

EdgeThe Third Culture

[Home](#) [About Edge](#) [Features](#) [Edge Editions](#) [Press](#) [The Reality Club](#) [Third Culture](#) [Digerati](#) [Edge Search](#)

A SELF WORTH HAVING [6.30.03]

A Talk with Nicholas Humphrey

What I'm now thinking — though it certainly needs further work — is basically that the point of there being a phenomenally rich subjective present is that it provides a new domain for selfhood. Gottlob Frege, the great logician of the early 20th century, made the obvious but crucial observation that a first-person subject has to be the subject of something. In which case we can ask, what kind of something is up to doing the job? What kind of thing is of sufficient metaphysical weight to supply the experiential substrate of a self — or, at any rate, a self worth having? And the answer I'd now suggest is: nothing less than phenomenal experience — phenomenal experience with its intrinsic depth and richness, with its qualities of seeming to be more than any physical thing could be.



Nicholas Humphrey *Edge* Video [Broadband](#) | [Modem](#)

Introduction

Nicholas Humphrey is a research psychologist whose interests are wide ranging: He studied mountain gorillas with Dian Fossey in Rwanda; was the first to demonstrate the existence of "blindsight" after brain damage in monkeys; and is the only scientist ever to edit the literary journal *Granta*. Thirty years ago he breathed life into the newly developing field of evolutionary psychology with his theory about "the social function of intellect." His more recent ideas concern the nature of phenomenal consciousness.

Unlike Daniel C. Dennett, who sees the role of philosophers as disabusing people of their "primitive" ideas about the nature of consciousness, Humphrey believes that we should take these primitive intuitions at face value. If people say that the problem is what it "feels like" to be conscious, then the problem is indeed to explain "feeling." Humphrey and Dennett are a pair of bookends. Humphrey has been described as a "romantic scientist", who believes in the heuristic value of stories that go beyond the limits of established facts. But he would probably not agree that there is a hard and fast line between facts and stories. "I'm me," he says. "I'm living an embodied existence, in the thick moment of the conscious present. I'm trying to work out why."

—JB

NICHOLAS HUMPHREY, School Professor at the London School of Economics and Professor of Psychology at the New School for Social Research, is a theoretical psychologist, internationally known for his work on the evolution of human intelligence and consciousness. His books include *Consciousness Regained*, *The Inner Eye*, *A History of the Mind*, *Leaps of Faith*, and *The Mind Made Flesh*. He has been the recipient of several honours, including the Martin Luther King Memorial Prize, and the British Psychological Society's book award.

[Nicholas Humphrey's Edge Bio Page](#)

[THE REALITY CLUB: Responses to Nicholas Humphrey by Kevin Kelly, Daniel C. Dennett, Jaron Lanier, Dylan Evans, Daniel Moerman, George Dyson, Thomas Metzinger, Jesse Prinz, Alva Noë, Rupert Sheldrake, Natika Newton, John Skoyles, Geoffrey Miller; Kevin Kelly; Nicholas Humphrey replies](#)

[A SELF WORTH HAVING?](#)

(NICHOLAS HUMPHREY:) Forty years ago I wanted to solve the problem of consciousness. It seemed to me it would be a shame to leave it for the next generation to get the prize. Consciousness presents the greatest ever challenge to science; so great, that unless we find an answer soon, science itself is in danger of being humbled. Consciousness — phenomenal experience — seems in many ways too good to be true. The way we experience the world seems unnecessarily beautiful, unnecessarily rich and strange.

I've had various goes at it: approaching the problem from different angles — through neurophysiology, through animal behavior, through social science, through philosophy of mind. My guess is we'll need all these approaches, and more, before we see what consciousness really is and what it's for.

Recently I've been toying with a rather grand idea about why we may need to have conscious qualia in our lives. My idea is that we need them in order to realize just how important we are. Our experience of being conscious encourages us as nothing else could to take ourselves seriously as *se/ves*. It dramatically affects our whole attitude to the kind

of people that we think we are. We find new value in our lives and, just as important, in the lives of other people.

I've come to this on the back of my earlier ideas about the nature of sensation. Some time ago I proposed a theory of how sensations work and why they have their qualitative properties. I argued that sensations derive their characteristic phenomenology from the fact that they are — in evolutionary origin — a kind of bodily action, involving reaching back to the stimulus at the body surface with an evaluative response. Conscious feeling, I suggested, is a remarkable kind of "intentional doing". Feelings enter consciousness not as events that happen to us but as activities that we ourselves engender and participate in.

When a person smells a rose, for example, he responds to what's happening at his nostrils with a "virtual action pattern": one of a set of action patterns that originated far back in evolutionary history as evaluative responses to various kinds of stimulation at the body surface — wriggles of acceptance or rejection. In modern human beings these responses are still directed to the site of stimulation, and still retain vestiges of their original function and hedonic tone; but today, instead of carrying through into overt behaviour, they've become closed off within internal circuits in the brain; in fact the efferent signals now project only as far as sensory cortex, where they interact with the incoming signals from the sense organs to create, momentarily, a self entangling, recursive, loop. My theory was that the person's sensation, the way he represents what's happening to him and how he feels about it, comes through monitoring his own signals for the action pattern — as extended, by this recursion, into the "thick moment" of the conscious present.

Now, I still think this is a pretty good idea. Especially because of its potential to explain the underlying functional architecture — even the neurophysiology — of phenomenal experience: the "what it's like" to live in the subjective present of sensations. The sensory loops I identified could create an "as-if" time dimension, so that every moment of consciousness lasts — paradoxically — longer than it actually lasts in physical time.

But there was a puzzle that I had pushed aside. I'd produced a model of reentrant circuits in the brain which might possibly provide the basis for the phenomenology of consciousness. I'd proposed an evolutionary story about how these circuits originated as a kind of bodily activity. But, if truth be told, I'd done nothing to explain *why* evolution had taken this remarkable course, at least I certainly hadn't explained the crucial final stage when the activity in the sensory circuits became self-resonant.

Let's be clear: this final stage can hardly have been an accident. In fact it must have required very fine tuning of the circuits to produce just the right degree of feedback — which is to say, to produce just the right degree and quality of temporal thickening of consciousness. But what's the point? Why ever should natural selection have gone to so much trouble to create a thick subjective present? Why don't we let conscious time slip by like physical time does? What can be the biological advantage to us of experiencing our own presence in the world in this magically rich way?



Humphrey, Minsky, Gould, Dennett, Brockman (August, 1995)

So that's what I'm working on now. And what I'm now thinking — though it certainly needs further work — is basically that the point of there being a phenomenally rich subjective present is that it provides a new domain for selfhood. Gottlob Frege, the great logician of the early 20th century, made the obvious but crucial observation that a first-person *subject* has to be the subject *of* something. In which case we can ask, what kind of something is up to doing the job? What kind of thing is of sufficient metaphysical weight to supply the experiential substrate of a self — or, at any rate, a *self worth having*? And the answer I'd now suggest is: *nothing less than phenomenal experience* — phenomenal experience with its intrinsic depth and richness, with its qualities of seeming to be more than any physical thing could be.

Phenomenal experience, surely, can and does provide the basis for creating a self worth having. And just see what becomes possible — even natural — once this new self is in place! As subjects of something so mysterious and strange, we humans gain new confidence and interest in our own survival, a new interest in other people too. We begin to be interested in the future, in immortality, and in all sorts of issues to do with co-consciousness and how far consciousness extends around us.

This feeds right back to our biological fitness in both obvious and subtle ways. It makes us more lively, more fascinating and more fascinated, more determined to pursue lives

wherever they will take us. In short, more like the amazing piece of work that humans are. Lord Byron said that "the great object of life is sensation — to feel that we exist, even though in pain." That's the raw end of it. But, at a more reflective level, what keeps us going, gives us courage, makes us aim high for ourselves and our children is the feeling that as human selves we have something very special to preserve.

None of this would have happened if it weren't for those sensory circuits in the brain developing their special self-resonance — a development that was pushed along by natural selection for metaphysics. As I once put it (imitating a famous passage of Rousseau): "The first animal who, having enclosed a bit of the world's substance within his skin, said 'This is me' was perhaps the true founder of individualized life. But it was the first animal who, having enclosed *a bit of time* within his brain, said 'This is my present' who was the true founder of subjective being."

~ ~

I've had the good fortune to be involved as a researcher in opening three different doors onto the problem of consciousness: through neuropsychology, ethology and aesthetics.

When I was a PhD student in Cambridge in the 1960s, I was at the right place at the right time to make a wonderful discovery: the phenomenon that later became called "blindsight." There was a monkey in Larry Weiskrantz's lab, called Helen, who had had the primary visual cortex at the back of her brain completely removed in a surgical operation. The operation had been done a couple of years earlier, and during the two years since the monkey had seemed to be almost completely blind.

However, there were reasons to think this might not be the whole story. And so, one week when I had time on my hands and the monkey wasn't involved in Weiskrantz's research, I decided to find out more. We were both at loose ends. Over several days I just sat by her cage and played with her. And an extraordinary thing happened. I realized that this blind monkey was interacting with me with her eyes. I would hold up a piece of apple and wave it in front of her, and she would reach out and touch my finger and try to get it from me. Within a few days she was transformed from a monkey sitting around listlessly, gazing blankly into the distance, to a monkey who had suddenly begun to be interested and involved in vision again.

I persuaded Larry to let me go on working with Helen. Over the next seven years I took her with me from Cambridge to Oxford, and then back to Cambridge. And she and I developed a remarkable relationship. I was her tutor and she was my apprentice. I encouraged her and coaxed her, trying in every way to help her to realize that actually she *wasn't* blind. I took her for walks in the fields and woods near the laboratory at Madingley near Cambridge. And slowly but surely I taught her to see again. In the end she could run around the room picking up crumbs off the floor, she could catch a fly as it passed by. If you didn't know this monkey had no visual cortex, you would have assumed she had completely normal vision.

Yet I was pretty sure that actually her vision wasn't normal. I knew her too well; we'd

spent hours and hours in this strange interaction, with me wondering what it's like to be her. And, though I found it hard to put my finger on what was wrong, my sense was that she still didn't *really believe* that she could see, that she herself was unaware of her capacity for vision. There were telling hints in her behavior. For example, if she was upset or frightened, she'd stumble about as if she was in the dark again. It was as if she could only see provided she didn't try too hard.

In 1972 I wrote a popular paper for the *New Scientist*, and on the front cover of the magazine they put the headline under Helen's portrait, "A blind monkey who sees everything." But that surely wasn't right. Not *everything*. My own title for the paper inside the magazine was "Seeing and Nothingness," and I went on to argue that basically this was a kind of seeing we'd never had any inkling of before. Could it be there was no phenomenal experience, no sensation accompanying it? With a monkey, who couldn't describe her inner world, there seemed no way of being sure.

Then, a couple of years later Weiskrantz, spurred on by what we'd found with Helen, moved the research to a new level by showing that a human patient with extensive damage to the visual cortex was equally capable of recovering some degree of vision. But now, with this human patient, it was possible to have him tell the researchers what it was like for him. And, to everyone's astonishment it turned out that, yes, this was indeed *unconscious vision* — blindsight. The patient believed he was blind, and reported no sensation, and yet he could still *guess* the position and shape of objects in the blind part of his visual field.

As I say, I was lucky. It was a remarkable break for a young student to have, and helped shape me both scientifically and personally. It was a transforming experience: day by day to watch in Helen the emergence of an "impossible" capacity. It was like being a midwife to a miracle. It made me feel good. But those seven years also left a different sort of mark on me. After such an unusually intimate experience, I no longer wanted to do research that involved brain lesions in monkeys. I still respect and admire those who continue to do this kind of work, but I myself wanted to go in a different direction.

~ ~

In 1974 I had the chance to go and work with Dian Fossey, studying mountain gorillas in the Virunga mountains of Rwanda. Dian was working for a Cambridge PhD, under Robert Hinde at Madingley, and I was nominally in a position of some authority over her, because I was assistant director of the lab at the time. I went to stay in her camp for three months, to help her with her research, to answer some questions, and to give advice if I could (although of course it really wasn't my place to give advice to Dian Fossey).

I was in an unusual position: a lab-based experimental psychologist, now given the chance to observe the behavior of apes in the wild. Those were the days when everything was much more relaxed than it is today, when I could set off alone at dawn to track down a particular group, and spend the day, even sometimes the night, with them.

In the grandeur of the mountains, half-accepted into the gorilla family, watching and

watched by a dozen black eyes, far from any other person, left with my own thoughts, I began musing about an issue that has fascinated me ever since: What's it like, for a gorilla, to be a gorilla? What does a gorilla know about what it's like to be me? How do we read minds?

When we're engaging with other human beings, we hardly notice the extent to which we are involved in mind-reading. We take it for granted. But the issue comes into much sharper focus when you find yourself doing the same with other animals who are similar to humans but perhaps not similar enough. It's a real challenge to know whether you're getting it right.

I was trying to understand what it was like to be a gorilla, living in a gorilla family in the forest. The gorillas were, maybe, trying to understand what it was like to be me. Puzzling about what was going on between us, I began to wonder about the special role of *introspection* and reflexive consciousness.

When we imagine what it's like to be another person, we project feelings, sensations, beliefs, and wishes into their minds. But of course we can only do this because we've experienced these very states of mind ourselves. Then, could this perhaps provide a clue as to why it's so important to us to be able to introspect? Could it be that the biological function of introspection — the reason the capacity evolved — is precisely that, by introducing us to how our own minds work, it helps us to read the minds of other people?

It dawned on me that this could be the answer to much that is special about human evolution. We humans — and to a lesser extent maybe gorillas and chimps too — have evolved to be "natural psychologists." The most promising but also the most dangerous elements in our environment are other members of our own species. Success for our human ancestors must have depended on being able to get inside the minds of those they lived with, second-guess them, anticipate where they were going, help them if they needed it, challenge them, or manipulate them. To do this they had to develop brains that would deliver a story about what it's like to be another person *from the inside*.

Later, I would call this new organ of reflexive consciousness the "inner eye".

~ ~

Since student days, I've been interested in aesthetics, in value. In fact my next laboratory-based project after finishing the research with Helen was to investigate whether monkeys have aesthetic preferences. I had a hunch that somehow the study of value must be relevant to understanding consciousness — though I wasn't then sure how.

Here's something to think about. Suppose you were to be turned into a sensationless "zombie": someone who is in every respect exactly like a normal human being except for not having phenomenal consciousness (and all that follows from it) — someone for whom the subjective present never lights up. Would life be worth living any more?

Early on in my career I got involved in another remarkable case study, that threw unexpected, and tragic light on just this question. A 27-year-old woman came to London from abroad in 1972 to have an operation to remove cataracts from her eyes. She'd been blind since the age of three. The doctor who operated on her had promised her that there was a good chance of being able to see normally again. I met her several months after the operation and found her in a state of great despair. She was convinced the operation was a complete failure; she couldn't see any better than she could before.

Unfortunately, it seemed all too likely that, as the result of years of lack of use, her visual cortex had in fact atrophied, so that she was in effect in much the same condition as my monkey, Helen. And yet, if this were the case, perhaps not all was lost. Perhaps she would be capable of learning to see again as Helen had.

I decided to try some of the same things with her. I took her out into in St. James's Park and around London. We walked through the gardens while I described the sights and held her hand. And soon enough it became clear that she did indeed have a capacity for vision that she wasn't aware of. She could point to a pigeon on the grass, she could reach for a flower, she could step up when she came to a curb.

It seemed that, after all, the operation had not been a total failure: her eyes were working again, to a degree. But was this what she was hoping for? No, it only proved the more traumatic. For the awful truth, she let on, was that her vision — just as in blindsight (and very likely it *was* a kind of blindsight) — still lacked any qualitative dimension. She'd been living for 20 years with the idea of how marvelous it would be if only she could see like other people. She had heard so many accounts, stories, poetry, about the wonders of vision. Yet now here she was, with part of her dream come true, and now she simply *couldn't feel it*. She was desperately disappointed, almost suicidal. In the end she dealt bravely with her situation by putting on her dark glasses again, taking up her white cane and going back to her former status of being blind.

This case stayed with me — to remind me, if I should ever forget, how much consciousness *matters*. Even to the extent that mattering may be one of the main reasons why consciousness exists. What if it's consciousness that gives us a reason for waking up every day, and going out into the world —to experience the qualia of a rainbow, the sunset, music, interactions with our friends, sex, food? What if consciousness provides such an incentive for living that, as human beings, we would not — and probably could not — do without it?

Of course human beings find meaning on lots of other levels. But the more I try to make sense of it, the more I come back to the fact that we've evolved to regard consciousness as a wonderfully good thing in its own right — which could just be because consciousness *is* a wonderfully good thing in its own right!

~ ~

You asked me to explain how I've developed as a scientist. I'll confess I've long had the ambition to make a difference — so that I leave the world a different place than it would

have been without me. But it's a funny thing: this kind of personal ambition can actually lead to an anxiety about being a scientist. The problem is that scientific truth is thoroughly impersonal. The answers to the questions scientists ask are in a sense already out there in the book of nature, waiting for someone — anyone — to find the key to reading them. If one particular scientist doesn't find the answer today, we can be pretty sure that another one will tomorrow or the next day — probably rather quickly given the way in which science is moving. So, if we're honest, we have to admit that even though we may have great fun in getting to the answer, and maybe great success and fame at having been the first to get there, in the end our personal contribution hardly matters. Even worse, perhaps what we are doing is to make some other scientist miserable — because we beat him to it.

This surely makes science a rather different enterprise from other great enterprises of our culture. Consider painting, or writing, or making music, where it's certainly arguable, that every creation is an individual work, which has the stamp of the person and the personality that made it. If Shakespeare hadn't written *Hamlet*, nobody would have written *Hamlet*. If Picasso hadn't painted the *Demoiselles d'Avignon*, nobody would have painted it. If Metallica hadn't composed their heavy metal music nobody would have done it.

Nonetheless, science does at least have this advantage over the arts: we need not doubt that the questions we are asking are important. Everybody wants to know the answer to the riddle of consciousness, or the origin of the universe. By contrast not everybody wants to know the answer to whatever it is that Metallica is trying to do. Or even, Picasso. Shakespeare? Well, perhaps Shakespeare's in a different league. Everyone *does* want to know the answers to the problems posed in *Hamlet*.

The ideal life, maybe, would be to create science, but in a style and with a way of presenting it which does have some of the qualities of great art, but which nonetheless has the security of providing verifiable answers to the big problems. Could someone be the Metallica of science? Maybe that's not the best model. But I'd say someone could certainly be the Dostoyevsky of science.

But, then, does it really matter whether your contribution has your personal stamp on it? I'd be the first to agree there are ways of making a difference — perhaps nobler ways — where who cares whether it's you or someone else. As scientists we have unrivalled opportunities to do things which, through their practical effects, make the world a better place to live in. And in my own work I'm afraid to say this practical element has been very much missing. I've made people interested and excited about ideas, but I can't claim to have done much to change people's lives for the better in any material way.

Maybe it's not too late. Recently I've been involved in research on the placebo effect, coming at it from a mixture of philosophical and evolutionary perspectives. The placebo effect is a very important aspect of all medicine. A large part of medical cures are effected by the patients themselves, when the medical procedure allows the patients to bring their own resources to bear to solve the problem. In the classical placebo case, you give a sugar pill and the patient uses this as an excuse to cure himself. But placebos are actually

present in every kind of medical treatment. To the extent the patient believes the treatment is going to work, he allows himself to deploy his own healing resources in a way that he wouldn't have done otherwise.

How should we understand this? What questions should a science of the placebo effect be asking? Of course it's important to investigate the brain mechanisms that underlie these effects, and lots of researchers are already beginning to home in on the problem at the level of neurophysiology and immunology. But it's no less important to look at the bigger picture, and ask: Whatever is going on here, from a *functional* standpoint? If a placebo is releasing in people an ability to cure themselves, why don't they just get on with it? Why ever should anyone withhold self-cure? You'd think that when you're sick you should just get better if you can; you shouldn't need to wait for permission from a doctor, a shaman, or a psychotherapist to utilize your own resources.

It's this level of question that has set me looking for some possible evolutionary explanation. Why should humans and other animals hold healing resources in reserve? What can be the advantages of *not* getting better when you actually could? As I've looked further, I've found many examples of it.

People may die from cancer when they have immune resources still waiting in reserve which could have been deployed against the cancer. People die in head-on car collisions because they don't apply the brakes hard enough. When athletes are running a marathon, they may reach the end of what they can do and collapse from fatigue, when, in fact, their muscles still have significant reserves left in them.

What's going on?

You'll have guessed the way I want to go with this: my idea is that nature has *designed* us to *play safe*, and never to use up everything we've got — because we never know what might still lie around the corner. When we reach the end of a marathon there may still be a lion waiting at the finishing post that's going to suddenly give chase. When we're sick with an infection and respond with an immune reaction, we may still be hit by a further infection the next day. Remember the story of the wise and foolish virgins and their lamps: it's always wise to keep something in reserve.

I'm now thinking in terms of there being what I call a "natural health management system", which does a kind of economic analysis of what the opportunities and the costs of self-cure will be — what resources we've got, how dangerous the situation is right now, and what predictions we can make of what the future holds. It's like a good hospital manager who has to choose if and when to throw resources against this or that problem, to hold so much back, to decide if it's essential to build up this area or that area — basically to try to produce an optimal solution to the problem of maintaining health with enough left over to meet coming challenges.

If this is right, it makes the placebo effect fit into a much larger picture of homeostasis and health management. And it converges with ideas being developed by researchers coming from quite different disciplines. I've been particularly struck by the work of the

South African physiologist, Timothy Noakes, who has come up with the idea of there being what he calls "a central governor" in the brain which regulates just how far the body should be allowed to go in meeting the demands of extreme exercise.

These ideas are big, because they are producing a new perspective on how we and other animals have evolved to manage our internal healing resources across the board. But it already goes much beyond mere theory.

There's a phenomenon, well known to sports physiologists and athletes called "interval training". If you want to improve your prowess as an athlete, one highly effective method of doing it is to build up in the following way. If you're a sprinter, for example, you sprint for two minutes and then relax and jog for five minutes. Then you repeat this pattern again, and again. The result is that you soon find you can run about 15% better than you could before.

Why does this work? According to Tim Noakes, what may be happening is this. In order to improve peak performance you need to persuade your central governor to let you go beyond your own self-imposed limits, when otherwise "cautionary tiredness" would kick in and say, "No more." And one way of doing this is by teaching your central governor that the risks are not actually so great after all. Through interval training you can teach your own brain that you are not going to get into trouble by pushing yourself a little further than you might otherwise have done.

Noakes' theory is a clever way of looking at how to stretch the limits of athletic performance. But what about applying the same idea in other areas? In particular, what about the possibility that we could have interval training for the immune system? If people are not deploying their immune resources to maximum extent, so that they don't get better when they could have, could we teach them by a similar schedule of exercise for the immune system that it's safe to do so?



Smolin, Pinker, Guth, Himpfrey, Dennett (June, 1997)

Here's the experiment. Let's do it in mice before we try it in humans. We give a mouse a bacterial infection. The mouse gets sick, and throws its immune resources against the infection — but only so far as it dares. Twenty-four hours later we follow up with antibiotics, and the mouse gets better. So the mouse's health management system gets the message that it's safe to go at least this far. Now, a week later, we repeat this pattern of infection followed by relief. Then we do it again, and again. And what I'd hope we'd find is that the mouse's health management system will learn that it can afford to use more of its resources than it otherwise would have dared to, because every time it goes to its own self-imposed limits it discovers it's followed by safe recovery.

Now, suppose we take one mouse which has been put through this regime, and another mouse which hasn't, and we inject them both with a carcinogen. I predict that the mouse which has been through interval training for its immune system will survive the cancer in a way in which a mouse that hasn't done won't.

If this were to work with people, imagine how it might turn medicine around! It might prove to be one of the best ways ever of achieving one of the main goals of modern medicine, which is to get people to use their own healing resources to greater and better effect than they usually do.

I have to say I really like the idea. Maybe this interview will be remembered as its first airing (I hope!).

*Edge*The Reality Club

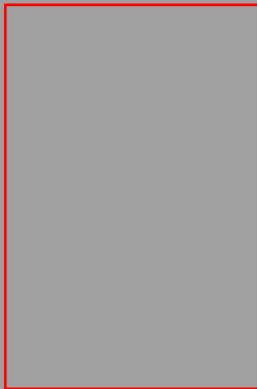
[Proceed to The Reality Club Discussion...](#)

EDGE SUMMER READING

"An engrossing treat of a book...crammed with hugely enjoyable anecdotes... you'll have a wonderful time reading these reminiscences." — *New Scientist*

"...inspiring glimpses into the minds of today's scientific intelligentsia" — *Publishers' Weekly*

Curious Minds: How a Child Becomes a Scientist

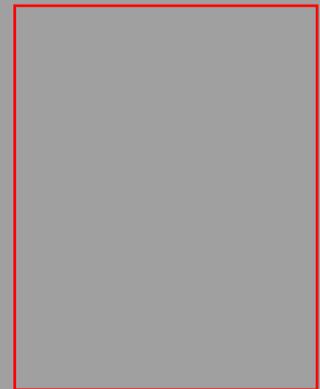


U.S. (August)

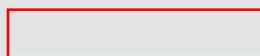
Original essays by Nicholas Humphrey • David M. Buss • Robert M. Sapolsky • Mihaly Csikszentmihalyi • Murray Gell-Mann • Alison Gopnik • Paul C. W. Davies • Freeman Dyson • Lee Smolin • Steven Pinker • Mary Catherine Bateson • Lynn Margulis • Jaron Lanier • Richard Dawkins • Howard Gardner • Joseph LeDoux • Sherry Turkle • Marc D. Hauser • Ray Kurzweil • Janna Levin • Rodney Brooks • J. Doyne Farmer • Steven Strogatz • Tim White • V. S. Ramachandran • Daniel C. Dennett • Judith Rich Harris • edited, with an introduction by John Brockman

[\[Click here for table of contents and contributors\]](#)

When We Were Kids: How a Child Becomes a Scientist



U.K. (September)





[Home](#) [About Edge](#) [Features](#) [Edge Editions](#) [Press](#) [The Reality Club](#) [Third Culture](#) [Digerati](#) [Edge Search](#)

*Edge*The Reality Club

Re: [A SELF WORTH HAVING: A Talk with Nicholas Humphrey](#)

Responses (in order received) by [Kevin Kelly](#), [Daniel C. Dennett](#), [Jaron Lanier](#), [Dylan Evans](#), [Daniel Moerman](#), [George Dyson](#), [Thomas Metzinger](#), [Jesse Prinz](#), [Alva Noë](#), [Rupert Sheldrake](#), [Natika Newton](#), [John Skoyles](#), [Geoffrey Miller](#); [Kevin Kelly](#); [Nicholas Humphrey replies](#)

Kevin Kelly: I really enjoyed Nick Humphrey's original posting on "A Self Worth Having", and was further impressed and enlightened by the many notable responses to it, but it wasn't until Nick unloaded his detailed counter-reply that I finally thought I understood what he was trying to say. In particular, the latter half of his reply, where he comes clean with his larger vision — that's when I lit up. Something clicked and then I combined this aha with a throwaway point of Marvin Minsky's about the necessity of obscurity. [\[more\]](#)

Daniel C. Dennett: If only more scientists were as willing as Nick is to show us the "early" parts of their thinking—the thinking you do before you present your findings, or design your experiments, or formulate your hypotheses, or even choose your field! — the wishful, impressionistic explorations in imagination that uncover the hunches that then actually motivate your life's work. [\[more\]](#)

Jaron Lanier: Bravo! The beauty of Humphrey's work is that his mind is active on so many levels at once. He quixotically attacks the problem of "Consciousness," but in the process he generates huge new ideas that are not only testable, but potentially useful. [\[more\]](#)

Dylan Evans: I share Dan Dennett's misgivings about Nick Humphrey's belief in the "wonderfulness in its own right" of consciousness. Wonderfulness is surely a relational property — a two place predicate — so nothing, not even Nick's consciousness — can be

"wonderful in its own right". [\[more\]](#)

Daniel Moerman: Nick Humphrey and I have worked together for the last several years in the "Placebo Group" organized by Anne Harrington through the Harvard University Mind /Brain /Behavior Center. We have often found ourselves with different positions on things, always with great interest and good humor. [\[more\]](#)

George B. Dyson: We're lucky Nick Humphrey did not solve the Consciousness Problem 40 years ago. We might have missed out on some of the wonderful details he's been exploring along the way. [\[more\]](#)

Thomas Metzinger: A Metallica of science! This one is going to haunt me for a while. Can we have a Patti Smith and a Sid Vicious of science too? And another Hendrix, some Primal Scream and a bit of early African Headcharge? Kerouac? A Cezanne of science? A Buddha? [\[more\]](#)

Jesse Prinz: There aren't enough idea people in science. There are dangers in being overly data driven. Too many good people spend careers tweaking experimental paradigms rather than charting out new terrain. Too many people are slaves to statistical significance. Nick Humphrey is always refreshing. [\[more\]](#)

Alva Noë: Nick Humphrey's work on consciousness is important. It is original; it is also sometimes maddening. [\[more\]](#)

Rupert Sheldrake: "A self worth having" is a fascinating synthesis of Nicholas Humphrey's ideas over the last forty years. [\[more\]](#)

Natika Newton: Dennett congratulates Humphrey on showing us the "early parts" of his thinking in this essay. But there are even earlier parts that I wish were reflected more here. [\[more\]](#)

John R. Skoyles: Is Nick Humphrey a fox or a hedgehog (to use Isaiah Berlin's famous division of intellectuals)? His diversity of ideas hints at him being a fox. But maybe Nick under his fox-like guise is a hedgehog with a big vision. [\[more\]](#)

Geoffrey Miller: Consciousness as a corporate pep rally. [\[more\]](#)

Kevin Kelly: I really enjoyed Nick Humphrey's original posting on "A Self Worth Having", and was further impressed and enlightened by the many notable responses to it, but it wasn't until Nick unloaded his detailed counter-reply that I finally thought I understood what he was trying to say. In particular, the latter half of his reply, where he comes clean with his larger vision — that's when I lit up. Something clicked and then I combined this aha with a throwaway point of Marvin Minsky's about the necessity of obscurity....So now I have my own version of what I think Nick is saying. Hopefully he can correct me if I'm way off. [\[more\]](#)

Nicholas Humphrey Replies: If Dan Dennett and I are bookends, then we are bookends that meet up round the back. Though I may be more impressed than he is by

the "magical qualities" of consciousness, I count myself as good a functionalist and anti-mysterian as anyone — and have learnt many of the philosophical moves from Dan.

[\[more\]](#)

[Kevin Kelly](#)

I really enjoyed Nick Humphrey's original posting on "A Self Worth Having", and was further impressed and enlightened by the many notable responses to it, but it wasn't until Nick unloaded his detailed counter-reply that I finally thought I understood what he was trying to say. In particular, the latter half of his reply, where he comes clean with his larger vision — that's when I lit up. Something clicked and then I combined this aha with a throwaway point of Marvin Minsky's about the necessity of obscurity.

So now I have my own version of what I think Nick is saying. Hopefully he can correct me if I'm way off. The summary as I see it:

The self is a survival mechanism for high intelligence. A rapidly expanding intelligence growing in power and dimension would be awfully confusing to a growing mind, particularly if it offered multiple views, and well... out of the body experiences. An emerging sense of self would be calming, focusing and attractive, a place to rest from wild ideas (like death and pain), allowing a more stable and effective human to survive.

An increasing intelligence would also reach a point where it became aware of its own intelligence — and that's a highly dangerous spot because an intelligence that was naked and transparent would be susceptible to intellectual manipulation. The first thing that a mind smart enough to see itself would do is start to hack itself. This is Minsky's point. A "thick" self that was unexplainable, obscure, and mysterious would prevent an emerging mind from subverting itself. This shroud of mystery would help the mind to grow more powerful, while keeping it attractive to its host (there's a lovely me inside here!), and protect intelligence from its own inclinations to "improve" things. Those humanoids with selves would survive longer, and the selves over time allow intelligence to spin in many directions.

Intelligence is the mind's worst enemy — a power so lethal it must be cloaked inside something unintelligible so it can't be easily broken by understanding, and at the same time be so emotionally appealing that a mind gravitates toward this cloak of unintelligibility because it feels more and feels good. A phenomenal self is that enabling cape. We don't see selves in other animals because their intelligence has not yet reached a point that it needs cloaking.

If the thickening self is a way for intelligence to protect itself from intelligence, then it may be but a phase, a kind of scaffolding that can be shunted away as the mind advances to explore safer ways to manage the

disruptive power of intelligence. A self may not be needed for an AI beyond a certain level, although it may be necessary to reach that level.

My bet is that should we ever engineer that choice for ourselves (to clear the cloud of "I"), we'll keep a mysterious self because, as Nick points out, we've co-evolved so that we (I mean our selves) see the self as the most beautiful thing in the universe. We really do. And of course the self tells us that we are crazy to let it go.

And we probably are.

kk

[Daniel C. Dennett](#)

If only more scientists were as willing as Nick is to show us the "early" parts of their thinking—the thinking you do before you present your findings, or design your experiments, or formulate your hypotheses, or even choose your field!—the wishful, impressionistic explorations in imagination that uncover the hunches that then actually motivate your life's work. If you want to know why scientists disagree so vehemently with each other, resist each other's arguments so pigheadedly, persist for so long in the effort to refute the competition, you often have to delve back into the way their imaginations have formed their unspoken setting of the issues. Nick tells us not just how he got there but why he wants the science of consciousness to turn out his way, and I daresay he is not alone in much or all of his motivation, so he manages to uncover a hidden agenda that probably drives a lot of others as well. Since I have been gamely resisting much of this drift for many years—with scant success!—I find it is always valuable to get a glimpse into what may be driving it.

"What can be the biological advantage to us of experiencing our own presence in the world in this magically rich way?" This question compresses several different questions into one "magical" lump: Who is "us"? *Homo sapiens*, or primates, or mammals, or vertebrates? (Do fish "experience their own presence"? Do birds?) Nick says "we humans gain new confidence and interest in our own survival" so perhaps he is reserving this explanation for *human* consciousness, but I don't see any reason to believe we are any more intent on our own survival than, say, a fox that gnaws off its own leg to escape from a trap. And why would the advantage have to be "to us" in the first place? Might consciousness (of the human sort) be an affliction that spreads because it can? Perhaps the benefit in fitness is to the parasitic habit we spread to our children when we raise them with language, an energy-wasting, distracting price we must pay to be communicators! (Remember: genetic fitness isn't about pleasure or delight; it's about how many offspring your offspring have.)

Other remarks of his suggest that he means to include primates, at least, in the ambit of his thesis, but then his remarks about the "first animal who . . . said [sic] 'This is me'" raise a problem. And—most crucial of all—the *motivation* that Nick imagines qualia to underwrite has to be *explained by a functional model that deconstructs qualia into*

something else, something that is not a content of consciousness. If qualia are presupposed to be *wonderful* (set aside "magical," which is just a rhetorical flourish) or *delightful or delicious*, or anything of the sort, then we're stuck with something like "intrinsically delicious" properties, which couldn't explain anything (any more than "intrinsically awful" pains could explain pain). The deliciousness or awfulness, the importance, in short, that makes selves worth having, as Nick says, has to be constructed out of dispositions to behave, to seek or avoid (and the convoluted variations thereof), not the other way around. But it sure doesn't seem that way. Here Nick goes along with, actively endorses, what I consider a prime example of a pre-scientific perspective that must be challenged.

I don't know if Gottlob Frege ever observed that a first-person subject has to be the subject of something; he is sometimes cited as making the mirror-image point: that an experience has to be someone's, some subject's. I myself am dubious of the utility of both versions of the point, since unless they are deemed definitional, they leave problem cases in their wake. The "pains" that cause you to shift position while you sleep—is there a subject of those pains? Are they not experienced because there is no subject at the time? Could a robot be a subject, and if not, why not?

Robots, many think, would have to be "zombies," and Nick asks you to imagine you were turned into a "sensationless 'zombie': someone who is in every respect exactly like a normal human being except for not having phenomenal consciousness (and all that follows from it)—someone for whom the subject present never lights up." We are asked to imagine being turned into a person who was like Nick's blind woman squared, cubed—sensationless in *every* modality. A pitiful state indeed, but this is not what a philosopher's zombie is. This is a being that would be hugely different behaviorally from a normal human being, an unmotivated, listless vegetable, while a philosopher's zombie, by definition, is as lively and (apparently) motivated as anybody could be. It is this (inadvertent) sleight-of-hand, this mis-imagining of the zombic hunch, that keeps the zombie concept alive (well, undead) when it should have lapsed into oblivion years ago.

Nick knows all this. But he persists, and he tells us why. He wants us not to forget "how much consciousness matters," and he proposes that we've evolved to regard consciousness as a wonderfully good thing "because consciousness is a wonderfully good thing in its own right!" Think about what this "in its own right" could mean. Are we to imagine the possibility of creatures that *had* consciousness (that wonderfully good thing in its own right) but were oblivious, somehow, to their good fortune? Would they be deeply wrong not to treasure it? Would they leave less descendants than cousins who somehow began to get an inkling of how wonderful consciousness actually was? What would be wonderful about it if they were ill-equipped to enjoy it? (Imagine the Planet of the Rolexes—inhabited only by tortoises that happened by cosmic coincidence to be born with exquisite Rolex watches embedded in their shells. But they don't tell time, have no need to tell time, and often scrape off their Rolexes in the course of mating. Dumb tortoises! If only they knew what they had! But in what sense would they be wrong?) If wonderfulness must co-evolve with the capacity and disposition to appreciate and exploit it, then it isn't wonderfulness "in its own right." Is it?

Nick and I have been going round and round on these topics for almost twenty years,

and I've been amazed and provoked by his ingenious ways of putting his points—but never convinced. I'm pretty sure I still don't get it, but it's always a joy to play another round.

[Jaron Lanier](#)

Bravo! The beauty of Humphrey's work is that his mind is active on so many levels at once. He quixotically attacks the problem of "Consciousness," but in the process he generates huge new ideas that are not only testable, but potentially useful.

Since Humphrey's style is almost Proustian, I'll briefly report that while reading his new idea of Placebo, I experienced the internal flux of contradictory reactions that collectively form the intuition that his is not only a "Big" new idea, but one with legs.

First you're certain that something so elegant and obvious in hindsight can't possibly be new. Surely you've even thought of it yourself. Alas, not. Then the flash of jealousy. Then the rush of associations. See how it compares with ideas in economics about perceptions of future values in individuals and markets. Draw a metaphor between the immune system and Muhammad Ali's "Rope-a-Dope." Hurry, hook it up to Csikszentmihalyi's idea of "Flow."

We already knew our paranoid genes build up our fat even when famine is unlikely and we'd prefer they butt out, and now it turns out they might withhold vital immune responses as well. A metaphor presents itself in which national policy maps to the body. Why so many uninsured in America? Because of an excessively conservative philosophy! What if this is a moment when the Polis can afford to heal itself after all?

The "Self" problem ought to be differentiated from the "Consciousness" problem. A "Self" can be thought of as one part of a system which models, predicts, and to some degree directs the rest of the system within its environment. There are purported states in which human subjects are said to have lost some significant degree of self while retaining some significant degree of consciousness, including some forms of hypnosis, non-metaphorical (Haitian) zombiehood, and recreational drug use. I don't claim to be an expert on these phenomena, but I do think there's something to be lost if people think of their selves as merely being functional parts. (I've described my pragmatic objection to a cousin of self-blindness, the artificial intelligence worldview, elsewhere on [edge.org](#). Briefly: AIs will tend to create poor quality user interface designs, because they are willing to reduce their human standards to achieve parity with whatever it is computers can do at the time.)

Dan Dennett has his own charming blindsight, in that he certainly displays a robust sense of self but seems to either not experience himself internally or be immune to awe at that sense of experience. As always, I suspect he is secretly as mystified by his internal life as anyone else, but enjoys the reactions he can provoke by pretending to be a sophisticated zombie.

Dylan Evans

I share Dan Dennett's misgivings about Nick Humphrey's belief in the "wonderfulness in its own right" of consciousness. Wonderfulness is surely a relational property — a two place predicate — so nothing, not even Nick's consciousness — can be "wonderful in its own right".

Still, I think that Dan's objections are largely concerned with the way that Nick has expressed his hypothesis, and fail to engage with the substance of that hypothesis. Perhaps Nick invites such responses by choosing to employ a particularly flowery style of writing, but he is at least practising what he preaches — aiming at "the ideal life", in which one can "create science, but in a style and with a way of presenting it which does have some of the qualities of great art".

Nick's hypothesis is that we humans (and perhaps some particularly gifted chimpanzees) find the extra motivation we need to compete in the great Darwinian race in the peculiarly rich kind of conscious experience that only we are blessed with. This hypothesis begs a whole bunch of questions that need answering — for example, why do we need some extra kind of motivation to survive and reproduce when foxes don't? Maybe Nick will share his answers to these questions with us later; I have a feeling that he's deliberately withholding some of the missing pieces of this jigsaw so that he can astonish us again later. But whatever his answers turn out to be, they won't go any way towards meeting my deepest objection to his hypothesis.

This objection is just a hunch, really — a gut feeling. So I won't argue for it; I'll just spell it out. It strikes me that Nick is simply reading back into our evolutionary past things that may have emerged much later. The magical/wonderful/delightful quality of consciousness that Nick waxes lyrical about seems inextricably wedded to the development of all those cultural technologies — poetry, art, philosophy and so on — that only emerged a hundred thousand years or so after *Homo sapiens* first walked the earth. In other words, I think Nick has mistakenly assumed that the greatest jewels of human consciousness are products of natural selection when they are most likely products of history and culture.

There's too much individual variation in the capacity for magical/wonderful/delightful phenomenal consciousness for it to be the result of natural selection. How many people get the chance to be the "natural philosophers" (in both senses of the phrase) that Nick thinks is the common lot of humanity? Nick may be overly generous in ascribing to the whole human race the supercritical kind of consciousness which he so clearly possesses in abundance.

Daniel Moerman

Nick Humphrey and I have worked together for the last several years in the "Placebo Group" organized by Anne Harrington through the Harvard University Mind /Brain / Behavior Center. We have often found ourselves with different positions on things, always with great interest and good humor.

I have not thought about consciousness the way Nick Humphrey has; as an anthropologist, I've always assumed it as the essence of human being-ness, and gone on from there. And, thinking about it now as Nick sticks his intellectual elbows into my virtual ribs to do so, I come again to a conclusion with which he probably won't agree. I find it hard to believe that there is any adaptive explanation for consciousness.

Consciousness is a gift, and perhaps one from the Devil. It makes no sense. Five thousand other mammals from platypus to dolphin manage without anything remotely like a human system of consciousness, language, meaning, recursion, uncountable sets, aesthetics, etc. Yes, all animals (mammals and planaria) probably have some sense of self (although in some cases, like slime molds, it's hard to know where it would reside); all sexual animals, at least, communicate at least once in a while (well, oysters do it without much communication that makes any sense to me; so let me change it to "most sexual animals"). Some stuff may mean things to primates; although that obviously depends on the definition of "mean," something that would be hard to discuss with the wisest chimpanzee (which is, I guess, the point). And, of course, I know lots of human beings who have utterly no sense at all of aesthetics, even if they can "talk," in some sense of the word.

All this says that human beings are the result of an evolutionary process, and that we have close animal kin. But we only need look at the anatomy of the hand, or the liver, to know that. When it comes to consciousness, to the "individualized life" as Nick calls it, we have, in my view, no peer, nothing even close; this is the land of saltation. What I've said about the lack of non-human consciousness is obviously controversial (see Dan Deanne's comments), but among the most successful living forms on the planet are plants, in the broad sense — ferns and redwoods, mosses and roses, alga and fungi — where I doubt much controversy exists regarding consciousness; I once marveled to see a veil of green alga growing around the glowing blue radioactive rods of a nuclear power plant under 20 feet of water; they will outlive us all.

So, given a) the astonishing persistence of non-individualized life, of life free of human-style consciousness (for tens of thousands of animal species, and hundreds of thousands of plant species), and b) the damage that we consciousness-rich persons have done to the whole ecosystem, to the evolutionary system which has been going on for a billion years (more damaging than a streaking asteroid, than a billion volcanos, than the drifting of continents; or whatever), it seems to me that we have to look at consciousness as not an evolutionary (and specifically adaptive) development (which Nick notes is incredibly hard to account for, in the way that we can account for other adaptations, like sickle cell anemia, or tool making), but an accident, or a gift, or both.

Similarly, I find it hard to account for, say, art (e.g., Chauvet, Lascaux, MOMA) in such terms; art seems to me a decisive element in human evolution, but it's not something which can be accounted for in terms of differential reproduction.

It makes good sense to argue that mammals replaced dinosaurs because they were better adapted to early life: warm blooded animals could do this and that which dinosaurs couldn't. That's what I learned in school. But that's not what happened; the dinosaurs (and others) were destroyed by the C/T asteroid, and mammals evolved subsequently: an accident with an astonishingly small probability. Rewind the clock to the big bang, add one more dust mote somewhere in our galaxy, the asteroid would have missed, and, chances are, the mammals would have continued on as a minor element in the ecosystem: no consciousness. A scenario like that seems to me as plausible, or more plausible, than adaptation as an explanation for consciousness, even though after it appears, it is a marvelous and incredibly powerful device for living on (and destroying) earth.

Also a comment on Nick's notion of the "natural health management system," and Timothy Noakes' "central governor." These quite similar notions seem to me very English and very European, the homunculus version of the National Health Service. To an American, even a very left liberal humanist American, my insides seem to me much more chaotic than that. There are winners here, and losers there (heart up, liver down), there's precious little planning (or if there is, no one pays attention; if someone did, we wouldn't be obese, or have high blood pressure or elevated cholesterol), and there's never a cop when you want one. There is quite a bit of cleaning up after accidents and upsets (the immune system is always off doing its thing, often unfortunately to excess, creating rheumatoid arthritis and lupus), and so on. In any event, if someone is in charge in there, it's time for a revolution. But I don't think there's anyone to revolt against; rather like the Wizard of Oz, there's no one of substance behind the curtain.

So, as it seems often to happen, Nick and I disagree, but, we disagree about the most interesting things, and things which are extraordinarily worth thinking about.

[George B. Dyson](#)

We're lucky Nick Humphrey did not solve the Consciousness Problem 40 years ago. We might have missed out on some of the wonderful details he's been exploring along the way. In my opinion, the Consciousness Problem (and a Theory of Everything for physics) is Nature's way of keeping us humble while drawing our attention to some of the details that really count. One of these is the placebo effect.

A few years ago I was in Phoenixville, Pennsylvania, a small town where many buildings were vacated when the steel industry declined. I was visiting the last company still manufacturing punched-card data processing machines, and they were doing a booming business (this was just before Y2K). But there was even more activity in the factory next door. Their product? Placebos. For almost any drug on the market, they produce an inert replica: round pink pills, triangular red ones, blue ovals, yellow tablets, the entire range. We know they work, but we do not understand why.

As I drove back into New Jersey on Highway 1, past one pharmaceutical giant after another, I pondered the irony that the drug companies have built research lab upon research lab in expensive Princeton, hoping to discover new products, while the placebos are made in an abandoned factory with no research labs (and no lawyers) because they have one product, no side effects, no patents — and it works. How it works is still a mystery. I look forward to Nick Humphrey shedding new light on this.

[Thomas Metzinger](#)

A Metallica of science! This one is going to haunt me for a while. Can we have a Patti Smith and a Sid Vicious of science too? And another Hendrix, some Primal Scream and a bit of early African Headcharge? Kerouac? A Cezanne of science? A Buddha?

Of course, no such things as selves exist in the world: nobody ever *was* or *had* a self. A self could never be something you *have* — like a bicycle or book by Dostoyevsky. You could only *be* one. What the organism as a whole can have is an internal model of itself as a whole: all that ever existed were conscious self-models that could not be recognized *as* models. The phenomenal self is not a thing, but a process — and the subjective experience of *being someone* emerges if a conscious information-processing system operates under a transparent self-model. You are such a system right now, as you read these sentences. Because you cannot recognize your self-model *as* a model, it is transparent: you look right through it. You don't see it. But you see *with* it. In other, more metaphorical, words you constantly *confuse* yourself with the content of the self-model currently activated by your brain.

This is not your fault. Evolution has made you this way. On the contrary: arguably, until now, the conscious self-model of human beings actually is the best invention Mother Nature has made. It is a wonderfully efficient two-way window that allows an organism to conceive of itself *as a whole*, and thereby to causally interact with its inner and outer environment in an entirely new, integrated and intelligent manner. Consciousness, the phenomenal self, and the first-person perspective are fascinating *representational* phenomena that have a long evolutionary history, a history which eventually led to the formation of complex societies and a cultural embedding of conscious experience itself.

Nick Humphrey is right that introspection is absolutely central for social cognition: it allows you to discover that you are a representational system at all, that you generate internal states which may be true or false. At the same time allows you to *inspect* the results, which your low-level mind-reading systems — your mirror neurons in F5 of your premotor cortex — automatically spit up all the time. It allows you to understand that your mind is driven by other agents in the environment, and that you can drive *them*. But in order to "intro"spect, you need a coherent, first-order self-model in place first! Then you can start to fool other people.

Placebos work via the human self-model too: it has a conscious tip-of-the-iceberg, which is neurocomputationally anchored in low-level processes of constantly drawing a self-

world boundary, like the immune system or homeodynamic autoregulation in the upper brainstem/hypothalamus. Change the representational content of the conscious self-model in the right way, get some unconscious microfunctional output, get an effect for free.

Believe in the magic of consciousness, get an effect for free as well. It is obvious how false beliefs can be highly advantageous: just think of the idea that one can be permanently happy with members of the other sex or that children automatically give meaning to your life. Here is some terrible philosopher's jargon for you: functional adequacy is not epistemic justification. And it would be great to live in a world where *Magic Happens!* Come on, Dan! Admit it: there are aspects of the scientific world-view which may be damaging to our mental well-being, and *that* is what everybody intuitively feels, and *that* is why people look for back doors and placebos.

The incentive Nick speaks about is not in consciousness as at such, but in the transparency of the conscious human self-model: the new functional property is that it makes a system maximally egotistic, a true believer in itself, in whatever its reward system and its emotions currently tell it. Why? Remember: it cannot recognize its own self-model *as* a model. Nick is also right about the relevance of the thick present, the phenomenal Now: computationally speaking it is the simplest form of explicit time-representation, it allows for explicit predictions in time, creates a virtual window of presence and allows an animal for the first time to represent the fact that *a world is present* (my own minimal notion of consciousness, by the way).

The system is *in time*. Things explode when you combine these two: have a coherent world model integrated with a thick moment and then put a rich, transparent self-model into it. What you get is *Being Someone* - the phenomenal experience of a self currently being present in a world. Is that worth having? Is it in our interest to have *this* form of consciousness?

Flowery placebo or not, the merit of Nick's contribution lies in drawing attention to a truly deep, highly relevant and constantly neglected issue. It is not at all clear if the biological form of consciousness, as so far brought about by evolution on our planet, is a *desirable* form of experience, an actual *good in itself*. Let me further provoke Nick by playing the Gloomy German here.

The theoretical blind spot of current philosophy of mind is the issue of conscious suffering: thousands of pages are being written about color qualia or the contents of thought, but almost no theoretical work is devoted to ubiquitous phenomenal states like physical pain or simple everyday sadness ("subclinical depression"), or to the phenomenal content associated with panic, despair and melancholy — let alone to the conscious experience of mortality or of losing one's dignity. There may be deeper evolutionary reasons behind this cognitive scotoma, but I am not going to pursue this point here (didn't Jaron Lanier talk of "death-denial" some years ago?)

The ethical/normative issue is of greater relevance. If one dares to take a closer look at

the actual phenomenology of biological systems on our planet, the many different kinds of conscious suffering are *at least* as dominant a feature as are color vision or conscious thought, both of which appeared only very recently. Evolution is not something to be glorified. One way — out of countless others — to look at biological evolution on our planet is as a process that has created an expanding ocean of suffering and confusion where there previously was none. As not only the simple number of individual conscious subjects, but also the dimensionality of their phenomenal state-spaces is continuously increasing, this ocean is also *deepening*. For me, this is also a strong argument against creating artificial consciousness: we shouldn't add to this terrible mess before we have truly understood what actually is going on here. I admit that there exists unfathomable beauty in phenomenal experience.

Nick Humphrey's "magical richness" may not be intrinsic, and it certainly is open to functional explanations. But nobody seems to see how we pay a *high* price for that beauty — which raises the normative issue if it is a value, a good "in itself". Personally, I have my doubts that the conscious self-models we have today are worth having. We better stop glorifying our own neurophenomenological *status quo*, and find the courage to think about positive alternatives in a rational way.

[Jesse Prinz](#)

There aren't enough idea people in science. There are dangers in being overly data driven. Too many good people spend careers tweaking experimental paradigms rather than charting out new terrain. Too many people are slaves to statistical significance. Nick Humphrey is always refreshing. To read him is to experience the rush or vertigo that inevitably accompanies any trip out on a limb. Should we follow Nick through his latest forest of tenuous stalks and twisting branches? Perhaps we should — as long as we don't grasp for those nimble leaves that cluster fetchingly at the tips.

Might consciousness have evolved to give us a sense of self? That depends. The self is not any one thing. We have biographical selves, self conceptions, and self-defining long-term goals. It's pretty likely that consciousness is functionally independent of these things. Indeed, the idea of a single, enduring self, which these notions tend to presuppose, may be a cultural construct. Is my biographical self really the same person as my self of aspirations? Do I really have just one biographical self? The Egyptians believed in seven souls. The Greeks (like Freud) believed in three. Which one is me? Perhaps we need conscious minds to raise such questions, but to suggest that consciousness arose for this lofty purpose (or for the purpose of postulating other minds) is an engrandizement of experience.

There is, however, a more minimal sense of self. It is captured in a feeling of agency. We sometimes do things deliberately. We sometimes exert executive power over mental events. These "controlled processes" constitute a thin kind of self-ownership. If consciousness is linked to that skinny self, then Nick's idea may be vindicated. I think there is reason to believe in such a link. I have been defending the view that consciousness arises when perception becomes accessible to working memory by means

of attention. If that is right, then consciousness arises under just those conditions when perception becomes available for deliberative response. Human blindsighters (what to say about Helen?) do not respond spontaneously to visual inputs, because those inputs usually don't get to working memory. When blindsighters have phenomenal experience in the blind fields when and only when stimulus parameters are such that working memory can be entrained.

In a word, I think consciousness is a precondition for control. If minimal me-ness consists in control, then consciousness is a precondition for one kind of self. Moreover, it is plausible that consciousness was selected for this purpose. I think consciousness came on the scene with working memory. If working memory is an adaptation, then so perhaps is the process by which information becomes available to working memory.

In sum, I think Nick is right to pursue the connection between consciousness and self. He may also be right about the placebo effect. And, if the minimal-me is a control hub, then there may be much to gain in bringing our native healing powers into consciousness. I hope that Nick's speculations take root.

[Alva Noë](#)

Nick Humphrey's work on consciousness is important. It is original; it is also sometimes maddening. His key insight that consciousness depends on action continues to bear fruit. More specifically, Nick is on to something fundamental when he proposes that a crucial station in the evolution of consciousness is the appearance of creatures who can decouple motor response from sensory stimulation. That's what simpler organisms, with their simpler bodies, cannot do. For the amoeba there is a direct bio-chemical linkage between stimulus and response. For the spider, or squirrel, or person, there is not.

The importance of Nick's idea lies, I think, in the way it opens up the possibility of a genuinely biological study of mind. The evolution of bodies of increasing complexity is something we have a theoretical grip on. But this very increase in bodily complexity brings with it, necessarily, more complicated patterns mediating stimulation and response. For Nick, there aren't two stories, the story of the evolution of the body and that of the evolution of the conscious mind. There is one story: the evolution of complex, embodied, sensorimotor systems. An aspect of this approach that I like is that it isn't brain-centered: it is the animal — the sensorimotor system — that is the locus of consciousness, not the isolated brain or nervous system. Too many discussions of consciousness today — encouraged perhaps by technologies of brain imaging — seem to think the brain (in a rather Cartesian way) is the seat of consciousness.

This brings me to the maddening part. Does Nick really believe that the "self-entangling, recursive loops" in our neural circuitry somehow explain phenomenal consciousness? How could such neural structures or episodes explain anything at all about the felt quality of a tickle, or the experienced character of redness? Is there something magical about reentrant, self-resonating neural circuits? Is it their loopiness perhaps? It is high time we give up the search for this sort of neural substrate of experience. Nick's basic

insight — that consciousness belongs to the sensorimotor system — provides the resources we need to move beyond just this sort of picture.

Colin McGinn has argued that humans may not be smart enough to explain the phenomenon of consciousness. We are just not evolved enough. If Nick is right, consciousness may have evolved precisely thanks to our inability to explain it! Consciousness, with all its mystery, is what gives us a self worth having!

A neat idea. But could it be true? It is microbes, not animals, who rule the earth, and they do so, I dare say, without phenomenal consciousness or a sense-of-self. And as for animals, true, we've got the most inflated sense of ourselves, but not because we have more phenomenal consciousness than the fox has, or the elephant. And then there is the fact — illustrated beautifully by Humphrey's own work on blindsight — that phenomenal consciousness is not necessary for the perceptual guidance of action. Indeed, much of what humans do best — the four Fs, athletics, playing instruments — we do automatically and without the mediation of reflection.

Hume famously denied that the self is anything other than a bundle of fleeting sensations, perceptions and images. There is no self, in the sense of a persisting Cartesian subject of thought and experience. Hume also denied that it is any part of our naïve experience that we take ourselves to know such a self! If you introspect, Hume urged, you will not encounter a self. Rather, you will meet only this or that fleeting mental item. Hume suggests that we would be truer to our experience of ourselves, if we give up the idea that there is a self at all!

In the present context, reflection on Hume reminds us that the very idea of a sense of self is little more than a place holder. (The same goes for the "phenomenally rich subjective present.") What is a self worth having, in an evolutionary context? If Hume is right, and the self is just a bundle, then one can reasonably wonder why we need phenomenal consciousness for that.

But I bring up Hume for another reason as well. John Brockman's comparison of Nick with Dan Dennett needs comment. Yes, Dan disabuses people of their naïve ideas about the mind, while Nick is more willing to take seriously what people say about their own minds. But we shouldn't overemphasize this difference. Consider the possibility of a Humean synthesis: that careful first-person reflection on one's own experience reveals experience to be precisely what Dan says it is! What then?

Rupert Sheldrake

"A self worth having" is a fascinating synthesis of Nicholas Humphrey's ideas over the last forty years. Although I have known Nick throughout that entire period, I've never seen such a clear summary of his various approaches to consciousness.

What struck me more forcibly now than ever before was the resemblance between Nick's ideas about the nature of sensation and the philosophy of Henri Bergson (1859-1941).

Conscious feeling, Nick suggests is a "remarkable kind of intentional doing". Feelings enter consciousness not as events that happen to us but as activities that we ourselves engender and participate in. The basis of Bergson's view was that "perception is entirely directed towards actions". Bergson, like Nick Humphrey, saw perception as concerned with virtual actions. "I see that my perception appears to follow all the vibratory detail of the so called sensitive nerves; and on the other hand I know that the role of their vibrations is solely to prepare the reaction of my body on neighbouring bodies, to sketch out my virtual actions. Perception, therefore, consists in detaching, from the totality of objects, the possible action of my body upon them." [*Matter And Memory*, 1911 edition].

Again, like Nick Humphrey, Bergson thinks a key ingredient is the enclosure of time within the brain, through memory. "The growing complexity of the nervous system shunts the excitation received onto an ever larger variety of motor mechanisms, and so sketches out simultaneously an ever larger number of possible actions." In relation to memory, "By allowing us to grasp in a single intuition multiple moments of duration it frees us from the movement of the flow of things, that is to say, from the rhythm of necessity...The memory of a living being appears indeed to measure, above all, its powers of action upon things." Perhaps it is this resonance with Bergson's philosophy that makes some people see Nick as a romantic.

However like Dan Dennett, I was not convinced by Nick's argument that consciousness gives us a reason for waking up every day. All other species wake up every day and get on with their lives without our kind of consciousness.

Nick emphasizes that "our experience of being conscious encourages us as nothing else could to take ourselves seriously as selves. We find new value in our lives, and, just as important, in the lives of other people." What he does not discuss is that in practically all traditional human societies, if not all, the experience of being conscious is associated with a belief that human consciousness is part of a much wider realm of consciousness which includes spirits, gods and God. This sense of being special goes beyond the nearly human realm. Human beings feel special because they share in something that goes beyond their usual limitations, linking them to other realms of consciousness, experienced in moments of insight and ecstasy.

In Nick's interesting theory of the placebo effect, he draws attention to the importance of a mechanism for keeping something in reserve, which the placebo effect partially overrides. When I worked in the field of tropical plant physiology, I came across an interesting parallel for this process. In annual plants, which die after fruiting, there is no

need to keep anything in reserve. They give all they have got, continuing to form fruits until they run out of resources, with the result that the later-formed fruits get smaller and smaller. By contrast, in perennial plants, which need to keep something in reserve for the next year, the fruits formed early and late in the season are more or less the same size. Perennial plants yield less than their full capacity, because they hold back reserves for the following season. We are like perennials.

[Natika Newton](#)

Dennett congratulates Humphrey on showing us the "early parts" of his thinking in this essay. But there are even earlier parts that I wish were reflected more here. He refers to his theory that "conscious feeling ... is a remarkable kind of 'intentional doing'." Conscious feelings, he holds, are activities, not events that happen to us. Humphrey mentions possible mechanisms underlying the experienced "thick moment of the conscious present": that reentrant circuits constituting conscious states originated as bodily activities, and that evolution has modified these circuits so that they are now internal. The thickness of the present emerges from an "as-if" time dimension in which the signals for the fossil-like action patterns are monitored.

Now Humphrey leaps ahead of that proposal to a teleological account of temporal thickness: it evolved because it enhances our appreciation of a valuable selfhood. It may do that, but there is so much more of interest to be said about action circuits themselves, that it seems both premature and disappointing for him to turn to adaptationist speculations at this point (the related objection by Dylan Evans is well-taken). It also seems, at this stage, unnecessary. If having conscious experience is acting, then there is no deep mystery about how (or why) consciousness evolved. The structure of current action, taken together with the reentrant loops of perception in creatures like humans who rely on perception to create action plans, predicts a temporally extended conscious present.

There may be in internalized remnants of overt perceptual acts, which contribute to the subjective feel of consciousness. But even without these, if consciousness itself is an activity, then that activity seems sufficient to account for temporal thickness.

The claim that conscious feeling is an activity can be ambiguous. I believe, and think Humphrey agrees, that at every real-time moment in conscious experience we are engaging in some action or other, which constitutes the experience. If so, the structure of the action is the structure of the experience. The action need not be bodily. In attending to any stimulus, one is motivated in such a way that the stimulus is salient with respect to some current goal. Conscious attention to the stimulus would be a central part of the ongoing action, during which the subject/agent has in view both the previous stages of the action and the desired or anticipated goal. Perceptual response to the stimulus may involve internalized ancient action patterns, but it also essentially involves current action patterns.

One has a goal — to find food, to satisfy curiosity drives, even to protect a restful state

against perturbations — to which the stimulus is relevant, with a positive or negative valence. Consciousness of the present includes retention of the immediate past (at minimum for the purpose of context and continuity) and also anticipation of the immediate future, as well as alternative longer-range futures (at minimum for evaluating the affordances of the stimulus). If what we are conscious of is our activity — our agency, some say - then all three essential parts of the action must be contained in our experience: past, present, and future. It is the past and the future of the current activity of conscious agency that creates the temporal thickness that Humphrey, in my view, correctly identifies as the basis of consciousness. If our experience of the "now" is extended in such a way that qualitative properties of perceived objects can take on a substantive and enduring presence, then it is at least plausible that this (illusory) presence is what gives consciousness its ineffable and unique features.

It is important to recognize the motivation behind identifying the thick present with action stages. It is just as Humph says: we need to understand "why evolution [has] taken this remarkable course." The function of causing us to value our selves cannot be the whole story; we know that evolution builds on existing structures, so what structure could underly the tripartate nature of present experience? If present experience is the experience of agency, then the structure comes built-in. We are not forced to hypothesize tiny internalized remnants of ancient perceptual activity, although these may well be components of ongoing experience.

[John R. Skoyles](#)

Is Nick Humphrey a fox or a hedgehog (to use Isaiah Berlin's famous division of intellectuals)? His diversity of ideas hints at him being a fox. But maybe Nick under his fox-like guise is a hedgehog with a big vision. Blind sight, sensation, Machiavellian intelligence, the nature of consciousness — and now placebos — below the surface of Nick bubbles an imagination and curiosity befitting a Hobbes, a Kant or a Descartes. Nick, I am sure will be embarrassed by that suggestion, as he will by the following one: that he is staking out a path in the study of the mind and body that might do what heliocentrism did in an earlier age for the study of the heavens. But rather than put the Earth on the periphery, Nick's new 'heliocentrism' does this for inputs and outputs.

This is a very different vision from making them primary as in bottom up models. Descartes does this when he explains pain as working mechanically like a "man pulling a bell rope and ringing a bell in the tower". This image from the physiological exterior to the mental top has been the "intuition pump" (to use Daniel Dennett's fine phrase) that has ruled the imagination of those examining the mind for nearly the last three and half centuries. An input X activates some physiological process Y that ends up with a mental result, Z. A "fire particle" to use one of Descartes' examples, triggers a pain receptor which transmits an impulse that then gets felt as a sharp burn. A bacterial antigen activates a T cell that produces interleukin-2 that results in us feeling sick. Cocaine gets snorted, transferred to the brain and activates dopamine receptors — and we feel a pharmaceutical high. And so on.

But bottom up mechanics is not the only way to describe a mechanics of the mind.

Engineers have created top down machines such as the Segway self-balancing scooter, and psychology textbooks have long illustrated perception with pictures of black splotches on white backgrounds that suddenly when given a title jump out as a face or a Dalmatian dog standing amongst leaves. Neuroscientists are indeed increasingly revealing the brain to be a top down not a bottom up machine: the lateral geniculate nucleus (the way station from the retina to the primary visual cortex) receives more axons from the visual cortex than it projects up to it. Zapping the parietal cortex at the right moment with TMS can block processing downstream in the visual cortex. But when theorists think mechanically (which is the only proper way to think if we are to enlarge science) they do so bottom up. Nick does not.

I have only space to show this for his work upon placebos and the "health management system".

At present, physiologists and neurophysiologists study the body and brain in terms of mechanical bottom up processes, not mechanical top down ones. Mental beliefs and expectations thus have no place in their models about the processes that lead to such things as pain, feeling unwell or the pleasure of an illegal drug. As a result, the fact that people's beliefs about medical inert pills prescribed by doctors effect their body and brain gets ignored. It just is not in the physiological scientific script much in the way that the fact that the coasts of Africa and South America snapped together like separated jig-saw puzzle pieces did not fit in with that of pre-plate tectonic geology. But yesterday's overlooked anomaly can turn into today's boring triviality. And while Nick's health management system theory is no plate tectonics (at least yet), his top down approach to placebos gives them a plausible place in science.

Nick's revolution is to put together facts with some commonsense.

First, much of what we call illness is not due to physiological dysfunction but the body's own self-healing. Temperature rises as part of a strategy to kill pathogens — pathogens if they had a choice would keep temperature stable. Likewise, we feel sick to conserve energy — pathogens do not in themselves make us feel groggy. "Sickness" therefore is not so much a physiological input as a physiological output.

Second, such top down self-healing is organized by goals that have been finely tuned by evolution to maximize our long term survival. The body, for example, needs to be run on the assumption that it can dip into reserves if the unexpected happens — hungry *Panthera leo* tend to turn up when they are least expected. And the body might not bother with self-healing if it can see it will get better by doing nothing. There is thus not one path to health but a variety of options. The health management system makes a top down choice upon them — and that introduces the possibility of such decisions being shaped by beliefs.

If you get infected with a rhinovirus, your body has choices. It could decide to go on an all out defence that ups body temperature, diverts energy resources into T cell attack, and ends up with fatigue and 'queasiness'. But it might also decide to tolerate the bug and gamble it will not cause major damage. That latter option is not as rare as might be expected 20% of two year old infants in one study (van Benten et al., 2003, *Pediatric Allergy and Immunology*, 14, 363) were found to have rhinovirus infection but none of the physiological responses, 'illness', that go with clearing respiratory viruses. The

health management system can decide to fight or to leave alone. Thus if a person's health management system predicts the body will get rid of a virus without needing to fight it physiologically then they might decide to abort having a "cold". It does not take much for that to happen — swallowing a few tablets prescribed by a doctor is enough. It does not matter that in reality that they are inert sugar pills — if the health management system thinks it has had external help (even if it does not — as with a placebo), it can decide to turn off having a cold.

Obviously much more exists to Nick's approach the main point is that he proposes a mechanical top down model where previous theory, in so far as it has existed, has been bottom up. As I have noted Nick's other work such as upon consciousness, sensation and the Machiavellian mind also has this top down way of tackling problems. It is his hedgehog vision. Functional brain imaging is increasingly revealing a brain with a top down modus operandi. Thus we can expect Nick's theoretical explorations to be viewed increasingly as important — if not for what they say, then for where they point.

[Geoffrey Miller](#)

Consciousness as a corporate pep rally.

I've argued in *The Mating Mind* that many of our communicative abilities — language, art, music, humor — evolved as a sort of genetic marketing strategy. They show potential mates how great our genes are, how well our brains work, how ingenious and charming we are. Brains that sell themselves best in the mating market leave the most copies of their genes, and that's enough to drive mental evolution. Our minds are there largely for marketing our genes.

Yet my mind-as-marketing argument missed something. True, every large corporation needs marketing experts to lure potential consumers, investor-relations experts to lure potential investors, and public relations experts to influence the media. But corporations also need a sense of subjective well-being, an internal sense of morale formalized in their mission and vision statements, embodied in the esprit de corps of employees, and reinforced by within-company newsletters, intranet web sites, team-building retreats, and corporate pep rallies.

Nick Humphrey's "A Self Worth Having" nicely makes the point that human consciousness may have evolved to serve this sort of motivational function in the multi-billion-neuron corporation that we call the human brain. Consciousness may be the brain's ongoing corporate pep rally, in which its diverse agents and adaptations are continually reminded that they work for the common good of the company (the human phenotype), its profitability (reproductive success), and its stakeholders (genes).

Rupert Sheldrake, Dan Dennett and Dylan Evans are doubtful that the human brain needs such an aggrandized sense of subjective richness to prosper, when all other species do so well without such puzzling phenomenology. Do we really need subjective qualia to keep our neural circuits from becoming lazy and phoning in sick, like disaffected employees? If selfish genes can build adaptively selfish brains in every other

species, why would we need the extra corporate propaganda provided by a unified subjective consciousness thickened in time, ornamented by sparkling qualia, super-charged by a sense of self?

One possible answer concerns the selfishness-eroding power of advanced mind-reading abilities and human empathy. Insofar as humans evolved uniquely potent ways to understand the beliefs and desires of other people — especially friends, family, mates, and offspring — our self-interest tends to get spread pretty thinly through our perspective-taking. Social understanding brings compassion, and compassion may dilute adaptive self-interest. Perhaps subjective consciousness evolved to keep our own interests salient, front and center, when our mind-reading abilities would otherwise undermine our genetic narcissism.

Consider the phenomenology of extra-marital lust, for example. Ideally, married humans develop deep empathy with each other and a sense of shared interests that promote efficient partnerships and parenting. Yet with such benefits come possible genetic costs: over-developed spousal empathy may deter adaptive affairs that produce more or better offspring with other partners. The hominid who felt total psychic unity with her long-term mate may miss valuable opportunities to copulate with superior short-term mates. What she needs to override her strategic social mind-reading is a consciousness that is salient, self-centered, and seducible — a private romantic getaway from her social, spousal, and parental duties. She needs a sensorium in which a new lover's sweet talk can make her forget her husband's potential jealousy. Likewise, her husband needs a sensorium in which his empathic fidelity to Hilary can lose itself in the warm, wet immediacy of Monica's mouth.

Equally, both sexes need a subjective consciousness in which immediate pain and distress can over-ride empathy with a lover's irritabilities, addiction, and rages. Who will survive better: the battered wife whose mind-reading leads her to over-identify with her abuser, or the battered wife whose subjective pain makes her forget the drunken bastard's point of view and empowers her to leave? In these ways, one could imagine an evolutionary arms race between altruistic mind-reading empathy and selfish sensory consciousness.

I'm not sure this consciousness-versus-empathy model really works, though. There must be simpler, cheaper ways to turn off mind-reading when it's maladaptive. There must be simpler, cheaper ways to remember that one's genetic interests are centered in the experiential opportunities that surround one's own body.

Maybe the corporate newsletter is a better metaphor for consciousness than the corporate pep rally. Maybe consciousness is about the consistency of one's marketing messages, rather than the intensity of one's motivation. Most companies now use the IMC (integrated marketing communications) approach to selling products, which emphasizes the need for a unified corporate and brand image across all of one's promotions, packaging, publicity, advertising, sponsorships, public relations, employee relations, investor relations, and government lobbying. The company's right hand must know what the company's left hand is doing, and this is hard to achieve in multi-national companies with thousands of diverse employees. If BMW's South African advertising portrays the company as technically proficient, BMW's customer care call center in Delhi

better have a phone system that works well. If Monsanto's British press releases portray genetically modified crop seeds as magnanimous, pro-social benefits to humanity, its security sub-contractors protecting its experimental crop fields in Kansas better not beat up any anti-GMO protestors there. The more complex the company, the more communication channels there are, and the more diverse its stake-holders are, the harder it is for every part of a corporation to stay "on message".

The human mind has evolved the greatest complexity, the most diverse communication channels, and the most diverse set of social relationships of any animal. Therefore, we face the greatest challenge in maintaining an integrated marketing communication strategy — a credibly consistent way of presenting ourselves to family, friends, mates, and other stake-holders in our well-being. Consciousness may help solve this IMC problem by working as the brain's corporate newsletter. Consciousness needs subjective time to be "thick" because it needs a news cycle of optimal length - not too fast to integrate a decent amount of information and analysis, not too slow to respond to emergencies. Consciousness needs to feel like a unified "Cartesian theater" so the mind's diverse adaptations all stay "on message", responding to the same set of challenges and opportunities with a unified action plan that stays consistent with the corporate image and mission statement — the sense of self.

[Kevin Kelly](#)

I really enjoyed Nick Humphrey's original posting on "A Self Worth Having", and was further impressed and enlightened by the many notable responses to it, but it wasn't until Nick unloaded his detailed counter-reply that I finally thought I understood what he was trying to say. In particular, the latter half of his reply, where he comes clean with his larger vision — that's when I lit up. Something clicked and then I combined this aha with a throwaway point of Marvin Minsky's about the necessity of obscurity.

So now I have my own version of what I think Nick is saying. Hopefully he can correct me if I'm way off. The summary as I see it:

The self is a survival mechanism for high intelligence. A rapidly expanding intelligence growing in power and dimension would be awfully confusing to a growing mind, particularly if it offered multiple views, and well... out of the body experiences. An emerging sense of self would be calming, focusing and attractive, a place to rest from wild ideas (like death and pain), allowing a more stable and effective human to survive.

An increasing intelligence would also reach a point where it became aware of its own intelligence — and that's a highly dangerous spot because an intelligence that was naked and transparent would be susceptible to intellectual manipulation. The first thing that a mind smart enough to see itself would do is start to hack itself. This is Minsky's point. A "thick" self that was unexplainable, obscure, and mysterious would prevent an emerging mind from subverting itself. This shroud of mystery would help the mind to grow more powerful, while keeping it attractive to its host (there's a lovely me inside here!), and protect intelligence from its own inclinations to "improve" things. Those humanoids with

selves would survive longer, and the selves over time allow intelligence to spin in many directions.

Intelligence is the mind's worst enemy — a power so lethal it must be cloaked inside something unintelligible so it can't be easily broken by understanding, and at the same time be so emotionally appealing that a mind gravitates toward this cloak of unintelligibility because it feels more and feels good. A phenomenal self is that enabling cape. We don't see selves in other animals because their intelligence has not yet reached a point that it needs cloaking.

If the thickening self is a way for intelligence to protect itself from intelligence, then it may be but a phase, a kind of scaffolding that can be shunted away as the mind advances to explore safer ways to manage the disruptive power of intelligence. A self may not be needed for an AI beyond a certain level, although it may be necessary to reach that level.

My bet is that should we ever engineer that choice for ourselves (to clear the cloud of "I"), we'll keep a mysterious self because, as Nick points out, we've co-evolved so that we (I mean our selves) see the self as the most beautiful thing in the universe. We really do. And of course the self tells us that we are crazy to let it go.

And we probably are.

[Nicholas Humphrey](#)

If Dan Dennett and I are bookends, then we are bookends that meet up round the back. Though I may be more impressed than he is by the "magical qualities" of consciousness, I count myself as good a functionalist and anti-mysterian as anyone — and have learnt many of the philosophical moves from Dan. In fact if I were to set out to criticise my own position, I would do it in very much the terms that he does here (which is why, in so far as I continue to disagree with him, it is not for want of seeing his side of the argument).

The commentaries are provocative and helpful. (What have I done to deserve such a kind reception? Dan Dennett *and* Rupert Sheldrake both being positive!) Since Dan's commentary came first and set the tone, let me respond to him and pick up others' contributions along the way (though there's no way I will get to all of them).

Dan says I have read more into Frege than is warranted, and maybe I have. True, Frege never actually said that a subject has to be the subject *of* something. What he did say, in "The Thought: a Logical Inquiry", is:

It seems absurd to us that a pain, a mood, a wish should rove about the world without a bearer, independently. An experience is impossible without an experient. The inner world presupposes the person whose inner world it is.

So, I agree, Frege's point (like Kant's, much earlier) was that subjective experience requires a subject, not that a subject requires subjective experience. Still, I'd say the second point is not only just as logical as the former, but in some ways more important — at any rate more psychologically salient. Because, for anyone who reflects on it, it leads beyond the dull analytic conclusion that "I have such and such experiences, therefore I am" to the much more interesting revelation that "I am because I have such and such experiences." And, at the level of individual psychology, this surely is — or can be — headline news: the discovery that "*This* is what it means to be me!"

I'm not saying of course that most people (let alone most animals) articulate this discovery. But the revelation — at whatever level it gets through — that "This is what I am" is nonetheless something with the potential to inspire new forms of self-interest and self-respect. Thus I can (and indeed do) *like* being the subject of my experience. I can be amazed, proud, tickled to be "the person whose inner world *this* is".

I don't agree for a moment with Alva Nöe, Jesse Prinz or Thomas Metzinger (or Hume) that "no such things as selves exist in the world". I think selves exist just in so far as there are subjects who recognise certain sets of intentional states *as their own* — which is to say, since in my view it all boils down to action, subjects who represent themselves as the *agents* or *authors* of those states. In fact (though I realise this begs some questions) I'm happy to make this the definition of a "self": namely, that my self *is* the representation in my mind (which is certainly part of the world) of "I", the author of my intentional states, continually-updated.

I'd ask you to note, *pace* Nöe, that this set of mental states -- the states of which *I* am the author, as distinct from the states of which *you* are or *he* is — constitutes much more than a *mere* Humean "bundle". In fact this set of states is quite obviously a well marked natural kind: every mental state that's mine is in it, nothing else is (which is arguably about as well marked and natural as any kind can get.)

The set may include, of course, a variety of different types of mental state: sensations, percepts, thoughts, desires, beliefs, volitions, higher level thoughts about sensations, thoughts about thoughts, and so on. I think it's fair to assume that a self of sorts could be constituted through the subject recognising his authorship at any of these levels. (In fact I've argued elsewhere that a single individual may originally, soon after birth, house several different types of self in parallel, and that they come together only as and when they turn out to be engaged in a single project). But it's the thrust of my new argument, as I outlined it in the Interview, that not all such selves would be equally capable of motivating adaptive thinking and behaviour — or as I put it, be selves "worth having". In this regard, I contend that the *author of phenomenal sensations*, the *subject of the thick moment*, is uniquely privileged.

Now, Dan and others challenge me to be more specific about how an interest in this type of self, let's call it the "phenomenal self", translates into biological survival, and at what level of the evolutionary tree I suppose that it kicks in -- earthworms, foxes, homo erectus, Bertrand Russell?

My answer is that I see it as happening in several stages, with the phenomenal self revealing latent strengths as new demands and crises have occurred. As the saying goes, there are horses — or rather selves — for courses. Different types of "I"s come into their own as and when different kinds of biological and cultural niches open up opportunities for them.

I've argued before that the neural architecture that is required for "thick sensations" evolved only with the mammalian brain (though there may have been a similar development in birds). So earthworms are out. But foxes are surely in. I'd guess that a fox does indeed have a phenomenal self, and that this self is an asset to the fox's planning for its future. Nonetheless the "I" of a fox has a much less demanding role than the "I" of a modern human being.

Still, Dan , endorsed by Dylan Evans and Rupert Sheldrake, chides me for arguing that, even at the most basic level of the life-game, human beings are in a different league. What's good enough for foxes, they suggest, ought to be good enough for us. Why should human beings need *novel* incentives to, as I put it, "get up in the morning and get on with life"? But the *reductio* of this suggestion would be that what's good enough for worms is good enough for foxes, and what's good enough for bacteria is good enough for worms — so that all that's ever been required for an organism to get on with life is a basic "life instinct". And this is absurd.

In reality, new challenges — and new temptations to abandon the struggle — must have arisen with each advance in biological complexity. The fact that we and other living species are here today is testament to the fact that we have each evolved to find, on our own terms, reasons to carry on. And I think it's obvious that *our* terms, *human* terms, are quite unlike any other.

Dylan hints that I may have a hidden agenda when I talk about human "reasons to live". And (as he knows, since he and I have discussed it) actually I do. But this is not the place to start on my ideas about the deep (million year deep!) history of existential anxiety (which soon enough I hope to turn into another book).

Suffice it to say that I believe that about 50,000 years ago, the human species faced a crisis: human beings were in danger of becoming victims of their own mental evolution. Under pressure from *ideas* they were beginning to *lose heart*.

Human minds had been extending progressively — and safely — into areas never yet visited by our ape ancestors. But everything changed once intelligence and culture crossed a certain threshold. The critical event was the development of a mind that, on the one hand, demanded *meaning* and, on the other, was capable of the dreadful realisation that *human existence ultimately has none* — that life ends as nothing. From that point on, no one was safe from the destructive self-questioning: *why bother? what's the point?*

What saved us? I can do no more than tease you with my answer here. To balance the reasons for embracing death, human beings had to discover a good new reason to value life. And they found it, I believe, right in front of their nose: through reflection on the nature of selfhood.. In short, we were saved by phenomenal consciousness — or, at any rate, by a new-found relation to it.

The philosopher Tom Nagel has put it like this: "There are elements which, if added to one's experience, make life better; there are other elements which, if added to one's experience make life worse. But what remains when these are set aside is not merely *neutral*: it is emphatically positive. Therefore life is worth living even when the bad elements of experience are plentiful, and the good ones too meagre to outweigh the bad ones on their own. The additional positive weight is supplied by experience itself, rather than by any of its contents."

But *why* is conscious experience emphatically positive? Why *should* it be? Ah, that's the question (or, as Dan might say, the wrong question).

In the Interview I said, perhaps incautiously, "the more I try to make sense of it, the more I come back to the fact that we've evolved to regard consciousness as a wonderfully good thing in its own right — which could just be because consciousness *is* a wonderfully good thing in its own right!"

Dan and Dylan take me to task for this talk of "good in its own right". Values, they point out, are *relational*. Something is good or bad in the estimation of someone, not good or bad intrinsically. I agree of course. So let me explain better what I meant.

Because values are indeed relational, there can be *two* ways by which something gets to be valued in the course of evolution: there can be evolution of the evaluator and/or evolution of the valued thing. But if and in so far as it is the second way — the thing as such has evolved — then I think it fair to say the thing has come to be valued "in its own right."

Consider, for example, why honeysuckle is attractive to bees. The answer may be (a) that bees have evolved to like the look and smell of honeysuckle, *and/or* (b) that honeysuckle has evolved to have a look and smell that appeals to bees. In the latter case honeysuckle would be attractive to bees in its own right.

So, then, consider again why consciousness is seen as a wonderfully good thing by human beings. The answer may be (a) that human beings have evolved to like the peculiar qualities of consciousness, *and/or* (b) that consciousness has evolved to have peculiar qualities that appeal to human beings.

What excites me is this second possibility. (Dan — but probably nobody else — will see how it might fit with his own idea of consciousness as a self-serving meme) . Lets suppose that human beings, far back in time, had minds predisposed to regard anything with counter-intuitive, magical (yes, "magical") qualities *as special and worthy of preservation* . [Listen to GM Hopkins' haunting words: "All things counter, original,

spare, strange; Whatever is fickle, freckled (who knows how?) With swift, slow; sweet, sour; adazzle, dim. He fathers forth whose beauty is past change. Praise Him."] Then, lets suppose that genes that help create phenomenal consciousness were to have got a foothold in such a predisposed-for-mystery human mind.

The human beings who carried such genes for phenomenal consciousness would, *ex hypothesi*, have taken themselves — as hosts to something so remarkable — more seriously as selves . Indeed the more mysterious the qualities of consciousness, the more seriously significant the self. In which case, it's easy to see how the very qualities of consciousness that philosophers such as Colin McGinn and David Chalmers deem "metaphysically problematical" would have been the occasion for consciousness becoming a runaway evolutionary success. In fact these qualities would soon have been *designed in*.

And yet, contra-Dennett, I'd stick with my point that, because humans who valued their consciousness would have enjoyed greater biological fitness, this would have been a symbiotic mutualism between the host organism and the new genes — not a parasitic exploitation.

So, the long and the short of my argument is this. (I'm glad to see that Alva saw it coming). *There has very likely been positive selection in the course of evolution for precisely the qualities of consciousness that we have most trouble in explaining..* In the jargon of computer programming, the baffling qualities of phenomenal consciousness aren't "bugs", they're "features".

Back to [A SELF WORTH HAVING: A Talk with Nicholas Humphrey](#)

EDGE SUMMER READING

"An engrossing treat of a book...you'll have a wonderful time reading these reminiscences"

— ***New Scientist***

"...inspiring glimpses into the minds of today's scientific intelligentsia" —

Publishers' Weekly