Tony Chadwick Essay Prize 2006 Winner
Can we Save 'Qualia'?
(Thomas Nagel and the 'Psychophysical Nexus')

By Eileen Walker

§1. Introduction: The problem of causal exclusion

If our minds are part of the physical world, mental states must be physical states. But are they the same thing as brain states? If so, our rich inner lives are simply electrochemical processes in the brain – a reductive view many reject. For this reason most philosophers have opted for various versions of non-reductive physicalism or property dualism.[1] However they have as much difficulty as did Descartes in trying to explain how our minds interact with the physical world.

A physical event, if it has a cause can only have a physical cause, since the physical world is a causally closed system.[2] If our minds bring about physical change in the world, which experience of human behaviour tells us they do, then our minds must be part of the physical world. However, if our behaviour is caused by the neurophysiological states of our brains, this alone is sufficient for causal closure. There appears to be no role for phenomenal experience. So-called 'qualia' or 'what it is like'[3] become redundant or 'epiphenomenal', a mere by-product of evolution like the red colour of blood. This is what Kim calls the problem of 'causal exclusion'.[4]

If mental and physical processes turn out to be *the same thing*, causal exclusion is no longer a problem. This was the solution proposed by the 'type-type' identity theorists of the 1950's.[5] Each mental token (instance) of a specified type is identical with a neurophysiological token of a specified type. Anyone with a pain will be in the specific brain state identical with that particular pain. The usual example given is the identity of pain with the firing of C fibres in the brain[6]. The argument runs as follows:

Scientific or 'theoretical' identities like *water = H_2O* are examples of *contingent identity*.[7] The identity is contingent because we cannot know its truth analytically or a priori like the truths of mathematics. The chemical constitution of water could have turned out to be something entirely different, hence the contingency. We therefore have *one* phenomenon here, but under two descriptions. The same is true of the identities which obtain between our mental states and our brain states, such as *pain = the firing of C fibres*. It is hoped that science will eventually discover these theoretical identities.

---

1 *Property dualism* : the view that brain states have two types of property, mental (subjective) *and* physical.

2 The sum total of physical causes of a particular physical event, all acting in conformity with the laws of nature, provides a sufficient explanation for that event. A *non-physical* event, if it were to intervene in the system sufficiently to bring about change, would have to violate the laws of nature. (See discussion in Heil, pp. 23-26). A *natural* law is inviolable. If one appears to be violated, this is evidence that we got the law wrong or chicanery is going on.

3 Recall that Nagel's first paper on this topic was 'What is it like to be a bat?', 1974.

4 Kim, op. cit. throughout. The view that physical causal closure leads to the exclusion of mental causation

5 It is the view of J.C.C. Smart which is considered here.

6 Heil points out that empirical evidence no longer supports this view, but for the argument, any specified neural state will do.

7 Smart p. 57-59 This is the view with which Kripke will radically disagree.

Kripke attacks the claim that theoretical identities are contingent.[8] Convinced that they must be *necessary*, he argues that since the relation between mental and physical states *is* contingent, it cannot be one of identity. Firstly, to demonstrate the necessity of theoretical identities, Kripke uses his notion of the 'rigid designator', a labelling device whose function is to pick out the *same* individual or natural kind in every possible world (every way our world might be or might have been). In the example *water = H_2O*, each term flanking the identity sign *independently* designates one and the same stuff in every possible world. This identity is not an analytic statement (like 'a vixen is a female fox') whose necessary truth is knowable *a priori*. It is nonetheless a *necessary* empirical truth, discovered *a posteriori*, like all scientific or theoretical identities. Both 'water' and 'H_2O' are rigid designators. They have different uses or connotations – an ordinary speaker would use the former, a chemist the latter – but they designate the same natural kind.

Although Kripke argues that scientific identities are necessary, and regards 'pain' as a rigid designator since it is *necessarily* painful, he claims that the mind-brain relation is *contingent*. Since one can imagine feeling pain in the absence of the firing of C fibres, mental and physical states are *logically* separable. This alone grounds his claim that the relation between mind and brain is contingent, so cannot be one of identity.[9] The difficulty here is that one can imagine all kinds of plausible but impossible scenarios – a bonfire on the moon, a flesh-and-blood zombie, just like you or me, but without qualia.[10] One cannot reach conclusions as to what *is* solely on the basis of what is conceivable. Because dualism is conceptually possible, it does not follow that it is true. The tantalising possibility that a *necessary* identity could obtain between a brain state and its corresponding mental state is one which Kripke does not explore, though Nagel will.[11]

Meanwhile, a more convincing - and non-Cartesian - argument that the relation between mental and physical states is contingent is provided by the notion of *multiple-realisability*[12]. A mental state must be realised in a physical state, but *which* physical state is not specified. According to type-type theory, anyone with a pain will be in the *specific* brain state identical with the type 'pain'. But in certain cases of brain damage, another part of the human brain takes over the function of the injured section. Creatures apart from humans suffer pain. It is unlikely that their brains will have evolved to produce the same brain states as ours to specify their pain. The identity theory must be modified to claim that every individual who suffers a pain will at the same time be in some kind of *token* physical state. The relation between pain and the physical state which 'realises' it looks contingent after all, supporting Kripke's view.

But things are not quite so simple. The characterisation of *pain* envisaged by the proponents of multiple-realisability is a functional one. Mental states are states whose *function* is to convert

---

8 What follows reflects the arguments of Lecture III of *Naming & Necessity*, though Kripke prepares the ground in Lecture II.

9 The argument is reminiscent of Descartes' thought experiment in Meditation 2

10 This is one of David Chalmers' key arguments in *The Conscious Mind.*

11 To be fair to Kripke, he made no claim to illuminate the mind-body problem which he regarded as deeply confusing. His concern was to refute the misconceptions of the identity theorists regarding necessity, contingency and identity.

12 Kim attributes the term to Putnam on p.3 of his book.

sensory inputs into behavioural output. *Pain*, like all mental states, is simply whatever fulfils the causal role of intermediary between environmental input (hand-burning) and behavioural output (withdrawal of hand from hot surface). Since a functional state is characterised by what it does, not by how it feels, sensations or 'qualia' do not enter into the picture. This is what Kripke cannot accept. His characterisation of *pain* is as a rigidly-designated sensation. Enter Thomas Nagel.

§2 Nagel's project

In a recent paper Nagel sets out the structural conditions for a new theory of mind which would not only establish a causal role for qualia but, more importantly, take account of their *intrinsic* nature. It would be a 'no priority' theory, incorporating both first and third person perspectives but privileging neither. Nagel's purpose is to realign our view of the logical relation between mental concepts, behaviour and brain states so as to see them all together in what he calls a *psychophysical nexus*. The bulk of Nagel's paper is taken up with challenging the logical basis of property dualism so as to prepare the ground for a new kind of identity theory. It is these arguments which are examined in what follows. The final part of his paper is more speculative, and on the basis of our newly-aligned mental concepts, Nagel hopes to erode McGinn's view that the likely fact of mind-brain identity is forever closed to us. That discussion must wait for another time.

Nagel uses Kripke's logical distinctions in his own postulates. He sets out the conceptual relations which must obtain between phenomenal, neurophysiological and behavioural states if we are to have any hope of coming up with a satisfactory theory of mind which is not reductive of qualia, and which does not lead to causal exclusion of the mental. However, he resists Kripke's conclusion that because scientific identities are necessary, the mind-brain relation cannot be one of identity since it is *contingent*. If the relation between mind and brain is to be regarded as a scientific identity – intuitively the most plausible view for Nagel – it must be one of *a posteriori* necessity. This is what a new theory would make clear to us, despite the lack of perspicuity between our first-person experience and third-person descriptions of observed behaviour and neurological processes.

§3. What is required for a new theory

(a) We must challenge arguments that claim to establish the *contingency* of the mind-brain relation. The most important of these are:

i) The multiple-realisability arguments of the functionalists. Mental properties are defined as properties whose function is to convert sensory inputs into behavioural output. There is no necessary link between a mental state and the physical state which realises it. The capacity to perceive an obstacle in front and move round it could be 'realised' by a human or a mechanical brain. Furthermore, on this view, mental states need not be conscious.

(ii) The thought-experiments of philosophers, notably Descartes, Kripke and Chalmers, that since mental and physical states are logically separable, there can be no necessary relation between them. Nagel claims that the appearance of contingency in these cases is illusory.

(b**)** If multiple-realisability holds, the neural states which cause a human, a dog or a robot to withdraw fast from a hot surface will not be the same. Pain may not be felt as 'pain'. The organism does not even have to be conscious. Phenomenal experience is irrelevant. We must therefore not only incorporate qualia within the functionalist framework, but explain how the *intrinsicality* of qualia fits into the overall conceptual picture. If qualia do have a function,[13] it is still a contingent fact that a phenomenal mental state is the state it is rather than a different state. Even if my *quale* of thirst gives content to my belief that I am thirsty and therefore induces me to go and get a drink, there is no reason why my thirst *feels* the way it does.

(c**)** If the theory is to embrace *both* first-person and third-person priority perspectives in concept-fixing, mental concepts must be redefined both in terms of their (third-person) functional role and their (first-person) qualitative nature.

To these questions we now turn.

§4 Nagel's proposals

Nagel's first draft of a realigned functionalism runs as follows:

Reference-fixing functionalism: The reference of our mental concepts to inner states is fixed by the contingent functional roles of those states, but the concepts apply rigidly to the occupants of those roles. It is neither necessarily true of a given mental state, nor analytically equivalent to its being the mental state it is, that it occupy a certain functional role, but that is how we in fact pick it out. Mental concepts rigidly designate states that are essentially physiological or phenomenological, or both.

To unpack this, let us take as example the mental concept of *thirst*.

(a) The state of thirst *happens to be* the mental state whose function it is to produce water-seeking behaviour.

(b) We stipulate the term 'thirst' as rigid designator of the mental state whose function it is to produce the said behaviour.

(c) It is a contingent fact that the feeling of thirst is what occupies the role of inducing water-seeking behaviour, but it is by this *intrinsic* phenomenal quality that we pick it out.

(d) 'Thirst' rigidly designates a mental state that is essentially phenomenal or essentially physiological, or both.

We thus have two reference-fixers of a mental state. One is the phenomenal quality which enables us *necessarily* to pick it out. The second is the contingent functional role by which it is initially 'named' or 'baptised'. Nagel points out that something further is needed to link these conceptually. The logical relations of mental concepts are thus spelled out as follows:

Though mental concepts cannot be analysed functionally, functional roles are needed to fix the reference of mental terms, because of the inextricable first-person/third-person character of mental concepts. It is a conceptual but contingent truth that each mental state plays its characteristic functional role in relation to behaviour. It is a conceptual and necessary truth that each conscious mental state has the phenomenological properties that it has. And it is a

---

[13] Sydney Shoemaker, 'Functionalism and qualia', *Philosophical Studies* 27, 1975, also considers this question.

non-conceptual but necessary truth that each conscious mental state has the physiological properties that it has.

Illustrating again using the concept of *thirst*:

(a) It is conceptually and necessarily true that *thirst* feels like thirst and nothing else. Recall that Kripke used the example of pain to make this point. We know this non-inferentially through first-person experience.

(b**)** It is a *contingent* truth that the feel of thirst plays the functional role of inducing water-seeking behaviour. (If evolution had gone otherwise, a different 'raw feel' could have fulfilled the same function. It is the multiple realisability problem.) However, it is a *conceptual* truth that thirst is what induces us to drink - we define thirst *as* thirst because of its behaviour-producing functional role. This is the third-person guarantee – the rigid designator the concept of thirst requires if we are to establish a relation of identity between its first-person phenomenal quality and its functional role in relation to behaviour. (This is the lesson of Wittgenstein's 'private language argument' – that private mental states require external corroboration)

Nagel's ingenious analysis does seem to offer a plausible solution to the notorious problems of functionalism – the contingent relation between mental state and functional role, and the failure to account for the intrinsicality of qualia. However, he has not yet tackled the far more intractable problem of the identity of mental and neurophysiological states. If there *is* a necessary connection between them it is opaque to our perceptions and understanding, so it cannot be a conceptual one. It is this very opaqueness which has encouraged some philosophers to argue that there *cannot* be such a connection. This is the position which Nagel challenges next. He argues that thought-experiments purporting to demonstrate that mental and physical states are contingently related are illusory. We look at these next.

§5 'What's wrong with the inconceivability argument'[14]

Thought experiments as to how things might be are very useful in enabling us to distinguish what is conceptually possible from what is not. Nagel gives the example of the inconceivability of a number having parents. We understand *a priori* why this is impossible – we see a contradiction between the conditions of numberhood and the biological conditions of parenthood. In the case of the mind and the brain, however, our first-person feeling of thirst is so very different from our third-person observation of the corresponding neural patterns on a brain-imaging screen, that we might be tempted to conclude that there cannot be a necessary connection between the two. If there were a necessary connection, we could not see it directly.

Philosophy is beset with such thought experiments as Descartes imagining he has thought but no sensations (intuitively highly implausible), Kripke imagining pain in the absence of the corresponding brain state, or Chalmers imagining a chocolate-tasting zombie isomorph of himself. Nagel's own examples are the zombie and the conscious robot. His arguments to support his intuition that these thought experiments are illusory run as follows:

Firstly, it was concluded from arguments earlier that there is a *conceptual* connection from an inner phenomenal state to its functional connections with behaviour. Nagel uses the example of smoking a cigar – we can perhaps use instead the taste of a glass of red wine! If I introspectively

[14] Nagel's own heading

identify this taste as I savour it, it is easy to *imagine* it in the absence of its functional and behavioural connections. However, this act of introspective identification is itself a mental act which cannot be detached from these functional connections, for example, the capacity to distinguish it from white wine or a nasty glass of plonk. If we can be prey to illusion in the simple case of an imagined separation of a mental event from its functional connections, it should warn us that Cartesian thought-experiments might be an equally unreliable guide as to the seeming lack of connection between mind and brain.

Secondly, Nagel points out that if we try to conceive what it is like to be a zombie, we put ourselves in the 'mind' of the zombie by a *sympathetic* act of imagination. We put ourselves in a conscious state *resembling* the thing imagined. This would be impossible without putting our brain in the corresponding brain state. We cannot detach the mind from the brain in imagination, because *if* the relation is necessary, we cannot refer to the one without referring to the other. If it is not, we cannot know that either. The lack of a perspicuous connection prevents us from seeing this. Of course this does not prove that the mind-brain relation *is* one of identity, but it does show that attempts to prove the contingency of the mind-brain connection are mistaken.

Conclusion

Thus far, Nagel has argued – plausibly - that one can establish a relation of identity between intrinsic phenomenal states and their behaviour-determining function, and we cannot discover *a priori* that the mind-brain relation is contingent. This opens the way to an exploration of mind-brain identity, but it is a long way from forming a conception that the relation might actually be necessary. What we need is a single referent identified rigidly by both its phenomenal concept and its physiological concept. Such a logical link could not be discovered directly, but would have to be established through the necessary link of each to the common referent.

The rest of Nagel's paper is taken up by a discussion of the logical, epistemological and empirical problems to be overcome if we are to come up with such a conception in a theory of mind whose laws would provide an entailment from the neurophysiological to the phenomenal and the behavioural. But that's another story……

Bibliography

David Chalmers	*The Conscious Mind,* OUP 1996
John Heil	*Philosophy of Mind: A Contemporary Introduction*, Routledge 1998.
Jaegwon Kim	*Mind in a Physical World*, MIT press 2000.
Saul Kripke	*Naming and Necessity* lectures 2 & 3, Blackwell 1980.
Thomas Nagel	'What is it like to be a bat?', *Philosophical Review* 1974, 4: 435-50: in many collections
_____	'The Psychophysical Nexus', pp. 433-457 sections I-V, in *New Essays on the A Priori,* eds. Boghossian and Peacocke, OUP 2000.
J.C.C. Smart	'Sensations and Brain Processes' 1959, *Philosophical Review*, AP 59; 68: 141-156; in many collections.