

Professor Bill Newsome

Lecture and Dinner/Discussion, Tuesday, 27th April 2010

Denis Alexander: Now we come to the real work of the evening. When we hold these evenings we ask someone to give a short response to our speaker and this evening Alasdair Coles, who is a University lecturer in neuroimmunology and consultant neurologist at Addenbrooke's Hospital, has very kindly agreed to do that for two or three minutes. That will also help to get the discussion going. So Alasdair thanks very much, over to you.

Bill Newsome: Alasdair is a neurologist as well as a neuroimmunologist, right?

Alasdair Coles: That's correct Bill, yes.

Bill Newsome: And we have a few other neurologists at the table?

Alasdair Coles: Yes, we do; so you know, Bill, we have a collection of people here who come from neurology, neurosurgery, neuroscience, psychology, and a few waifs and strays who are hanging on in there!

So, Denis can I first of all thank you on behalf of everyone here for inviting us. These are always very pleasant occasions and it's very good of the Faraday Institute to see the value in collecting people like us together and I hope we reward you with the quality of our conversation. And Bill can I thank you as well on behalf of everyone here for a very dense and thorough exposition of the range of difficulties that we face as we address the issue of brain and mind and God. As you say, this is not a small topic, so I have just got a few things to say, really to launch us, if I may.

Your talk was divided into three sections. Firstly, on the issue of the substance of religion and the substance of science and to what extent those are compatible. Can we take your lead and say that they probably are, and let's move on.

The second section, on the attitude or the approach to science and the approach to religious thinking, I thought was really very interesting and you ended with this conclusion – and I think I have got this roughly right – that the more important the subject and the issue is for us as humans, the less likely it is to be amenable to scientific enquiry. I thought that was a very interesting kind of provocative statement; I'm sure there's some tongue-in-cheek there, and I want to return to that if I may.

Then of course you had this prolonged exposition on reductionism and you went through the various proposed positions, though you're not happy with all of them and there were a couple of reasons for your dissatisfaction. One is the position I guess most of us would be comfortable with, which is that there has to be some traffic in both directions, "bottom-up" and "top-down". I guess any neuroscientist who has kids and sends them to school will have to accept that the environment influences our brain and our brain influences our actions. But the other thing that you said almost *en passant*, which I want to pick up, is that there are some propositions that you just felt unhappy about, you were just dissatisfied. I think you even said "It doesn't do it for me" – there's that kind of gut response – and it's brave of you to say that and I'd think we would all sympathise with that, but I want you to flesh it out a bit.

You ended with a proposition that there's causality in organisation and that then brings me back to my main question, which is around the issue of what we call "free will" or "autonomy" and just what that means. I guess in a minute, when we've had a chance to pool other comments, it would be good to hear exactly what you think that is, because I guess if you asked most people in the street out there about free will or autonomy they would want to say "I'm expressing my free will when I, as a whole person, feel able to choose from a number of different options and my whole person is choosing this one, but I could easily have chosen that one". There's a sense that this is a deliberate act that he or she, as a whole person, is doing. You may or may not agree with that, but if that's true, if that is what free will or autonomy means, then what would that look like if it was reduced to a neuronal

firing activity? How could that ever be expressed at any level other than the whole person and is this a debate that could ever be answered in a reductionist approach? That's one question.

My final question is to go back to this point about the most important decisions we make in our lives, about who we marry and whether we stay where we are or whether we uproot our family and go to a better job, the examples you gave; and indeed whether to live or to die. So when people talk about free will and autonomy I guess there's a sense of a very conscious decision based on an accumulation of facts and the best understanding of this situation and trying to come to some rational approach, but you're raising an altogether different level of decision-making. Could that ever be reduced to brain activity, or is that an altogether different thing that we need to distinguish from the more conscious, deliberate rationalising?

So that what was I took away from your dense and attractive and stimulating talk. But I know that there are other people here who inject drugs into people, who operate on people, who look at people with damaged brains, which I think is an interesting aspect to all of this, and there are those of us whose job it is to determine the capacity of people to make decisions and to exercise their free will. I'm sure they will have views on the talk, so thank you very much.

Denis Alexander: Thanks Alasdair, very much. I think it would be best just to throw it open. To give some sort of focus to the discussion I think it's probably good if we don't talk about everything all at once but there might be one particular thing that Alasdair has said that someone might like to pick up and comment on. If we could then have a few contributions we could then come back to Bill for a response. Would anyone like to carry on the conversation, perhaps starting with reductionism as that's our theme?

Daniel Wolpert: I wanted to follow what Barbara mentioned when you gave your lecture. There has been some really interesting experimental work done on free will from Libet and

other people following up on that work. I would like to know is whether they can do it for really complex problems. Would it really surprise you if things were predictable far back in time from what would happen in the future? What experiment would amaze you and how would it change what you have taught us tonight? Would those things ever happen? It's something I would like to discuss.

Bill Newsome: I'll think about the answer to that while we have the next point!

Denis Alexander: Your subconscious will work on it ...!

Nicholas Gibson: You didn't like free will and you wanted to use autonomy and self-determination instead but I have found myself struggling to know what your definition of those things would be and quite a lot hinges on how we define those things for understanding free will and whether or not we have it.

Barbara Sahakian: Just a couple of comments really that I think are interesting in the light of that. One is the concept of hot and cold cognition, hot and cold decision-making: the idea that dorsal lateral pre-frontal cortex is involved in more cold decision-making and planning, but the orbital frontal cortex and areas, which have strong links to the limbic system, are involved in hot-decision making. So for instance, cold decision-making would be planning out a schedule, what are we going to do, is there enough time to get to the bank, and so forth. But the hot-decision making is the kind of question that you raise: do I want to marry this woman, do I want to go out drinking with my friends when I know I have an exam the next day. (*Bill Newsome: Or how am I going to answer that question that Dan Wolpert has just asked ...!*)

Barbara Sahakian: So I guess it's kind of bringing up the idea that these sorts of things can all be incorporated. The big questions may be in the other less important questions, say in the concept of the more basic gut reaction that we have about making decisions.

The other comment I wanted to make was on the autonomy and the interesting questions you raised about responsibility, because for instance we know that a lot of work goes on now with the MacArthur Foundation network and so forth on substance abusers and whether they have a controlling responsibility, given that they are addicted to drugs; and will the courts view them differently because they are addicted, or do they have the same responsibility that everybody does. So these are really just comments and reflections on the subject.

Denis Alexander: I think we had better let Bill respond before the number of points becomes too great!

Bill Newsome: Well, I have a number of thoughts which tend to be scattered so I'm afraid I'm not going to be truly pertinent perhaps to anyone.

First of all, as a neuroscientist I am totally committed to the idea of mechanisms. This is why I don't like the term "free will" and I tried to emphasise that in the talk today because for so many people "free" means "uncaused", as though somehow our mental space is unhinged from the underlying neurons, and that decisions made in the mental space (where we *really* live) give commands which the neurons then follow. That's exactly what I hope to avoid; that's why I don't like the term "free will". So I believe totally in mechanisms and the fact that when I think something, when I decide something, there are neural mechanisms that underlie those decisions, whether the decision is "hot" (impulsive, or emotionally-driven) or "cold" (logical, sequential analysis). I think that 80 or 85, maybe even 90% of our behaviour is emotionally-driven; I just think that's who we are as humans. Sequential, rational analysis is a relatively small part of our decision-making behaviour. Of course, "hot" and "cold" aspects of decision-making are not mutually exclusive; we have emotional predispositions but logical analysis can operate on those predispositions. But one day, when we really understand these things, I am convinced that all forms of decision-making will have underlying neural mechanisms. So I am deeply committed to mechanism.

What I am resisting, and what we all probably need to resist if we are to make any sense of our lives at all, is the idea that once we understand the mechanisms underlying some phenomena, we can throw away the phenomena; that somehow the lower level mechanism comprises a deeper, more comprehensive truth that totally subsumes and makes irrelevant the higher level. This is a “fundamentalist” style of reductionist thinking that ultimately must reduce all phenomena to the wave equation of quantum mechanics to achieve comprehensive understanding.

My problem is that the wave equation knows nothing about the people in this room, their motivations for being here tonight, the smell of the food, the taste of the wine, etc. The wave equation only knows about the motions of atomic particles! If we believe that a comprehensive understanding of events in this room must include knowledge of people, motivations, smells and tastes, then we must abandon reductionist fundamentalism. The only way I see to take those high level entities seriously is to say that a “mechanism”, properly understood, is reciprocal – it includes both the phenomenon that we are trying to explain and the organisation and actions of the underlying parts. And causality runs in both directions in a mechanism. We can intervene at the behavioural level (change someone’s beliefs) and cause changes in the constituent parts (how muscles move during subsequent actions). Conversely, we can intervene in the constituent parts (electrically stimulate key parts of the brain) and cause changes at the behavioural level. This notion of mutual manipulability is what Carl Craver calls “making a difference”. This way of understanding mechanism and causality seems to me appropriate scientifically, and seems congenial humanly. I admit that I don’t have this idea spelled out in a terribly precise way; but my sense is the truth has to lie in this direction.

To Daniel’s question about what would really surprise me and make me change my notions, I don’t know how to answer that. Quite honestly I don’t know what to make of Libet’s experiments. Those experiments are flawed in ways that have been pointed out by many people, and subsequent experiments have difficulties as well. Part of me wants to go

and do some experiments along these lines myself because I think some fresh minds working on this problem might be able to take the field to a new level.

Some of our dinner conversation revolved around the issue of consciousness and whether it can be accounted for reductionistically. My own tendency is to put consciousness off to the side because it is profoundly mysterious, and because I can think more productively about the issues of causality, mechanism and the reciprocal nature of manipulation.

In fact I have no clue how masses of neurons firing action potentials can generate subjective consciousness. I tend to ignore it while I consider things that I do have some clue about. Roger Carpenter – are you still here Roger? (**Roger Carpenter: I'm biding my time!**) Roger and I had a longish discussion about this at lunch today and Roger would argue exactly the opposite, I think – he would say, although he can talk for himself in a moment, that we can't interpret any aspect of existence except through the prism of consciousness because that's the one thing that we're absolutely certain of, and that has to be the fundamental principle. I am trying to talk about causality, mechanisms, etc, while putting consciousness to the side. Roger would say – in fact he did say it to me, though not in so many words because he's a gentleman – that's a fundamental mistake.

Roger Carpenter: I don't think I said that really. If I may, what I would like to do is clarify in my mind at least one aspect which I found harshly compelling in what you are saying. (**Bill Newsome: I consider that high praise from you!**)

You said I think quite rightly that you can't just describe everything in terms of the bottom-up determinism of the bottom layer. What I would say is that there are other ways of looking at it, not *actually* alternative causative mechanisms but you describe them as *effectively* causative mechanisms and you can think of causality working at what *you* call the top autonomous layer. But this is what I wasn't perfectly clear about: are you saying that in that autonomous layer determinism no longer applies?

Bill Newsome: No, I don't want to say that – if I did I didn't mean to.

Roger Carpenter: Well in that case in what way does that differ from ..

Bill Newsome: I don't have a problem with determinism. I don't have a problem with causes and I'm trying to say that over and over again. For me, free will is not the same as "uncaused"; the critical question is what counts as cause. Is causality only operative at the fundamental level – the primeval forces of gravitation, and strong and weak nuclear forces – or is there a way to talk sensibly about non-fundamental causes, for example at the behavioural level? Let's be as clear as possible. Meaningful autonomy does not include random behaviour in the quantum mechanical sense. I *want* my behaviour to be determined precisely by the things that are core parts of me, such as my history, my memories, my background, the values that I have adopted, and the influences that all of you will have on me this evening. If I remember or am affected tomorrow by anything that happens here tonight, it will be because *you* changed the structure of my brain, right? That's causality and that's part of me and I embrace that. So I'm not arguing against causality, I'm arguing about what counts as cause. I think that mental states – that's a weird term that's a substitute for "higher levels of organisation in the nervous system" – are causal. And if those things are causal then I'm OK, because those things are me. Does that help?

Denis Alexander: Does anyone else want to come in on this?

Michael Shadlen: You made, and you are making again, several points about levels of explanation. Your approach was epistemological – essentially an argument about how we know, and how we express ideas. I hear you saying that some ideas are more conveniently understood at a macroscopic level. But there's a more forceful argument that's been made by Ernst Mayr – do you know of his work? (**Bill Newsome:** *Is he an evolutionary biologist?*) Yes he is. Mayr thinks these levels are more than simply convenient. They have more fundamental status in biology, although we tend to ignore this because the philosophy of science has been dominated by physics. Biology introduces its own, very different principles

like variation (rather than prototypical ideals), evolution, natural selection and so forth. Mayr thinks it's not just a matter of convenience when we choose to explain something at a macroscopic level – in the sense that it is merely inconvenient to explain billiard balls with quantum mechanics, although you could. It may be inconvenient to explain chemistry with quantum mechanics – but there is nothing lost. In biology, however (including biology of mind), there is no way to reduce some principles to lower levels – what you call fundamental reductionism – without losing something essential. I think if you buy his argument, which I am not attempting to flesh out here, then different levels of explanation really do have their own fundamental place at the table of explaining causal relationships. And if so, responsibility and perhaps free will may not be explained away or cast as merely emergent properties of the brain's nuts and bolts. Perhaps this is what you mean when you refer to higher-level entities, like mental states, as causative. The philosopher Robert Kane refers to “self-defining actions” in his writings about free will. When you and I study the neural basis of decision making, say how speed and accuracy are traded off, I think we are providing grist for theories like those of Kane.

So while I am sympathetic to your perspective on high-level causes, I do not follow the next step of the argument that connects to religion. When you consider your top level, why do you need to invoke religion? I recognize that neuroscience – not just simply because it's in its infancy but also because of the kinds of simple functions we focus on in the lab – does not provide explanations and answers to life's deeper questions. But I fail to see why this implies that religion provides those explanations and answers. I really don't know why Leo Tolstoy doesn't have the same status as John Chapter 6. To me they are equals; they are both very compelling sources to cite in an argument.

Bill Newsome: I'll respond to that if I may. So on point one about Mayr. In some ways I think that's in essence what I was trying to say only I wouldn't just make a distinction between biology and physics – I would make it about every level, that every level has its own causal integrity and its own conceptual integrity. I don't know if any of you have read this,

but there's a book out by Robert Laughlin called "A Different Universe" and the sub-title is "Reinventing Physics from the Bottom Down". Laughlin is a physicist at Stanford who was a co-recipient of a Nobel Prize for superconductivity. The whole thrust of his book – he's a maverick and enjoys it – is about how reductionism doesn't even work in physics. He claims that you can't even reduce statistical mechanics arising from the physics of individual atoms in a truly comprehensive way to the motions of individual gas molecules. I'm not physicist enough to evaluate these things, but I would tend to side with Laughlin in saying that fundamentalist reductionism is not just a problem in biology. I'm sceptical that that kind of reduction works even in any area of science.

Concerning your point about religion, religion doesn't answer questions for me about mechanism. Religion answers questions for me about *me*. My faith is a tapestry with many threads, all of which compose a deep intuitive judgment concerning what my life is about. The tapestry includes judgments about whether or not values and ethics have solid grounding in the ontology of what's real about this universe, and whether there is anything about our existence that is really worth believing in and caring about. You and I have had discussions about these issues in the past and we are unlikely to agree about them tonight, but I think that these ultimate beliefs about what is real in our world actually matter. If our universe is all contrivance, then it seems to me that values are not matters of ultimate concern at all, but are simply contrivances for getting along reasonably conveniently as social animals. Unless moral principle lies somehow at the heart of our universal reality, I see no way to argue that one set of values is better or worse than another, except for reasons of convenience. Of course, it may be true that no such universal principle exists, and that we must make our way through the chaos of our world as best we can, preferably not thinking too deeply about the futility of it all. But there is a way of believing and thinking about the Universe and existence that endows much more dignity than that, much more hopefulness than that and much more of a clarion call to an ethic that helps us all be better than we would be otherwise. So I think that religion makes a difference; it makes my life

better and it makes a difference in many other lives too. Religion can get sick and it can get idiotic as we all know, but in its best form it holds out a hopeful and a positive vision of existence and it helps me personally be a better person than I would be otherwise. Mike can make the counter argument, which he does very effectively.

James Rowe: To come back and pick up on one of the questions you mentioned just now in your talk which is what counts as a cause? And then the proposal, the challenge, that understanding the causal nature of autonomy or seemingly free choices is one of the most important challenges for a neuroscientist, at whatever level he chooses to operate. You then proposed this heuristic, multi-level, explanatory systems neuroscience, which has an immediate intuitive appeal to me as a neurologist. When one's work is sometimes at a pastoral level, or involved with psychological mechanisms of motivation or perception, and sometimes in terms of genetic predisposition and the disease or anatomy, one moves to the clinical and scientific gap between these levels. So one may then say well, that's heuristic and it's interesting, but actually if one's trying to chase down what counts as a cause, one can avoid the issue by hiding in these many levels. Whatever level you are choosing to operate at you could say, well, the answer lies above or below. In such a slippery hierarchy you can never actually make any progress that pins down the nature of causes.

How can you help me out about the problem, where would I find an answer?

Bill Newsome: That's a very penetrating question. That is the thing that worries me most – that talk of levels and reciprocal causality is still a bit vague and slippery. We conjure up a few concepts and metaphors but when we try to put our finger on the precise argument, it evaporates. So I worry about that and I accept it as a criticism, but the problem is that I keep coming back to it and having to grapple with it because I don't see a good alternative.

If I try to be a hard ass and go with the reductionist fundamentalist, saying that I'm only going to attribute causal efficacy to physical forces that actually do work and which I can measure in a laboratory, then I'm running down that reductionist slope to quantum

mechanics. When I get to the bottom of the slope, it's not a satisfying place to be because I've lost too much that I care about. I get the wave equation, yes. But as I said in the talk, I lose all sense of the people in this room and their intellectual curiosity. People are not in the wave equation – only particles and their probabilistic motions. So however much I try to be a hard ass reductionist like that, it doesn't work in the end, because it doesn't correspond to the richness of life as I experience it. So I have to rethink and ask myself exactly what I've lost with the wave equation and why. That kind of reflection, I think, provides a real clue as to how we should be viewing and talking about reality, causality, mechanism and reduction.

The talk I gave today is the closest I can get to an answer at this point, and it is not so much an answer as it is a potentially useful perspective. I'm not delighted with where I am, and I'm certainly not ready to write a book about it because it still isn't as precise as I want it to be. If somebody has a better idea that avoids the poverty of fundamentalist reductionism while still taking mechanism seriously, I'm all ears.

The most precise concept that I've been able to come up with – and I'm not good at this because I'm not good at this kind of math and this kind of talk – is the notion of organisation or, in the abstract, information. The organisation of causally-interacting parts gets to the heart of things in some sense. If the organisation and interactions of all the synapses and neurons in my brain could be reproduced in silicon, I think that silicon entity would be conscious and in some sense me. At the same time, however, I believe that the silicon entity, if detached from a body that permits interaction with the world, would *not* be me. The brain is fully embodied and I take the notion of embodied cognition quite seriously. But to return to the main point, the notion of organisation may be the key thing that is lost during standard scientific reduction. Perhaps that is why the person disappears somewhere during the reductionist march to the wave equation.

I don't know if that does any good but that's about the best I can do right now.

Denis Alexander: Maybe there will be some contributions from people just sticking on this organisation as it were as the causal agency. It would be interesting to hear from the psychologists.

Mark Kotter: As a scientist working with stem cells and a neurosurgeon, I am trying to comment on an area with which I am not very familiar. Nevertheless I would like to ask two questions. My first question is: do you propose that a specific architecture, or a program perhaps, is what defines us as a person?

And the second question is: what is your view on religion? Where does religion fit into this concept? Is religion part of a computer code – like a sub-routine of a program – and are the values that you've mentioned also part of a script? I suppose what I am referring to is Richard Dawkins's concept of "memes". If this were the case then one could potentially replace any particular sub-routine by a different set of beliefs and values. Or do you believe that religion represents and connects to absolute values and ideas and therefore is fundamentally different from the concept of "memes"?

Denis Alexander: It sounds like you should respond to that, Bill.

Bill Newsome: Well, that's another very interesting question. It's one that's been explored in science fiction and in movies, but I think we can do a little thought game that may illuminate our intuitions on the question of "self as architecture".

If I remove one neuron from my brain and substitute for it a silicon chip that perfectly mimics the original neuron's interactions with all other neurons in my brain, it seems to me clear that this hybrid organic silicon entity will still be me. My life won't be different just because one of my neurons is made of silicon, rather than carbon, hydrogen, oxygen and nitrogen. Then of course we just extend the thought experiment one neuron at a time, until all of my organic neurons have been replaced by silicon neurons. Is there a point where I am no longer me because my brain is now 51% silicon and only 49% organic molecules? I don't think so. If the essential organisation – or architecture – is maintained, then the

functionality will be maintained, including my mental life in all its richness. So yes, I tend toward the idea that the organisation, or architecture, is what's essential.

Mark Kotter: So is our identity basically a complex computer program?

Bill Newsome: I don't know if I like this word "program" because this implies pure information to me. Obviously, there is critical information stored in the brain, but there must also be a way for the brain to interact with the world. We are not only the information in our brain, we are an entire organism that is constantly interacting with the world by virtue of the information in our brains. If we simply isolate the information and remove the dynamism of interacting with the world, then we don't really have an organism any more.

Take Lakoff – is that his name, the psychologist at Berkeley? He's a linguist who keeps pointing out how much of our cognition is metaphorically based on actions. If we want to think about some problem, for instance, we say "I'm going to chew on that a bit". There's this analogy with chewing, an analogy between a bodily function and a mental function. Lakoff has dozens and dozens of examples of this sort of embodied cognition. The fundamental notion is that the brain is not up there in the cranium in isolated splendour, but it's constantly feeding back and forth with the rest of the body. That's the one thing that gives me pause about this whole idea of downloading my mind to a computer. I think we can reproduce intelligence in a computer, but the way that computer thinks would become different, using its own experience to create metaphors based on its interactions with the environment.

Roger Carpenter: So, paradoxically, if you replace your brain with silicon you're still you, but if you kept the brain as it was and replaced your body with nuts and bolts you're saying you wouldn't be you any more?

Bill Newsome: That's a really interesting point. I think there is something to that.

Michael Shadlen: I'm not sure I understand what you're trying to convey with the cell-chip replacement scenario. Replacement of one thing by something functionally equivalent may indeed have the status of a conceptual metaphor in the Lakoff argument you mention. The scenario is useful the way trolley problems are useful in ethical discourse. Both serve as rhetorical devices that invite clarification about what we mean when we use certain expressions like 'greater good' in the trolley problem. I hear your cell-replacement example as a tool to help us discuss what we mean by identity of the agent (what it means to be me vs. you). But why would such a replacement exercise have ontological significance? Why should we assume that the neurons (or their replacements) hold the key to this mystery? Why not, for example, the history of the actor that owns the neurons and neuron substitutes?

Bill Newsome: I think my point there, Mike, is that it's a way of emphasising the deep importance of organisation, that there is something about that organisation that is independent of the individual parts that contribute to the organisation. I would say that's an ontological point.

Denis Alexander: I'd like to bring in Ahmed here.

Ahmed Mohamed: I am interested in the relationship between neuropsychopharmacology and religion. There's been a lot of research that indicates that people with mental health problems, addiction for example, are helped by religion if they attend prayers or if they belong to some religious groups. For example, cocaine abusers tend to significantly reduce their cocaine intake if they belong to religious groups. In the same way, there are medications that significantly reduce cocaine intake in cocaine abusers so I am interested in how religion and neuropsychopharmacological drugs (medication) are somehow achieving similar aims. I mention this because someone mentioned that religion does not explain mechanisms but I'm interested in the fact that, similar to medication, religion does have significant impact on the brain and its neural wiring. I wanted to ask whether you agree, Bill, because you said that you can take an electrical node from your brain and replace it with

another metal one so I am thinking of the parallel connections between religion and medication (i.e. if you replace the aspect of what religion is doing in the brain to a chemical import, for example a drug, whether they will achieve similar outcomes. Do you see my point?

Denis Alexander: That's a great comment and we'll just hold that for a second. Fraser, do you want to take us off in a completely different direction? Barbara, do you want to come in here?

Barbara Sahakian: Yes, it was somewhat related to that because it goes back to the discussion just before and links into Ahmed's comment. There is this issue of personhood and attention deficit hyperactivity disorder where frequently the most common treatment would be a stimulant drug, usually Methylphenidate or Ritalin. Some children report that when they're taking the drug that they are themselves, that's who they see themselves as, because maybe they have friends at school and the teacher likes them better and they seem to think that it realises their potential, so that's how they see themselves. Whereas other children see themselves as the real people off the drug, and they take the medication because they are told to but that's not who they see themselves as. So it's a little bit like this flip state where you can actually be two people at once, depending on whether you are on or off the drug.

Bill Newsome: I think some addicts describe that as well. They can stay clean for some period of time but who are they really? Are they the person with a rush or with a community of other drug users, is that who they identify with most deeply? Or do they identify the person off the drug and working, doing a job and making an income, doing what all these other people do, the alternative reality. I think there are flip states like that.

I think that's a really interesting point about the psychopharmacology and religion. Again, I'm a mechanist at heart and I think that the effects of religion and religious faith are real and that anything that has real effects on people – their emotions, their aspirations and

their behaviour – is going to have an effect on the brain. There are mechanisms that mediate those effects.

I have no problem with the fact that there are connections between religious experience and brain function. It would shock me if an experiment showed that there were religious experiences that don't have correlations with brain states.

Dan Wolpert: Do you think there's a gene for religion?

Bill Newsome: A *gene* for religion? We don't even have a single *gene* for blood pressure, how can we have one for religion? Multigenic is the byword. The links between genes and religious behaviour are likely to be terribly obscure; the links between brain activity and religious behaviour should be less so. It's often pointed out that people with certain psychiatric illness, such as temporal lobe epilepsy, exhibit hyper-religiosity. One well known faculty member at Stanford will drop this factoid like a little bomb on a bunch of freshmen and sophomores and state that maybe all religion is just a disease of the temporal lobe, which I think terribly illogical. People with diseased brains can become obsessed with all sorts of things, which does not mean that any one object of obsession is a disease of the brain.

Dan Wolpert: There are experiments, aren't there, where you try to simulate that, where you can actually induce those effects in people?

Bill Newsome: That would not surprise me. For any aspects of human experience that are well documented and real, there will be an underlying neural mechanism, so that's not shocking. It would be shocking if it were any other way.

But let's stick on this point for the moment. The fact that religion can get sick, and that hyper-religiosity can be associated with pathological states of the brain, is no more surprising than the fact that romantic love can get sick, and that obsessive romantic involvement can be associated with different parts of the brain.

As far as I can tell, all areas of human endeavours can get sick if the brain gets sick. And individual brains can change and become sick if immersed in a cultural environment where crazy ideas prevail. Mob psychology is called mob psychology for a reason – the group effects of being in a mob change the way brains of individuals are working. People exhibit behaviour in mobs that they would never exhibit otherwise. The mob exerts causal efficacy! So I totally embrace the notion that if religion makes a difference in our brain and our behaviour, there are going to be neural mechanisms involved. I also embrace the fact that religion can get sick in diseased brains – just like patriotism and romantic love, and any number of other things. This is not a controversial area for me.

What's more interesting to me is this notion of "top-down" causality. I'd be interested in what some of you people in the psychiatry field think about the point I raised during the talk about cognitive behavioural therapy in combination with drugs being a more effective treatment for depression than either alone. We have a "top-down" treatment and a "bottom-up" treatment, and both are causally efficacious. I understand that there is currently an experimental programme going on right now with hallucinogens as a treatment for depression – there was recently an article on this in the *New York Times*. It was about inducing hallucinogenic states as a spiritual treatment for depression and a number of other diseased states, and people believing that their experiences on these hallucinogens in well-controlled experimental circumstances have been the most significant experiences of their lives. Have you seen this stuff? It's amazing.

Barbara Sahakian: I'll comment on the top down stuff. We do experiments where we look at depressed individuals and healthy individuals and we give them some negative feedback. When a healthy person gets positive feedback obviously they respond to that, but when they get negative feedback they try to take out the information from that negative feedback and try to ignore the fact that they have been told that they're wrong and they haven't done it right. They try to dampen down the emotional component of the feedback and just pay attention to the information.

What we find is that, in depression, people can't do that as they are just overwhelmed by the emotional part which you can actually see in the brain when we image people doing this task. Healthy people exert top-down control from the pre-frontal cortex so they have cortical control over essentially the medulla which they have to deactivate to dampen down the emotional component, so it's very clear that you are getting cognitive control over the emotional part of the brain in order to be able to keep your emotions properly regulated. So it's almost certain that things like anti-depressants can obviously restore your neurotransmitter serotonin if it's in SSRI and put you in the right disposition to re-learn and learn that things are better, but ultimately in cognitive behavioural treatment, or some other psychological treatment, you need to adjust your thinking processes and think of other strategic ways of approaching things or viewing things cognitively

Bill Newsome: So another thing that drives me back away from the reductionist fundamentalism is the clear efficacy of these top-down influences.

Nicholas Gibson: Can I respond to that? Like you, I like the idea intuitively that these top-down interventions work in a top-down way. Just because I like the idea, and can understand what I think is going on intellectually, it doesn't mean that that's how it actually is. (*Bill Newsome: That's very true*) So isn't it actually quite plausible that all of this is happening at the quantum level and I might not be able to describe it in ways that were meaningful to me? That doesn't mean that that's not how it is.

Fraser Watts: I would like to mention this "multi-level, systemic top-down, bottom-up" – I'm entirely on side with all that. But I'm uneasy when you start to introduce the concept of mental causation and I'm not sure that you really mean it. I think you probably ought to abandon it along with free will, if I've understood your position correctly. I don't think you're going to abjure this position in which you are saying there are physical causes, and then there are mental causes that are completely separate from that.

What I assume you are saying is that there is a certain category of physical cause, neural cause, which also gives rise to consciousness, beliefs and where other things are going wrong as well. So it seems to me there's a particular neuronal level that we need to talk about in a double aspect way – of course that's a more complicated way of talking about things and so mental causation is a convenient shorthand, but I don't think you really mean to talk about mental causation, I don't think you really mean to talk about the effects of beliefs as though that were something different from, separate from, the effects of the brain.

Bill Newsome: When I talk about mental causation or the importance of beliefs I am talking about them in their instantiation within the brain. I'm not talking about them as a totally separate realm, where logic gets worked out and then you have to figure out a way to get that down to the brain. So I agree with your position. Do you have a shorthand terminology for this? How do you talk about such things?

Fraser Watts: I don't have a shorthand, which I think is why we fall back on this other terminology. It's just useful to note that it's potentially misleading and can be highly misunderstood.

Bill Newsome: Yes, I take that criticism. Craver talks about non-fundamental causation and I lapsed into the mental causation because it was a title of a very influential book fifteen or so years ago that many of you know about, but I certainly don't mean to imply that that's in a different realm separated from the physical.

Denis Alexander: I think Nick's point was that although it's intrinsically attractive, at the end of the day there are not sufficient grounds for believing it. Do you want to comment on that?

Bill Newsome: Somebody said that there is pleasure in understanding, but unfortunately the pleasure in misunderstanding is frequently as good as the pleasure of understanding. I think that is what you are getting at. I certainly concede your point that the reductionist-fundamentalists might just be right, but I just can't convince myself that that's the case. Maybe I should try harder but it is very difficult for me to believe that once I reduce things

down to quantum level, I have really encompassed reality in a fundamental way that should lead me to abandon the epistemological conveniences of the higher levels.

Nicholas Gibson: It depends on what you mean by abandon. So if you mean your conscious mind abandoning them, no you wouldn't because you can't understand the meaning or belief that our minds are actually ..

Bill Newsome: Then I have no reason to take quantum mechanics seriously, right?

Nicholas Gibson: Just because we can't engage with those ideas at the fundamental level doesn't mean that they can't be expressed and ultimately reduced to the quantum level.

Bill Newsome: Here's another thing I would say about reduction. Recall the slide in my talk (adapted from Craver's book) that showed the various levels of organisation that are relevant to understanding spatial memory. The top level was that of the organism – a rat learning to navigate a maze. The next level was circuits of “place cells” in the hippocampus which we believe to be the immediate neural substrate of maze learning. Next was the level of synaptic modification at which mechanisms such as long-term potentiation (LTP) change the strengths of synapses, resulting in “hard” changes that underlie long-term memory. Then came the level of NMDA receptors and coincidence detection mechanisms that initiate LTP. And finally was the level of gene transcription, by which new receptor proteins are made and inserted into “strengthened” synapses. Each of these levels is highly relevant to the neuroscientist who wants to understand the biological mechanisms of spatial learning. But of course there are many more levels of organisation above and below this particular window. At higher levels, spatial learning contributes to organismal activities, such as finding one's way home at the end of the day and engaging in social activities with other people.

Below the level of gene transcription lie the vast levels of organisation of chemistry and physics. And here is the key argument. Even though in principle it might be possible to reduce the gene transcription and ion channel operations to lower levels of chemistry and

physics, this reduction will shed no further insight into the phenomenon of spatial learning. Lower level processes of physics and chemistry are, of course, necessary for the hippocampal circuits to implement spatial learning, but explicit knowledge of those lower level processes are in fact unnecessary. Trying to understand spatial learning at the level of quantum mechanics would not take us closer to the truth, but further from it. The key mechanistic insights concerning spatial learning are at a higher level of system organisation.

Nicholas Gibson: But isn't there a distinction between whether it's a useful level of description and where causality originates? (**Bill Newsome:** Yes) and judging by what you say, and I completely agree with you, you seem to be saying because it's not useful and because in fact it's actually intellectually impossible for us to understand the higher level phenomenon especially the psychological one at the level of physics, then that can't be.

Bill Newsome: So this is probably getting to the nub of the question in some sense. When I talk about causality or causal relevance at higher levels, I'm not talking about new physical forces. As far as I know, the only physical forces that exist are the four mentioned earlier – gravity, electromagnetism, and weak and strong nuclear forces. When I talk about causal relevance and causality at higher levels, I must return to the notion of the neuron. A neuron can produce action potentials because specific cellular and molecular components are precisely organised in space and time to generate the dynamic phenomenon that we call an action potential. It is this functional organisation for specific purposes that is lost in the quest to reduce all phenomena to fundamental physics. Most scientists, I think, would agree that we should not try to reduce everything to quantum mechanics because doing so would be far too inconvenient, and perhaps impossible in practical terms. But I think the problem goes much deeper than that, as I indicated earlier. Notions of organisation, function and purpose have absolutely no place in quantum mechanics. If we want to appreciate what an action potential is, and how it contributes to the functioning of a central nervous system, it is not merely a matter of convenience that we avoid reduction to quantum mechanics, it is essential for achieving the understanding we desire.

Daniel Wolpert: But all that comes from those four forces – there has to be some control of them if you want self-determination.

Bill Newsome: I agree, but I don't think we have to talk about self-determination to understand the need for "control of those four forces". We only have to think about action potentials, which we believe we understand pretty well. If I want to explain the action potential, I must invoke something in addition to those four forces, and that "something in addition" is system organisation – how molecular and cellular components are organised in space and time to enable the action potential. It is the system organisation that harnesses the fundamental forces to achieve higher level goals, such as transmitting information via the action potential from one part of the cell to another.

Daniel Wolpert: If you believe that potential exists or does not exist, there are two possibilities and, yes, then I think there's some control over those forces.

Bill Newsome: The control comes from the organisation. Am I not being clear?

Daniel Wolpert: I don't understand how that can set down different paths.

Bill Newsome: Yes, the organisation of the system will determine exactly how inputs will be transformed to outputs. We simply just adjust the action potential threshold for a single neuron. The same synaptic inputs will produce a different output. So that means that the organisation of the components "makes a difference" even though we are not postulating any novel forces of physics.

Daniel Wolpert: That's just saying the state of the systems affects the way it will respond. .

Bill Newsome: So you're using the word "state" like I am using the word "organisation": do we really have a difference here?

Denis Alexander: We're about to wrap up and will have to make this the last comment, I think.

Jean-Pascal Pfister: You use the word organisation. I would also use the term “information” or even “information processing”, which seems less static. Our brain is an amazing information processing unit: it processes the high volume of sensory information into sensible motor output commands. This level of description (information processing level) seems to me fundamentally different from – but not incompatible with – the biophysical description. Even if we can completely describe all the forces of physics in action, this does not tell us how information is processed. (This is the analogy you gave during the talk about the computer having the hardware and the software).

[The following passage was not part of the round-table discussion, but came out later in a private discussion with Bill Newsome:

As you stressed during the talk and during the discussions, the real question after all is: what counts as a cause? Currently, I see three complementary answers to this question, but since I am a beginner in this field I would definitely value any comment. First, there is what could be called the “natural cause”. My actions can be understood by external observers as a result of the interactions between my embodied brain and my environment through biophysical processes. This is the “it” story, because I am studied as a mechanistic object. Fundamental reductionists would stop here and claim this is the only valid cause to consider. Because I am not a fundamental reductionist, I think it is equally relevant to consider the “I” story. This is what could be called the “personal cause”. I can understand my actions as a result of my own conscious decisions, e.g. “I felt hungry, but I decided not to eat this pudding because I want to lose weight.” Because I am at the same time the agent and the observer (I am conscious of my decision), this “I” story is a fundamentally different description from the “it” story. Finally, if one believes in an omnipotent God, which I do, then one has to accept that nothing happens outside from God’s will. This could be called the “divine cause”. Of course, there is a lot to say about this and I am also aware that many would disagree with this last point, but after all this is a faith question. In my opinion, focusing only on one of those three causal aspects and neglecting the other ones gives an impoverished view of the human being.

Let's take an example to illustrate those three causal descriptions. Christians consider the death of the Christ on the cross as a central event. What was the cause? Well, the jealousy of the priests and the greed of Judas led the Romans to nail him on the cross, which caused his death through suffocation. This is the "it" story. Secondly, Jesus repeatedly explained that he deliberately agreed to give his life ("I" story) in obedience to his Father's salvation plan ("God's story"). If we remove any of those three aspects, then this central event loses its significance.]

Bill Newsome: I would go further than saying it's a description. This comes back around to the point that Mike made about epistemic function – that software is an epistemological shorthand for a bunch of hardware states inside the computer. But I think the software probably deserves ontological status. The program is real and it is multiply realisable. The program can exist in any of many computers, it can be stored as magnetic particles on a disk, or it can be printed out on a piece of paper. Whatever form it takes, it is the same computer program and I think that's not just epistemic convenience, I think that somehow or other the program is a "real thing".

James Rowe: Just to pick up on the discussion about control and organisation and the fact that distribution of different categories about belief are not acknowledged as different layers in this hierarchy. It's very helpful in many ways, but one is it allows added value and added process of organisation which may convey additional information over and above every single layer, but it depends whether the problem about what do you do when different layers of the hierarchy disagree. For example, the creationist beliefs accompanying the religious system of the highest strata conflict with evolutionary principles at the intermediate strata and you say somewhere you do believe in it. Is there an actual or isn't there; is there creationism or is there evolution? These levels in hierarchy may conflict. How do you then get a control principle or resolution?

Bill Newsome: I don't know! Levels of a hierarchy conflicting – I suppose that's true, I must think about this. When you say levels of hierarchy for example, I immediately go back to the examples of spatial learning, the hippocampus, LTP and so forth, and I guess it is true that one could create conflicts within this system. We could perturb the system by inserting signals at lower level using drugs or electrical stimulation, or some other tool that mucks about with the innards of the system. But I suspect that any conflict would be resolved by accompanying changes at the higher level of the system. I am not sure how this would apply to your example of creationism and evolution.

Denis Alexander: I think that that's a good moment to wrap up. I think we ought to thank everybody for all their contributions. We've really worked Bill very hard and I think we ought to thank him very much.

Bill Newsome: I want to thank you all. I'm just astounded that a roomful of highly competent people would sit down and listen to what I have to say about these topics, so it's very humbling and gratifying. Thank you so much for being here and for this conversation which is very helpful to me. Thank you.