emissions?", page 112). But this needs to be seen in context. We are still emitting almost 42 gigatonnes of carbon dioxide each year. Change is not yet happening fast enough or on a large-enough scale to meet the world's growing energy demand.

Besides, closing coal mines and investing in renewables for electricity generation is the easy part. Generating electricity accounts for only a quarter of global greenhouse gas emissions, with emissions from agriculture, forestry, industry and transportation making up the rest.

Oil, the primary fuel for transport, is particularly difficult to replace. Cars and buses can be made to run on electricity, but powering planes will require the large-scale development of renewable, sustainable jet fuel. The current global production, mainly biofuel made from fermenting crops, is minuscule compared even with the annual US consumption of 90 million litres of jet fuel.

Of course, we could choose to fly less and cut down on other behaviours that have high carbon footprints, such as eating meat. The question is whether we will – and how soon. The aspiration set out in the Paris agreement is just that: an aspiration. Based on the concrete commitments for emissions reductions by 2030 made by individual countries, it seems likely we will continue to overshoot the trajectory necessary to hit zero in 2070 by some way (see diagram, right).

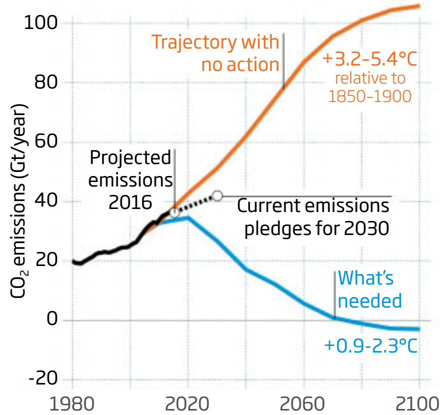
This means that not only will we have to wean ourselves entirely off fossil fuels, but we may have to suck carbon dioxide out of the atmosphere on an industrial scale to hit the 2 °C target. Assuming industry and agriculture continue to produce greenhouse gas emissions in the near term, we would have to suck about 600 gigatonnes of carbon out of the atmosphere this century.

The main contender for achieving this is bioenergy with carbon capture and storage (BECCS), in which plants are grown and then burned to produce energy. The CO_2 they sucked out of the atmosphere while growing is released when they are burned but captured and shoved underground before it can escape back into the atmosphere.

On the upside, the biofuel could power aircraft. The catch is that to have a meaningful effect, BECCS would have to happen on a massive scale, with thousands of industrial facilities across the globe. Not only is the technology untested on those scales, but we don't have enough spare land to grow the crops while continuing to feed ourselves. Staying below 2 °C would mean planting crops solely for the purpose of CO_2 removal on around 500 million hectares – equal to a third of Earth's arable land surface or half the area of the US – competing with ecosystems and food crops. In short, the consequences of negative emissions technology could be worse than overshooting 2 °C.

To stay below 2 °C, we need a radical change in commitment, where nations make climate change their top priority and the transfer to renewables is accelerated. With our present policies, we are on track to reach 3.6 °C of warming. Even with the promises made by world leaders in Paris, which have yet to materialise, we will exceed 3 °C. Then again, Donald Trump may be the spur the world needs to galvanise action (see "What has Trump changed?", page 114). Olive Heffernan

CO_2 emissions must start decreasing soon and **reach zero by 2070** to hit the Paris target of 2°C global warming



SOURCE: GCP/CDIAC/IPCC/FUSS ET AL 2014/ROGELJ ET AL 2016

BACK TO BASICS

CLIMATE IN NUMBERS

1.1 °C

CURRENT GLOBAL WARMING ABOVE PRE-INDUSTRIAL TEMPERATURES

SOURCE: wmo, bit.Ly/2sahwfc

2.0 °C

MAXIMUM WARMING ASPIRED TO IN PARIS CLIMATE DEAL

UNFCCC, bit.ly/2szPQfH

3.6 °C

LIKELY WARMING ON CURRENT GOVERNMENT PLEDGES

Climate Action Tracker, bit.ly/2snPKaK

41.9 GIGATONNES

GLOBAL ANNUAL CO_2 EMISSIONS, INCLUDING DEFORESTATION

GCP, bit.ly/2t1tJMF

2020

WHEN ANNUAL EMISSIONS MUST PEAK FOR US TO HIT 2 °C

MetOffice, bit.ly/2snCR0I

2.4%

ANNUAL INDUSTRIAL CO_2 EMISSIONS GROWTH 2004-13

GCP, bit.ly/2t1tJMF

<1%

ANNUAL INDUSTRIAL CO_2 EMISSIONS GROWTH SINCE 2015

GCP, bit.ly/2t1tJMF

10%

PROPORTION THE WORLD'S ELECTRICITY FROM RENEWABLE SOURCES IN 2016

IEA, bit.ly/2t1I6Sa

LIVING WITH CLIMATE CHANGE

3 IS RUNAWAY CHANGE LIKELY?

NASA

IN ANTARCTICA, the giant Thwaites glacier is in fast retreat. Ditto the Jakobshavn and Zachariae Isstrom glaciers in Greenland. Climate researchers worry they may have passed their tipping points, beyond which change feeds on itself and cannot be stopped. If the three glaciers melted fully, they alone would commit the world to more than 2 metres of sea level rise.

The Intergovernmental Panel on Climate Change has warned that rapid warming could take key Earth systems beyond their tipping points, part of the worst-case scenario of climate change. Tim Lenton of Exeter University, UK, says a threshold was passed in 2007 when the summer melt of Arctic sea ice accelerated. The fear is that, with less ice cover, the ocean will absorb more heat and prevent winter refreeze, locking the system into perpetual decline.

This is not the only system at risk. Historically, as temperatures have gone up, changing amounts of sea ice at the poles have caused ocean circulation to flip. A new flip could lose us the Gulf Stream and collapse the Asian and West African monsoons, affecting the livelihoods of billions. So far, annual changes in sea ice have not disturbed overall global ocean circulation. But the Atlantic leg has already weakened markedly, which Lenton says may mean it is inching towards its tipping point.

The trouble is that although there is plenty of historical evidence that tipping points exist, we don't know what the warning signs are. Take the methane trapped in the permafrost of Siberia and North America, both of which are expected to thaw rapidly this century. Methane is a potent greenhouse gas. It doesn't stay in the atmosphere for as long as CO_2, but if a large volume of it were released in one go, this could trigger runaway warming. Right now, methane is escaping from the permafrost, but it is minimal and nobody knows whether this is new or normal.

We certainly shouldn't be reassured by the apparent lack of runaway change. Experiments with biological and chemical systems show that they become sluggish when they approach tipping points. They also show that tipping points are sometimes passed without immediate impact. Unfortunately, this is an area where the uncertainties are great and the risks much greater.
Fred Pearce

BACK TO BASICS

HOW WE KNOW CO_2 IS AT FAULT

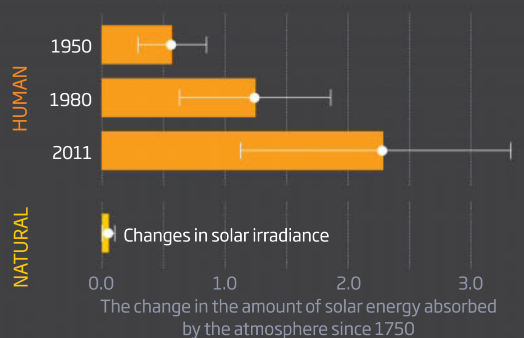
It makes up just 0.04 per cent of the atmosphere, but carbon dioxide is a small molecule with a big bite. We have known the mechanism behind this for more than 150 years.

In 1861, John Tyndall discovered that CO_2's three atoms vibrate when hit by certain photons. Photons from the sun pass straight through the atmosphere, unhindered by CO_2, but when they reach Earth's surface they bounce back as infrared photons - or heat, in other words.

Instead of passing through the atmosphere, these photons are absorbed by CO_2 and released again, only this time they fly off in random directions. As a result, more heat stays within the atmosphere than goes back out into space. This is the greenhouse effect, a double-edged sword that fosters life, but is also causing rapid global warming.

The other greenhouse gases act in the same way: methane, nitrous oxide, CFCs, ozone and water. Because there is so much of it, water vapour is the biggest heat trapper, but human activity is not directly increasing the atmosphere's water content. What we are doing is digging up fossil stores of carbon in the shape of coal, gas and oil. When we burn them to produce energy, we release CO_2 and methane. Some CO_2 is taken up by plants and returned to the soil, some is absorbed by the oceans, and some is stored in rocks that react with CO_2 when exposed to air. But these natural processes cannot keep up with the rate at which we are releasing greenhouse gases.

The final blow is CO_2's extraordinarily long life. It can stay in the atmosphere for thousands of years, so each molecule we produce adds to atmospheric concentrations and thickens Earth's blanket. "If we want the temperature to fall," says Ed Hawkins at the University of Reading, UK, "we will have to invent a way to remove CO_2 from the air on a huge scale. In the meantime we'll need to adapt to a warmer world." Julia Brown



SOURCE: IPCC/AR5

4

LIVING WITH CLIMATE CHANGE

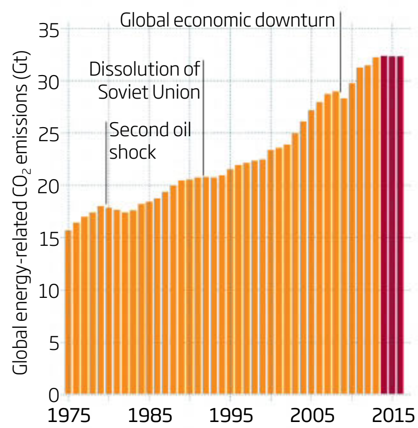
HAVE WE REACHED PEAK EMISSIONS?

HUMANITY'S appetite for energy has driven up the amount of greenhouse gases in the atmosphere. But here's the good news: last year, emissions from energy use stayed flat for the third year in a row. Overall emissions, including from industry, grew less than 1 per cent for the third year in a row.

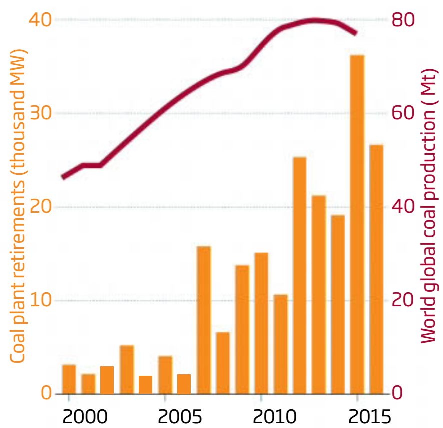
Energy emissions have stabilised or dropped at three other times in recent history, but only in economic downturns (see diagram, right). This time, the world economy is growing. At last count, 21 nations were seeing this "decoupling" of energy emissions and economic growth, including the UK, France, Germany and the US.

What's going on? For a start, king coal is dying. The biggest fall has come in the US, where this dirtiest of fossil fuels is being pushed out by gas and renewables. In China and India – growing economies with huge energy appetites – concern over air pollution are playing a part. Satellite images show that, in India for example, construction of some new coal power stations appears to have been abandoned.

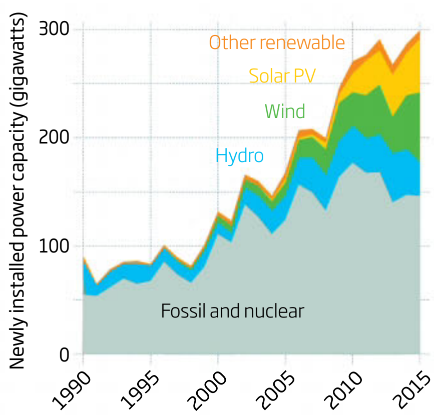
Renewables are also winning. Cheaper, more efficient turbines and photovoltaics mean that wind and solar energy cost the same or less to produce compared with fossil fuel power in more than 30 countries, even without government subsidies. According the World Economic Forum, this should extend to two-thirds of countries over the next few years. In 2016, the proportion of electricity from renewable sources other than large hydroelectric dams rose to 11.3 per cent, according to the UN Environment Programme, while renewables accounted for 55 per cent of the new capacity added worldwide. That in itself is turning into an economic win. According to the non-profit Environmental Defence Fund, solar and wind power in the US are creating jobs 12 times faster than the economy as a whole. Let's be honest: to stave off the worst of global warming, this should have been sorted at least a decade ago. But the trends do show that we can change our bad habits. **Catherine Brahic**



SOURCE: EA



SOURCE: EA (PRODUCTION) / ENDCOAL.ORG, BOOM AND BUST 2017 (RET REMENTS)



SOURCE: IEA/WEO

5

HITTING the agreed limit for global warming set in Paris is undoubtedly a tough ask, and it seems likely that the world will be 3 to 4 °C warmer by the end of the century (see "Can we limit warming to 2 °C?", page 110). That means loss and disruption. But starting to prepare now could mean that many people, perhaps even most, can thrive despite the rising temperature. That means more than building a few flood defences, however.

Take the challenge of feeding a growing population in a world where fertile land has been lost to sea level rise, extreme weather events are more common and vast tracts of land may be needed to grow biomass to burn as fuel.

Perhaps the highest priority here is to develop better crops that could feed more people using less land and fewer resources. Biologists are already developing crop plants that can capture more of the sun's energy, make their own nitrogen fertiliser and resist droughts, floods, salt, pests and diseases. But much more effort and money is needed.

It's a similar story with protecting our homes, businesses, roads and railways from extreme weather. In many cases we know what needs to be done, such as planting more trees on high ground or building bigger storm drains to deal with increased rainfall. We just need to work out how to pay for it.

Sea level rise is the most certain and potentially costly effect of climate change. It could be as much as 3 metres by 2100, and ultimately 20 metres or more. In 2016, the US allocated $48 million to move just 60 people from the shrinking Isle de Jean Charles in Louisiana. If the sea rises 1.8 metres, 13 million people would be displaced in the US, and hundreds of

LIVING WITH CLIMATE CHANGE

HOW TO COPE IN A WARMER WORLD

millions worldwide. There's a lot to be done to mitigate the risks, for instance banning building on land in areas such as Florida that could be lost to the sea in less than a century.

The most immediate danger posed by climate change, however, is financial. An estimated $882 billion of US property is threatened by sea level rise of 2 metres, for instance. Those properties could become worthless long before they flood if no one wants to buy them – a coastal property bubble waiting to happen.

That's nothing compared with the carbon bubble. An immense amount of wealth depends on the value of fossil fuel companies. It could be partly your wealth – most pension funds invest in fossil fuel stocks. "We are only just beginning to unravel how exposed everyone is to fossil fuel companies," says Anthony Hobley of the Carbon Tracker Initiative think tank.

Many investors believe that these companies are valuable because the firms can sell the coal, oil and gas in the reserves they own for a profit. But if we don't want to fry the planet, those reserves must stay in the ground, making them worthless. "Once climate change becomes a defining issue for financial stability, it may already be too late," warned the governor of the Bank of England, Mark Carney, in a 2015 speech.

Carney has since helped to set up a global initiative to get companies to assess their climate-related risks. The aim is to encourage companies with big risks to reduce them by, for instance, disinvesting from fossil fuels. It seems to be working: according to a 2016 report, some 60,000 individuals and 700 institutions, from companies to universities to pension funds, already plan to disinvest.
Michael Le Page

COCO ROBICHEAUX/ALAMY STOCK PHOTO

TALKING TO CLIMATE SCEPTICS

Some people reject the self-evident truths of climate change; others hold world views that don't easily find common ground with science. So how can they best be persuaded of the need for action?

Free market ideologues: "Saying climate change is the greatest threat to our world is a grab for global government by crazy catastrophists." This group may not deny basic climate science, but they deny its importance. They see calls to clamp down on emissions as a threat to the free market that drives capitalism.

Response: Ask why markets don't reflect the costs associated with climate change. Free markets need social and political stability, and so climate stability too. Big banks, insurance firms and oil companies have called for action on climate change. Government dilly-dallying is anathema to their bottom lines.

Christian ideologues : "The bible says humans have dominion over the Earth" and "it's all part of God's plan". Many Christians, particularly US evangelicals, say nature is for us to use as we see fit. It ties in with a political agenda opposed to collectivism, so reticent on issues that need collective action.

Response: Ask what happened to the strain of evangelism that sees "dominion" as meaning stewardship. Many other Christians say this gives us a moral imperative to tackle climate change. And climate change threatens the poorest most. Christian morals (and indeed the pope) say the fortunate should help those who are less fortunate.

Traditional conservatives: "The weather always changes, this is a green fad. Anyhow, the scientists don't agree. And none of my friends believe in it." This is an age-old drumbeat. During the latest UK general election, climate campaigners identified 18 MPs in the previous parliament who were publicly opposed to action on climate change – 16 were Conservatives.

Response: They can be persuaded with science. Point out that this is no fad. The greenhouse effect is 200-year-old physics. And climate models say more or less the same thing as chemist Svante Arrhenius calculated using pen and paper over a century ago.

The "we're doomed" brigade: "You can't change human behaviour, so you can't stop the emissions."This is not so much denialism as doomsday determinism, but it's odd how many people go from arguing that "there is no problem" to "there is nothing we can do about it anyway".

Response: Buy them a drink and explain how renewables are taking over. People do change their behaviour. The bar you are drinking in would have been full of smoke just 20 years ago. Fred Pearce

WHAT HAS TRUMP CHANGED?

DOMINICK REUTER/AFP/GETTY IMAGES

Donald Trump's pledge to pull the US out of the Paris climate agreement and renegotiate it has undoubtedly damaged the country's international standing. But its effect on global warming itself may be limited. By galvanising action in other countries, it might even end up producing a net win.

The US is the world's second-largest emitter after China, accounting for some 5 of the 42 gigatonnes of CO_2 emitted every year. Under the Paris agreement, nations set their own goals to cut emissions within the overall aim of limiting warming to 2 °C. The US had said it would reduce emissions by between 26 and 28 per cent in 2025 compared with 2005 levels .

It looks set to miss that target. But thanks to the shale gas revolution and the expansion of renewables, its emissions in 2014 were 7 per cent lower than in 2005. Such developments are likely to quicken. Barring huge additional subsidies, coal will continue to die a death. And states, cities and corporations have as much power to effect change as Washington does. In the wake of Trump's announcement, California – the world's sixth largest economy – and many more declared their commitment to the Paris goals. Trump is likely to be an obstacle, but not a bar, to progress.

The earliest the US can withdraw from the Paris agreement is 5 November 2020. Much can change domestically in the meantime, and other countries will march on towards a greener economy. In the wake of Trump's speech, many governments reaffirmed their commitment to the principles of the Paris deal. China and the EU agreed to cooperate more closely on climate, greener cities and transportation. No doubt the US will eventually join – but following, not leading. Catherine Brahic

6

LIVING WITH CLIMATE CHANGE

WILL WE HAVE TO GEOENGINEER?

DAVID KEITH is planning one of the first real-world geoengineering experiments next year. Only a decade ago, the idea that we might try to reverse global warming by artificially cooling the atmosphere seemed fanciful and dangerous. To most, it still does, but with the world on track for 3.6 °C of warming even if governments stick to all their promises (see "Can we limit warming to 2 °C?", page 110), researchers say we should at least understand what plan B looks like.

Keith and his colleagues at Harvard University intend to fly a balloon 20 kilometres up in the air, release bursts of tiny particles, and study how they deflect sunlight and interact with ozone. The particles will be benign – either ice, calcium carbonate or sulphur compounds that are naturally spewed by volcanoes – and only 100 grams will be released each time. Still, people have the right to feel jittery, says Keith. One concern is that even small-scale geoengineering experiments could send the message that there is a quick fix on the horizon. "The oil industry might say, 'See, there's no problem'."

As time drags on, the odds that we will reach for plan B get greater. "It seems likely to me that we will eventually have to resort to geoengineering because it's taking so long for a critical mass of people to take global warming seriously," says David Mitchell at the Desert Research Institute in Nevada. The risks are not to be sniffed at. Adding particles to the atmosphere to create a sunshade – a large-scale version of Keith's proposed experiment – could damage the ozone layer and models show that it is likely to mess up weather systems, causing severe droughts in some regions.

Another option is fertilising the oceans. The idea is that this would trigger algal blooms, which suck CO_2 out of the air during photosynthesis then drag it to the bottom of the ocean when they die. But large blooms could suffocate other marine species. Besides, there are doubts over how much of an impact it would have on the atmosphere.

A cheap, industrial-scale version, in which CO_2 is chemically removed from the atmosphere and stored underground, so far eludes us. Yet without such solutions, limiting warming to 2 °C may be out of reach. "People are trying to figure out a way to [remove CO_2]. But it's cheaper not to put it in the atmosphere in the first place," says Alan Robock at Rutgers University in New Jersey. Alice Klein

REUTERS/CHINA DAILY

LIVING WITH CLIMATE CHANGE

CAN I MAKE A DIFFERENCE?

7

"RIGHT now, your individual behaviour does matter," says Chris Jones. "Anyone can go carbon-neutral today. Better yet, you'll probably end up with more money in your pocket when you're done." Jones's group, the CoolClimate Network at the University of California, Berkeley, provides an online carbon footprint calculator. It estimates greenhouse gas emissions from transportation, housing, food, goods and services for any household in the US. Other groups provide similar calculators for many other countries. Despite individual and local differences, some broad generalities emerge. "Globally, the three main contributors to greenhouse gas footprints are cars, coal and cows," says Jones. And those three things are where individual choices can make the biggest difference.

For most households, especially in the US, transportation claims the biggest share of carbon emissions, about 30 per cent of the total. Most of this is from fuel, so buying a more fuel-efficient car can shrink your carbon footprint dramatically, especially if you currently drive a gas-guzzler. The other big carbon source, especially for the affluent, is air travel. "One flight will probably blow your carbon budget out of the water," says Stephen Cornelius, chief advisor for climate change at WWF-UK. Reducing air travel, by replacing business trips with teleconferencing for instance, can make a big difference. If you must fly, consider buying carbon offsets (see box, right) to balance the environmental impact of your flight.

Whether you can wring similar savings by improving your home's energy efficiency depends on where you live. In cold climates, better insulation can reduce the need to burn gas or oil for heat. But turning off lights, switching to LED bulbs and buying energy-efficient appliances only makes a difference if your electricity still comes from coal. If most of your electric power comes from renewable sources or nuclear plants, saving electricity has minimal effect on your carbon bottom line. In fact, the easiest way to green your home may be to buy your electricity from a renewable energy provider, says Jones.

MINT/PLAINPICTURE

One of the most effective places to reduce your carbon footprint is in the kitchen. Agriculture accounts for about a quarter of all greenhouse gas emissions, the vast majority from meat and dairy products. You can shrink your footprint dramatically simply by eating less meat, or none at all. The worst offender is beef, because cows belch large amounts of methane, a potent greenhouse gas. Chicken and pork avoid this problem, which may make them somewhat better choices, says Daniel Vennard of the World Resources Institute.

Because most households waste about a third of the food they buy – and because most affluent people eat more calories than they need – you can also reduce your food footprint simply by buying less and cooking it before it spoils. In contrast, buying local or organic food probably has very little effect, seeing as transportation makes up only a small fraction of emissions from agriculture, and the fact that organic farms generally have lower yields means more land must be ploughed, which releases carbon.

Anyone can take these steps today, and most of them actually cost less than the carbon-intensive behaviours they replace. According to Jones's calculations, these and similar measures can save US households thousands of dollars per year, while halving their carbon footprint. The money saved should be enough to buy carbon offsets to cover the remaining part of the footprint, thus making your household completely carbon neutral. That's a pretty good first step until government policies get more aggressive.

Bob Holmes ■

CARBON OFFSETS

People or firms with large carbon footprints can compensate for their sins by paying for someone to reduce emissions elsewhere. These carbon offsets, if done well, can be a good way to cope with unavoidable emissions from air travel or other costly behaviours. But to make a difference, they must pay for reductions that would not have happened without them. Certification by international oversight organisation Gold Standard helps ensure that offsets meet this requirement.

Offsets can be particularly helpful in mitigating the effect of air travel - easily the largest contributors to many people's footprints. "We purchase Gold Standard offsets for all the business travel we do," says Steve Kux, a climate policy analyst for the David Suzuki Foundation, an environmental group based in Canada.

TERENCE EDUARTE

POWER TO THE PEOPLE

The race to build a better battery will revolutionise the way we all use energy, finds Hal Hodson

WE ONLY notice them when they are about to run out. Icons glow red, warnings flash. The curse of modern mobility: our battery's about to give up.

It's a trivial, everyday annoyance. But the ramifications go far beyond just laptops and smartphones. Humans rely on two things to control their environment: information and energy. Shrinking transistors and the rise of microprocessors have given us immense control over the first: the capacity to store and manipulate data that we hold in the palms of our hands would have been inconceivable a generation ago.

But with energy, we're stuck in a rut. The development of electric cars stutters forwards thanks to the lack of ways to power them cheaply, efficiently and over long distances. And while we've made great strides in harnessing wind, wave and sun to generate cleaner electricity, again, the technology to store that juice lags badly behind.

Corporations and governments are pouring billions of dollars into improving existing battery technologies – with some success. But if we are to continue to compute and communicate with more freedom, while liberating ourselves from our dependence on fossil fuels, conventional thinking needs an overhaul. We're going to need a better battery.

The cutting edge of current energy storage technology is probably in your pocket right now – and 2 billion others around the world. The lithium-ion batteries that power most smartphones were born in the early 1990s as a quirk of the dying cassette tape industry. The rise of compact discs had Japanese company Sony casting around for something to do with old equipment for making tapes, says Jeff Chamberlain of Argonne National Laboratory in Chicago. Instead of coating the tape with magnetic film that could record data, they started coating it with goopy layers of an electrode that could store electric charge.

Fortuitous find

The first lithium-ion batteries contained rolls of these film electrodes, wound up in a cylinder like the spool of a cassette. They were instantly twice as good as anything else out there for compact energy storage. Existing nickel-cadmium and nickel-metal hydride batteries used chemical changes on the surface of two electrodes within them to shunt charge-bearing hydroxide ions and protons this way and that, and so charge and discharge. The new technology achieved the same by exchanging lithium ions, but slotting them into and out of nanoscale gaps within the material of the battery's electrodes in a chemical process called intercalation.

Because it is a light metal, lithium has a lot of charge-carrying ions for its weight, making for batteries that are smaller but more powerful. Lithium-ion batteries boomed from their serendipitous beginnings, driven first by the rise of personal electronic devices such as camcorders, and then mobile phones and laptops. Although they are still dwarfed in most respects by the bulky lead-acid batteries found in almost every car on the road today, in 2015, lithium-ion batteries accounted for around a third of the money spent on rechargeable batteries globally (see "Turn it on", page 118), and just under a sixth of the total energy stored, according to French research firm Avicenne.

At the same time, their performance has improved immensely: design tweaks have tripled the energy stored in a given volume since the technology was commercialised in 1991. Success has bred success, and lithium-ion batteries have found new and bigger ›

"Now lithium-ion batteries are a $15 billion business, big companies are taking notice"

applications, such as electric vehicles. For example, the Model S electric car designed by Tesla Motors, a company owned by serial entrepreneur Elon Musk, is powered by thousands of small lithium-ion batteries arrayed between the car's axles. It can go from zero to 100 kilometres an hour in about 3 seconds, and can travel about 430 kilometres on a single charge – although charging it can take many hours.

Tesla has no plans to stop there. Lithium-ion batteries are so important to the company that it has taken manufacturing into its own hands, building a "Gigafactory" just outside Reno, Nevada. By 2020, the company plans to produce as many lithium-ion batteries annually as the entire world produced in 2013 – enough for a fleet of 500,000 electric cars – and with a 30 per cent reduction in production cost per battery.

Although a cloak of secrecy surrounds Tesla's plans, achieving those goals probably means changing the way lithium batteries are made. For all their collective oomph, the thousands of batteries in a Model S are essentially just descendants of those first cylindrical film batteries. "For almost 25 years we've been using a suboptimal manufacturing process just because it was there," says Chamberlain. "Now that lithium-ion is a $15 billion business, big companies are taking notice."

And it's not just big companies. One small start-up, 24M, based in Cambridge, Massachusetts, has attracted more than $50 million of investment for an alternative manufacturing method. Instead of using ovens to dry the slurry that contains the battery's positive and negative electrodes, the company has found a way to keep the entire process wet, saving time, simplifying the design and boosting the energy density. The company also claims its method will cut the cost in half. "If it works, everyone else will do it instantly," says materials scientist George Crabtree of Argonne.

Too hot to touch

Crabtree's Argonne colleague Chamberlain is part of a consortium of companies and researchers with their own plans to improve lithium-ion batteries. As with Tesla, the details remain confidential, but the idea is to take coatings normally used to improve the longevity of artificial joints, and apply them to making batteries with dry electrodes. A number of processes are being tested, Chamberlain says, with the aim of finding what approach provides the most bang for the buck.

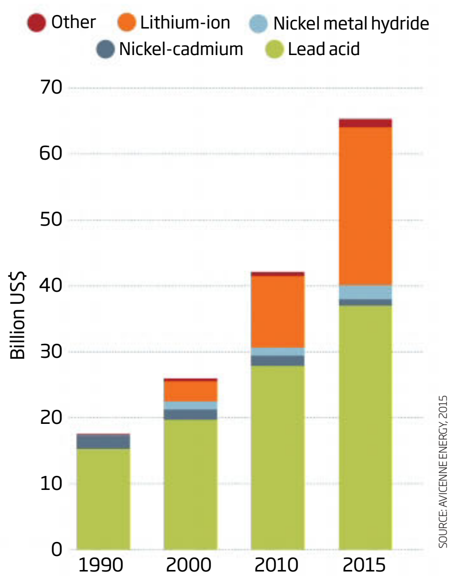

ROBIN HAMMOND/PANOS PICTURES

Bolivia's salt flats are the world's largest source of lithium

Lithium-ion technology has huge momentum, which means it's likely to form the backbone of our emerging energy-control infrastructure for some time. But it is not without problems. Lithium burns hotly, so batteries that contain it can be a fire hazard if their cells get overcharged. In 2006, Sony recalled 6 million laptop batteries that had been spontaneously combusting. In January 2013, batteries in one of Boeing's next generation Dreamliner aircraft caught fire while the craft was sitting empty at Boston's Logan Airport. Boeing has since updated its software, and modern systems generally have the problem under control – but a technology that requires tight oversight to avoid catching fire is hardly ideal.

TURN IT ON

Lithium's share of global investment in rechargeable batteries is booming



SOURCE: AVICENNE ENERGY, 2015

Other problems are less easily fixed. Lithium-ion batteries are approaching fundamental electrochemical limits on the density of energy they can store, while their cost is nearing its floor, too – something particularly problematic for larger-scale applications. "You might get it down by 30 per cent, but you're not going to get more than that," says Crabtree. "If you really want electric cars to compete with gasoline, you're going to need the next generation of batteries."

That means finding a new chemical basis for them, says Rachid Yazami of Nanyang Technological University in Singapore. Engineers have achieved incredible advances with lithium-ion batteries, he says – but not enough to meet increasing demands. "People want to charge electric cars very fast, and they want to run them for 500 miles," he says.

Lithium makes such an obvious choice for storing a lot of energy in a low mass that many alternative battery designs start off with the element, too. One is a lithium-sulphur battery that stores and releases energy by forming and breaking chemical bonds, instead of slotting ions into structural gaps. These batteries are less prone to catching fire, and although they're not yet commercially available, they have demonstrated energy densities three times those of the best lithium-ion batteries.

Dependence on lithium might not prove to be the best bet, however. For a start, plentiful as the element is, it's not always easy for international markets to get at. The largest identified resources are in Chile and Bolivia, which between them hold more than 40 per cent of the planet's known totals (see "On location", below). Here the lithium is found in a chloride brine, together with other metal salts, under the world's largest salt flats, the Salar de Uyuni. Processing lithium from brines may be cheaper and perhaps less environmentally damaging than mining it from rock. Bolivia has been reluctant to open up to any foreign mining companies, insisting instead that any lithium extracted there be used to make products – batteries and electric cars – within its borders for export. But in the face of intense interest, that stance has slowly softened.

Yazami is one researcher looking at alternative elements to build batteries from. He is coy about his own project, saying only that his lab is working on a system that uses materials far more common than lithium. "If I tell you I can develop a battery that can be charged in 15 seconds and last one week, you'd be happy," he says. "That's what we're doing."

In general, with an entire periodic table of elements for the taking, trial and error has long been the only way that better batteries have been found. "Traditional battery research is empiricism," says Chamberlain. "You use knowledge to investigate a material, then test that material."

That's changing, though, and we're beginning to use the spoils of the information revolution to spark a similar revolution in energy. Supercomputers are one such tool, making it possible to crunch through many combinations of elements in different proportions and optimise properties such as energy density and charging time. "We're using that data to find the needles in the haystack of possible materials," says Chamberlain.

One such system is the Electrolyte Genome, a program designed by researchers at Argonne in collaboration with Lawrence Berkeley National Laboratory in California that sifts through thousands of potential battery chemistries to find promising candidates. On the back of that, Argonne researchers recently built a prototype battery that uses magnesium ions instead of lithium to carry and store charge. Magnesium ions have two positive charges compared with lithium's one, doubling their capacity to store energy. However, they mostly appear only in bulky combinations with other elements – a problem the simulations are suggesting new ways to get around.

Another target for the computational approach are flow batteries, in which the storage material is all dissolved in solution, allowing for a particularly wide range of elemental mixes. Flow batteries have lower energy densities than other cutting-edge technologies, but are much cheaper. "We're looking for clever active elements or molecules to store and release the energy, organics that are really abundant, cheap and versatile," says Crabtree. So far 16,000 have been tested computationally.

Storing sunlight

Reducing the cost of reasonably good batteries might actually prove more important than improving their energy density, says Jeff Dahn of the University of Dalhousie in Nova Scotia, Canada. For instance, the Model S is a fantastic car in many ways, he says, and "the only issue with the thing is that it costs too much".

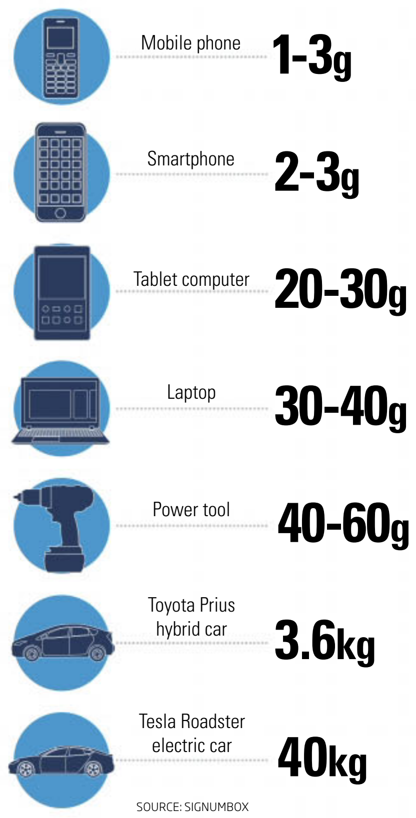
That certainly becomes true when we look beyond electric cars to an even more pressing energy problem: how to store electricity on the grid. Demand for electricity varies through day and night and through different seasons. Currently, the energy needed to meet peaks in demand is stored in the form of natural gas and coal. These fossil stores sit around in back-up power plants that ramp up when demand is high.

Renewable energy sources such as solar, wind and wave just compound this unpredictability: changes in local cloud cover, wind speed and the like produce irregular peaks and troughs that do not necessarily correspond with spikes in demand. "Shifting to solar and wind, we can't move the power up and down. We can't control it. We can't dictate when the wind blows or the sun shines," says Chamberlain. "So we have to store the sunlight somehow."

Our century-old model of electricity distribution is behind the problem, leaving us needing to use electricity as soon as it is ›

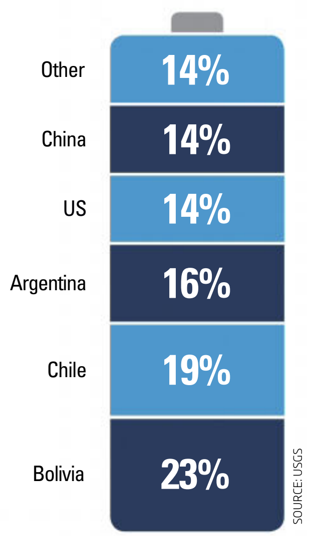
POWERED BY LITHIUM

The bigger the battery, the more lithium you need



ON LOCATION

Lithium resources are unequally distributed



"Good batteries are like needles in the haystack of all possible materials"

generated. On 20 March this year, for example, a solar eclipse knocked out two-thirds of Germany's solar generation capacity for about an hour. Grid operators, aware of the impending shortfall, spun up alternative generation from coal, gas and hydroelectric systems to meet it. With sufficient battery storage, though, the eclipse would have been a non-event from the start.

But batteries for grid storage will have to be very cheap to compete with fossil fuels as an on-demand energy source. "It's a very inexpensive way to meet our energy needs to drill a hole in the earth," says Chamberlain. "After years of engineering we've gotten to a place where that's a very profitable, effective way to get the energy we need."

Of all the ways to store energy, lithium-ion batteries are one of the most expensive. Yet, with some coaxing, grid storage is already happening using this technology. California signed a bill into law in 2014 that requires its energy companies to bring 1.3 gigawatts of

> **"Better batteries would enable the democratisation of electricity"**

storage to the grid by 2022 – about the same as the output of a large natural gas power station, and sufficient to cover about 1/40th of the state's average power needs at any one moment. Its purpose is to capture renewable energy at off-peak times, then feed it back into the grid when demand is high.

Responding to that initiative in 2015, utility company Southern California Edison commissioned Virginia-based start-up AES Energy Storage to build a gigantic lithium-ion array. Planned to be the largest battery ever built, it will be capable of delivering 100 megawatts of power for four hours, enough to supply 80,000 average US homes. "When you crunch the numbers, that procurement is half a billion dollars being spent on the grid, because of one state's policy," says Chamberlain.

Other companies such as Samsung and Siemens already offer products to back up grid power, but they are still expensive compared with fossil-fuel storage, and provide cover only for short periods. Doing the same thing across the entire grid is a mammoth task. "The scale is unimaginable," says Dahn, whose lab signed a five-year research contract with Tesla in June 2015. He calculates that storing the output of just his local utility company, Nova Scotia Power, for 24 hours would take the energy storage capacity of every battery made worldwide this year – and then half as much again.

In the end, the solution might lie on a smaller scale: giving everyone the power to store their own power (see "Help for heroes", page 121). Tesla is one company of several in this game: in 2015 it announced a device called the Powerwall, designed for homes and businesses. It uses the same batteries as electric cars to store energy, either from renewables or cheap night-time electricity, ready to be used during the day.

If such systems become commonplace, we might all become a little more aware of where our energy is coming from, and how our own behaviour affects its use and production, says energy researcher Philipp Grünewald of the University of Oxford. "Batteries would be a really helpful thing to give you a sense that you've got something you can trade," he says. He foresees a system where electricity providers put a small battery in customers houses for free, offering them cheaper rates in exchange for being able to manage that slice of energy storage for the good of the grid at large. That, however, would require buy-in from companies and consumers alike.

Chamberlain says it's hard to predict what changes the world will undergo if the battery revolution comes off – just as the consequences of the information revolution would have been hard to predict a decade or so ago. But he expects a similar empowerment as individuals gain the ability to produce, store and use electricity at will. "Batteries are a linchpin that would enable democratisation of electricity," he says.

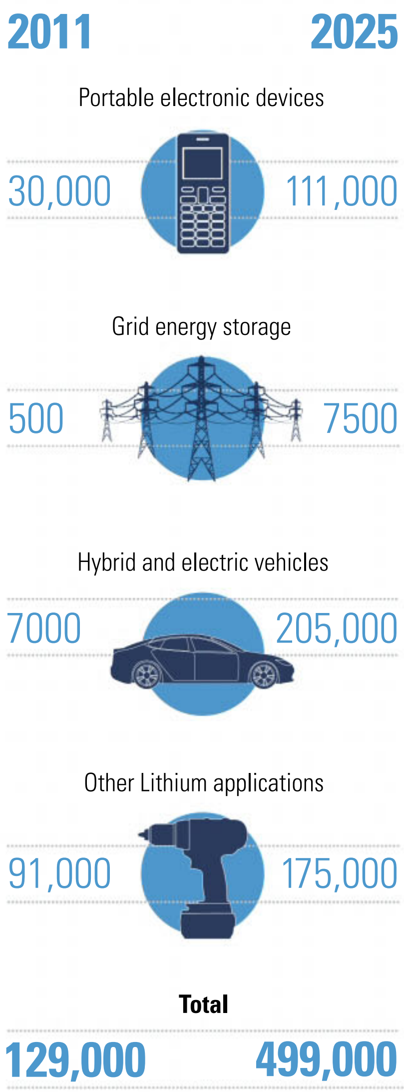
And as a growing global population demands ever more energy, the next generation of batteries can't come soon enough. "We are now at the edge of a new energy revolution," says Yazami. "We know fossil fuels are not a good solution. But without energy it's the end of the story."

Put that way, batteries become a technology of global importance, and not just to help avert dangerous climate change. "[The US] spends over a billion dollars a day to import energy in some form of petroleum," says David Howell of the Office of Vehicle Technologies in the US Department of Energy. "We ship a billion dollars offshore a day. That opens us to all kinds of vulnerabilities."

Grünewald agrees. "If we don't want fossil fuel based electricity in 2050, then storage will be absolutely vital," he says. "That starts today." ■

GLOBAL LITHIUM DEMAND

Tonnes of lithium carbonate equivalent used



HELP FOR HEROES

Teaming the battery with a nifty sidekick could transform how we use energy. Mark Harris reports

AGAINST the backdrop of the Nevada desert a gigantic factory is taking shape. Look at the artist's impressions of the finished building and you could mistake it for a Martian colony, its ranks of solar panels stark against the reddish dirt. But this is the Gigafactory, a sprawling edifice covering around 600,000 square metres. Here, electric car company Tesla has begun production of perhaps the single most important component of its vehicles: the battery.

A good rechargeable car battery will set you back around $10,000, for a product that is toxic, degrades substantially after a few years and must be carefully designed to avoid catastrophic overheating. The Gigafactory represents Tesla CEO Elon Musk's drive to make better batteries and so realise his dream of affordable electric cars.

Others are similarly exercised. In 2016, Samsung's woes with exploding batteries in its Galaxy Note 7 smartphone caused it to recall all the devices and cease production. "It will cost us so much it makes my heart ache," said Koh Dong-Jin, president of Samsung's mobile business. Better, cheaper batteries are top of the wish list for almost any technology that's not powered by fossil fuels.

Yet as Musk and others are finding, it's proving a long, hard road. Might there be a better way? That's the claim of researchers championing a long-overlooked device to store and supply energy. They think it could actually stand more of a chance of delivering the power we need, how we need it – and so revolutionise the way we use energy. Is it time to look beyond batteries?

Rechargeable batteries store energy by performing a reversible chemical reaction in which ions are stored in and flow between positive and negative electrodes. The right materials, such as the lithium compounds common to both Tesla and Samsung's batteries, can store lots of energy, but are slow to charge and discharge, and heat up when they do. In 2017 Samsung revealed what had caused the Note 7 fault, and it came down problems with the tiny separators used to keep components apart. The separators were simply too weak, creating a short circuit that heated other parts of the battery, causing a runaway reaction. Such safety concerns, plus the sheer cost of lithium batteries, have long had chemists casting around for something better.

But chemistry isn't the only way to store electric charge. In devices known as capacitors, energy is physically stored in an electric field between metal electrodes. Capacitors are sprinters to the battery's long-distance runner, charging and discharging in a blink, and doing this over and over again without their performance suffering. They are already used to power the flash on a camera.

But you can't run a car on a camera flash. A kilogram of petrol contains about 4000 Watt hours of useful energy, 30 times as much as the batteries in Tesla's current crop of vehicles. Traditional capacitors hold 1000 times less again, just 0.1 Watt hours per kg. If your car could drive 500 kilometres on a tank of petrol, it would run little more than 16 metres using the same weight of capacitors.

It's unthinkable, then, that a traditional capacitor could ever compete with a battery. But many have had that exact thought – even Musk. "If I were to make a prediction, I'd think there's a good chance that it is not batteries but capacitors" that will deliver a breakthrough, he said in 2011. In that reading, it's just a case of guiding the continuing evolution of the capacitor.

That evolution stretches back to 1966, when Robert Rightmire at Standard Oil of Ohio was part of a team considering the future of fuel storage. He knew that the charge a capacitor could store depended on the surface area of its electrodes. So why not make these surfaces more spongy, the better to cram in charge? He produced a capacitor where the electrodes were coated with thin layers of carbon chemically punctured with millions of tiny holes. This so-called activated carbon is typically used for jobs like decaffeinating coffee, and has an internal area about 100,000 times larger than its outside surface. And it worked. Rightmire's "supercapacitors" stored 10 times as much energy as traditional capacitors.

Ditch the coconuts

By the 1990s, small supercapacitors had become a commercial reality. They provided instant, short-lived back-up power to computers if the mains supply failed, so they could shut down safely. That's still a long way from powering a car. For a long time, not much changed. This was partly down to the curious source of that spongy carbon: coconuts.

"It's pure luck," says Aaron Feaver, chief technical officer at EnerG2, an energy storage company based in Seattle. "The coconut

"Badly designed batteries can explode, but that's the least of their problems"

didn't evolve to be an ultracapacitor electrode material, but it just happens to work pretty well." Leftover husks are heated to 600 °C in an oxygen-free oven to get rid of all elements except carbon, a process known as pyrolysis. The carbon is then treated with chemicals to etch in the tiny pores.

Coconuts were so cheap and convenient a source of carbon that no one thought much about other possibilities. At some point in the late 1990s, supercapacitors were rebranded "ultracapacitors", but the principle remained the same.

And they've continued to find new uses. Some wind turbine companies use them as an emergency alternative to batteries. Turbine blades need to be constantly adjusted to face the wind. If their electricity supply fails, the blades must quickly return to a neutral position to avoid strong gusts damaging or even destroying the turbine. That calls for a short power splurge – what ultracapacitors excel at. Plodding batteries are heavier and eventually need replacing. "Once you've put something into a turbine you're not going to want to go up and service it. You just want to forget it," says Kim McGrath from Maxwell Technologies, an ultracapacitor manufacturer.

That special ability of ultracapacitors to provide a short zip of power is useful in other places too. In China, fleets of hybrid diesel buses are equipped with ultracapacitors that charge up swiftly from regenerative braking systems, and later accelerate the bus until the diesel engine can take over.

Meanwhile, material innovations suggest ways to store more juice in capacitors. In the mid-2000s Joel Schindall, John Kassakian and Riccardo Signorelli at the Massachusetts Institute of Technology began to explore whether other types of carbon might perform better than the coconut husks. It just so happened that a nearby lab housed Mildred Dresselhaus, known as the "queen of carbon science" for her work on exotic forms of the stuff. She helped the trio build a forest of tiny carbon nanotubes, cylinders of pure carbon 10,000 times smaller than a human hair, that could boast over 2000 square metres of area per gram.

Ultracapacitors using nanotubes have gone on to be a success, notably through FastCap Systems, a firm founded by John Cooley, also from MIT. FastCap have produced capacitors that will help power NASA missions to Venus and deep space. Its best model can hold 10 per cent of the charge of one of Tesla's batteries, about twice as much as the next best commercial product.

Such nanotube designs are expensive, and in general ultracapacitor capacity is still not enough to put the Gigafactory in jeopardy – but that might not be the point. "We do not ever expect ultracapacitors to be the primary energy storage device in an electric vehicle," says Cooley. But if they can play the role of trusty sidekick, reducing the peak power load on tired batteries – the very thing that shortens their life – we could all benefit.

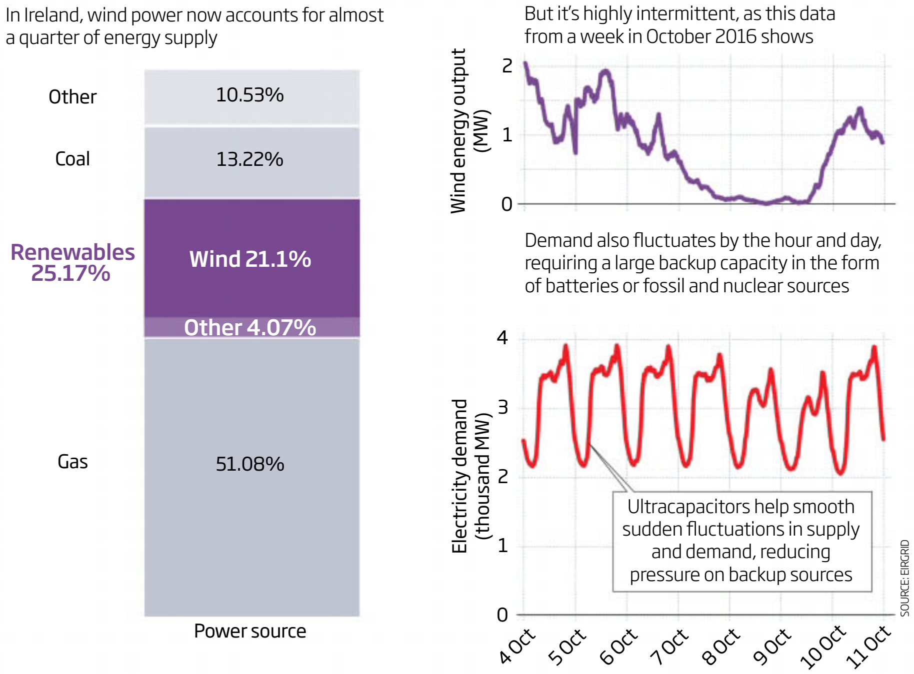
How so? While the idea of driving an electric car may or may not appeal to you, no one can ignore the problems facing electricity grids. We want energy supplies to be not just affordable, but reliable and green too. Ticking all those boxes is getting tougher, even for nations with highly developed economies. In October 2016, for example, the UK fell out of the top 10 nations in the World Energy Council's Trilemma Index, an energy security ranking.

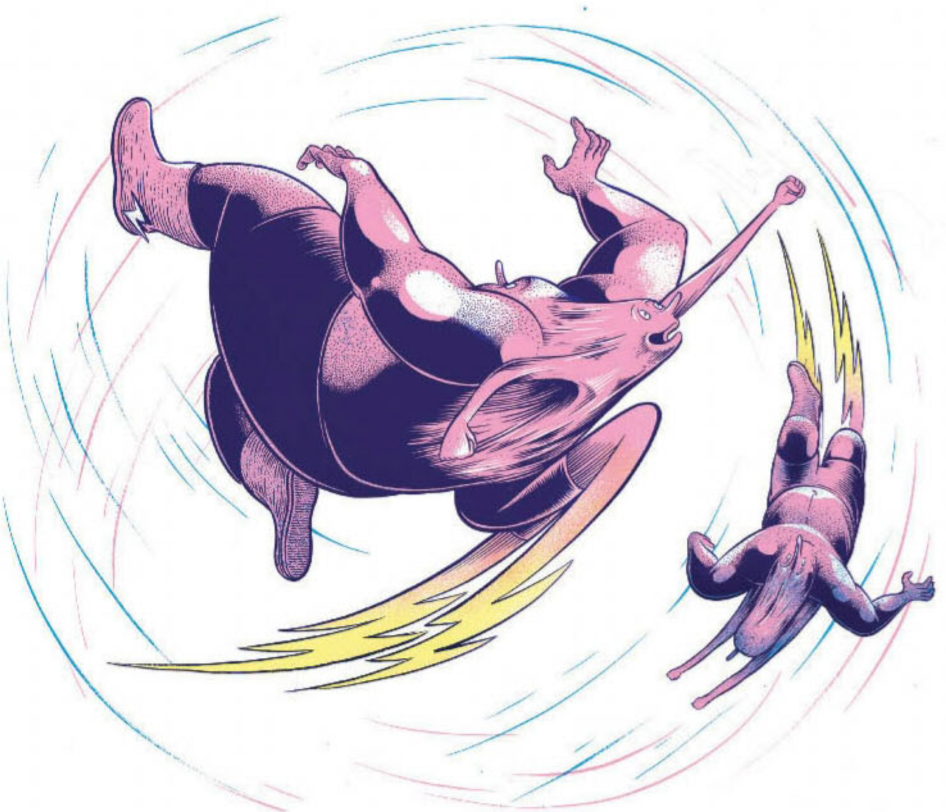
We have become serious about cheap green energy in recent years; renewables accounted for two-thirds of new generating capacity in the US in 2015, and over half worldwide, according to the United Nations. But on the one hand, demand for electricity varies widely and on the other, the supply of energy from renewables is intermittent. The wind doesn't always blow and the sun doesn't always shine (see "Supply and demand", left).

This problem has been met with the concept of the smart grid, where networks of sensors and switches constantly monitor and adjust the flow of energy from all sorts of generators to consumers. But this inevitably means storing the electricity, and those sluggish batteries are once again where we trip up.

Using batteries as the sole storage medium isn't ideal for two reasons. First, constant charging and discharging shortens their life. Second, batteries can't release all their energy quickly, so grids need excess battery capacity to cope with short surges in demand over and above normal fluctuations. Adding ultracapacitors instead of supersizing the

Supply and demand





RENAUD VICOURT

battery is a vast improvement. "The net effect is a reduction in the upfront expenditure and lower operating costs," says McGrath. "And the technology has now gotten to the stage where it blows the market open for us."

This year, Maxwell deployed two test ultracapacitor storage systems. One is in North Carolina, where the ultracapacitors are connected to a photovoltaic solar farm and a battery with a saltwater electrolyte. When the solar panels' output fluctuates due to passing clouds, the ultracapacitor goes to work. It can quickly supply nearly three times the power of the battery pack, but is exhausted in a couple of minutes. At that point, the battery, which holds about 40 times as much energy, steps in. The test is being carried out by Duke Energy, a utility company in the US with more than 7 million customers. It says the system is 10 to 15 per cent cheaper than a battery-only setup. "It should also slow down any degradation of the battery," says Duke's Randy Wheeless.

"It would be unwise to bet against ultracapacitors ousting batteries entirely"

Wind power is just as intermittent as the sun, and in the less balmy climes across the Atlantic it is the go-to renewable power source. In Ireland, wind power accounts for almost a quarter of electricity generation, and the country wants that to be 40 per cent by 2020. It is here that the second test is taking place, in an experimental smart grid in Tallaght, near Dublin. Ultracapacitors connected to local government office buildings have proved able to compensate for changes in frequency of the electricity supply within a fraction of a second. Klaus Harder of FreqCon, a German firm that supplied the ultracapacitor-battery hybrid storage unit, says the ultracapacitors are so far living up to their promise.

FreqCon has been testing a larger ultracapacitor-battery unit on the west coast of Ireland. But there is an ongoing challenge for the technology. Batteries may be imperfect, but they are still gradually improving. Ultracapacitors need to keep pace by increasing their capacity in tandem.

There is plenty of scope for that. Firms like EnerG2 say they will further improve the technology by using new sources of carbon to coat the electrodes. "Coconuts are cheap but they come with lots of natural contaminants and the activation process is toxic and expensive," says Feaver. EnerG2 designs its own carbon based polymers, similar to the resins used for laminating plywood. It then pyrolyses and activates them using a simpler, greener process.

EnerG2's carbon can also be tailored to different types of ultracapacitor. Those designed to quickly stop and start a car's petrol engine to improve fuel efficiency need a quick burst of power, but for smoothing a domestic solar panel's output, capacity might be more important than speed. Coconut carbon has pores whose size matches common electrolytes such as ammonium salts. But by adjusting its chemistry, says Feaver, EnerG2 can produce carbon with pores to match electrolytes designed for high power density, high energy density, or any combination of the two.

Practically invincible

Some think there could be greater leaps ahead if we break our attachment to carbon. William Dichtel, a chemist at Northwestern University in Evanston, Illinois, has developed polymer networks called covalent organic frameworks to work directly in ultracapacitors without needing pyrolysis. His team succeeded in producing a porous ultracapacitor material that approached the performance of a nanotube device but potentially at a fraction of the cost. "The caveat is that we're chemists doing basic research, not Tesla trying to put this in a car in a profitable fashion," says Dichtel.

There are concerns that exotic polymer-based ultracapacitors might not have the longevity of today's carbon systems. True, these ultracapacitors are not invincible, says Feaver. "But when you compare them with batteries, they might as well be." The battery in a cellphone or electric car is designed for 1000 charge-discharge cycles, whereas even Dichtel's experimental ultracapacitor was stable for at least 10 times as many cycles.

Ultracapacitors have come so far from their humble beginnings that it is tempting to wonder if they might graduate beyond their sidekick role and oust batteries entirely. We're far from that day, but perhaps it's unwise to bet against it ever arriving. We know that Elon Musk toyed with a PhD studying ultracapacitors before quitting for his first Silicon Valley start-up. And Tesla Motors' patents still make tantalising references to ultracapacitors. The man once so enamoured with ultracapacitors hasn't entirely lost faith, then. Maybe they are still evolving behind the doors of that huge factory in Nevada. ■

Going clean

Crack a simple chemical reaction and we don't have to kick our addiction to fossil fuels, says Jon Cartwright

SCARRED landscapes, billowing smoke, seabirds writhing in liquorice gloop: there's no denying fossil fuels have an image problem. That's before we even start to factor in the grave risk continuing to burn them poses to Earth's climate. But what's the alternative? Nuclear is expensive, renewables are unreliable, and we are a long way from making batteries that could power our fuel-hungry lifestyles. Realistically, we are going to be reliant on fossil fuels for a while yet.

What we need is a way to exploit them without emitting any planet-warming carbon dioxide. Alberto Abánades thinks he has the answer. He isn't a PR man for the fossil fuel industry, and nor does he have anything to do with various schemes to capture and bury carbon emissions after the event. He and his research team think they have cracked the problem using chemistry alone. By simply changing the way we liberate the energy trapped inside natural gas molecules, we can have all the benefits of fossil fuels – and none of the guilt. Too good to be true?

It's easy to see why we love fossil fuels. For a start, they are cheap and abundant. Discoveries of new resources and extraction techniques such as fracking mean reports of "peak oil" always seem exaggerated. They are reliable, too – you can shovel coal or pipe gas into a power station when the sky is cloudy or the wind's not blowing. And they can be portable – simply fill a car tank with petrol and you are good to go.

JACK HUDSON

We have tried to kick our fossil addiction before. During the oil crisis of the 1970s, all the talk was of hydrogen. The gas ticks a lot of boxes as a fuel: it is non-toxic and the most abundant element in the universe. It is clean, burning in air to create water vapour that falls harmlessly back to Earth as rain. It is energy-dense – you could drive the 600-odd kilometres from London to Edinburgh, or San Francisco to Los Angeles on a single tank. And it can be burned in power plants, even competing cost-wise with fossil fuels once carbon taxes are taken into account.

"It's easy to see why we love fossil fuels – they're cheap, abundant and reliable"

In practice, things aren't so simple. Being light and tiny, hydrogen has an annoying ability to wiggle through any material designed to contain it. Like petrol, it is flammable, yet burns with a near-invisible flame. Above all, it isn't abundant where and how we want it.

On Earth, hydrogen isn't a free agent. It is only found bound up in compounds such as water. Pure hydrogen can be generated by splitting water molecules using electrolysis, but that takes a lot of energy. Or you can extract hydrogen from coal or natural gas by heating them with steam, but that generates copious amounts of carbon dioxide.

So it came as little surprise when, in 2009, then US energy secretary Steven Chu, a Nobel prizewinning physicist, ditched funding for research into hydrogen-powered vehicles. In 2015, Elon Musk, CEO of electric-vehicle manufacturer Tesla, summed up many sceptical opinions when he labelled hydrogen an "incredibly dumb" alternative fuel.

Perhaps, though, we haven't been thinking about it in the right way. Natural gas is essentially methane, a molecule of one carbon atom and four hydrogen atoms. Rather than reacting natural gas with steam to liberate the hydrogen, Abánades, who is now at the Technical University of Madrid, and his team developed a deceptively simple plan. You "crack" the methane into its constituent atoms – pure, clean hydrogen, plus inert atomic carbon, or soot.

If it were that simple, it would already have been done. Breaking carbon-hydrogen bonds takes a lot of energy. They only start to crack spontaneously at temperatures above 550 °C or so; normally, temperatures over 800 °C are needed. But there is a bigger problem: the soot. This scuppered an early attempt to make methane cracking industrially viable: it coated the nickel-iron-cobalt catalyst used by chemists at the petroleum company Universal Oil Products to improve the reaction rate at lower temperatures. Their solution was to burn off the carbon – making carbon dioxide.

It's been the same lament with methane cracking ever since. Soot clogs things up and

BOOZE CRUISE

It is inherently difficult to compress flighty hydrogen gas into a fuel tank. The problem evaporates if you first convert it into a liquid alcohol, such as methanol. Aside from being easy to store, methanol can be used in regular internal combustion engines – where it can even perform better than petrol.

Compared with methane, methanol contains just an extra oxygen atom, but it is tricky to make from natural gas. It is much easier to create by combining hydrogen with carbon dioxide. The combustion of methanol in an engine releases carbon dioxide into the atmosphere, but if you use atmospheric carbon dioxide in the first place, the overall process is carbon-neutral. Eric Croiset at the University of Waterloo in Canada is hoping to work with a company to build a proto-plant that generates methanol in this way.

There are other options. In 2014, scientists at the Swiss Federal Institute of Technology in Lausanne reported a straightforward process for converting hydrogen into formic acid that can be fed into fuel cells, the battery-like power systems that drive hydrogen vehicles. The process is also reversible, so formic acid could be an alternative way to squirrel away hydrogen when regular storage is impractical.

the whole process grinds to a standstill. It's inevitable: the carbon has to go somewhere.

In his 20 years as an engineer, Abánades has worked on various types of energy generation, including nuclear and solar. His old group leader, Carlo Rubbia, first put him on to methane cracking in 2008. Rubbia had shared the Nobel prize for physics in 1984 for his part in finding the particles that govern nuclear decay, but, in his late 70s, he had long since turned his focus to energy innovation. "Professor Rubbia has always said to me, don't do what others do," says Abánades.

Bubble bath

Trawling back through the literature, they soon found something someone hadn't done. Back in 1999, Meyer Steinberg, a chemical engineer at Brookhaven National Laboratory in New York, and a veteran of the Manhattan Project to make a nuclear fission bomb, had proposed performing methane cracking in a heat bath made of molten metal. The idea, apparently never acted upon, was that the molten metal would improve heat transfer and allow the soot to float to the surface, avoiding clogging.

Abánades and Rubbia were then based at the Institute for Advanced Sustainability Studies in Potsdam, Germany. On the other side of the country, at the Karlsruhe Institute of Technology, was perhaps the best molten metal laboratory in Europe. By 2012 the two groups were collaborating on a 30-month fast-track project to see whether they could, well, crack methane cracking.

After two years of trial and error, they had what they thought was a viable reactor design: a vessel about the height and diameter of a hockey stick lined with quartz glass and stainless steel and filled with molten tin. Its external foil insulation made it look rather like a domestic hot water tank but it worked: they bubbled methane in at the bottom while raising the temperature of the tin up to 1000 °C, until hydrogen gas spouted continuously from the top.

But the real test was what it looked like inside. After two weeks, Abánades and colleagues switched off the reactor and peered in. Soot had indeed formed, but it had all floated neatly to the tin's surface, where it could be scraped away like the slag in an ore refinery. "We could even have operated the reactor for a couple more days," says Abánades. Last year, repeating the experiment at 1200 °C, the team managed to convert nearly 80 per cent of the methane they pumped in into hydrogen.

The notion that hydrogen can be continuously generated from methane, without directly producing any greenhouse gases, is enough to turn the heads of those in the field. "These are serious people," says Eric Croiset, a chemical engineer at the University of Waterloo in Canada, who performed a review of the state of methane cracking five years ago. "I wouldn't distrust their results."

We haven't reached the promised land yet, though. To heat their reactor, Abánades's team resorted to electricity from the wall socket – not necessarily the green option. A renewable source of heat, such as a solar concentrator, might do the trick, says Stéphane Abanades (no relation) at the French solar innovation lab PROMES, although there's a risk that when the sun sets or goes behind a cloud, the molten tin could solidify, damaging the reactor. "Supplying solar energy to such a reactor may not be an easy task," he says.

Alberto Abánades hopes that a future reactor could simply burn a little of the hydrogen it generates, perhaps 15 per cent of the total yield. This approach would generate similar low levels of carbon dioxide as hydrogen produced by wind-powered electrolysis of water, but would be cheaper, more reliable and more scalable, according to his team's preliminary analysis, performed in collaboration with RWTH Aachen

NICOLAS LORAN/GETTY

Clean hydrogen could transform our energy and crop production

University in Germany.

That still leaves the question of the soot. Scaling up methane cracking to terawatt-scale production – a reasonable extrapolation for a global hydrogen economy – would create a mountain of soot several cubic kilometres in volume every year. That is far less problematic than the carbon dioxide generated by directly burning fossil fuels, but still not an amount you can brush under the carpet.

"Other bits of the hydrogen puzzle seem to be coming together too"

Abánades is confident a cheap and abundant supply of pure black carbon will find its uses, given the element is already in demand for nanotechnology, steel production and as a filling for car tyres. "A new market could be opened up," he says. But first the carbon produced has to be of a higher quality. The methane cracking team believes its carbon is about 90 per cent pure, and could be improved either by tinkering with the reactor's chemistry or by purifying the carbon further down the line.

Is it full steam ahead for the hydrogen economy? Perhaps, especially as other bits of the puzzle seem to be coming together. For example, chemists are tinkering with ways to convert hydrogen into fuels that are easier to handle, such as methanol (see "Booze cruise" left). That might sound convoluted, but Abánades points out that oil is just as useless when freshly drilled from the ground. "Do we actually use crude oil? No, we transform it into gasoline. Hydrogen could be similar," he says.

Spurred on by cheaper hydrogen technology and the current range limitations of batteries, Toyota, Hyundai and Honda have all recently put cars powered by hydrogen fuel cells back on sale. In 2015, the European Union launched the Hydrogen Mobility Europe project, aiming to create a network of hydrogen refuelling stations across Europe. The UK government is providing small subsidies for fleets of hydrogen-powered vehicles. Croiset believes electric and hydrogen cars could address different markets, perhaps electric for inner city travel and hydrogen for longer distance commuting. "You will buy the vehicle that suits your needs," he says.

Others are less keen on the incentives that producing hydrogen-based fuels from natural gas create. The technology could commit us to more fossil-fuel infrastructure in the future, distracting from efforts to pursue renewable alternatives, warns climate scientist Ilissa Ocko from the Environmental Defense Fund, a New York-based non-profit that campaigns on global warming. What's more, the pipelines used to transport natural gas are known to leak a considerable amount of methane, a far more potent greenhouse gas than carbon dioxide. "Unless these leaks are plugged, it's possible that the warming from leaked methane will offset the climate benefits from methane cracking in the near-term," she says.

Abánades agrees that climate impact should be the deciding factor on which technologies to pursue. But in energy innovation, he says, it is tempting to view those working on different technologies as enemies, and it is easy to become tarnished by an association with fossil fuels. In the absence of a renewable silver-bullet, anything that limits the impact of fossil fuels has to be a good thing, he says. "Emissions should be stopped now, and that could be done through methane cracking. If they aren't, when it comes to controlling global warming, we will be too late." ■

THE REACTION THAT FEEDS THE WORLD

Should the futuristic hydrogen economy fail to materialise (see main story), hydrogen from methane cracking has a market ready and waiting: ammonium fertiliser. The Haber-Bosch process, which converts hydrogen and nitrogen into ammonia, generates most of the ammonium fertiliser used in agriculture. The reaction has been credited with fuelling the 20th century's population boom. It is so ubiquitous that it is part of you: over 80 per cent of the nitrogen that finds its way into the average person's tissue is thought to be as a result of the Haber-Bosch process.

Currently over 95 per cent of hydrogen production comes from traditional fossil-fuelled processes, mostly blasting natural gas with steam. In 2007 alone, the fertiliser industry generated a little short of 500 million tonnes of carbon dioxide, nearly 1 per cent of total global emissions. Re-supplying the Haber-Bosch process with methane-cracked hydrogen could drastically shrink this carbon footprint. With the world population expected to exceed 10 billion by the end of this century, that would be a significant step on its own.

New Scientist