

Notes on “Why science could never show that conscious events do not cause brain events” by Richard Swinburne

Epiphenomenalism is the theory that no pure mental events (and so no conscious events including intentions) cause physical events. (I understand ‘causes’ as necessary causal conditions.) It assumes that pure mental events are not identical to and do not (logically) supervene on physical events, and hence is a scientific theory.

Three fundamental epistemic principles: (1) The principle of credulity: what we apparently experience is probably so – barring counter-evidence; this includes what we apparently observe in the public world, what we apparently experience as conscious events, and the logical consequences which we apparently ‘see’; (2) The principle of memory: what we apparently remember having experienced, we probably did experience - barring counter-evidence; (3) The principle of testimony: what people apparently are telling us that they experienced, they probably did experience, again barring counter-evidence. Beliefs acquired by apparent experience, memory, and testimony are probably true – in the absence of counter-evidence (=defeaters). If we have inferred the occurrence of some event  $y$  from present evidence  $x$ , then an undermining defeater is evidence (making it probable) that  $x$  did not occur or is not good evidence for  $y$ , whereas an overriding defeater is new evidence that  $y$  did not happen.

My epistemic assumption (EA) that:

(1) A justified belief in a scientific theory (which is not itself a consequence of any higher-level theory in which the believer has a justified belief) requires a justified belief that the theory makes true predictions.

(2) A justified belief that a theory makes true predictions is (unless this is a consequence of some other theory in which the believer has a justified belief) provided by and only by the evidence of apparent experience, memory, and testimony that the theory predicts certain events and that these events occurred.

(3) Such justification is undermined by evidence that any apparent experience was not caused by the event apparently experienced, any apparent memory was not caused by an apparent experience of the event apparently remembered, and any apparent testimony was not caused by the testifier’s intention to report his apparent experience or memory.

The fundamental criterion (FC) behind (EA) is that justified belief that some event occurred requires the assumption that that event is one to which the believer has privileged access or causes effects to which the believer has privileged access (unless it is justifiably believed to be the consequence of some theory which predicts events justifiably believed to occur on grounds independent of that theory). Then justified belief that a theory makes true predictions requires (unless justified by a higher-level theory) the assumption that both a scientist's awareness of the calculations that the theory predicts certain events and the events predicted are accessible or cause effects accessible to the believer.

(Unless his belief depends for its justification on some higher-level theory of which epiphenomenalism is a consequence) a believer could be justified in believing epiphenomenalism in virtue of having evidence of when (relative to brain events) various conscious events occur, which I shall call ( $\alpha$ ) type evidence. But such evidence could only be had on the assumption that epiphenomenalism is false. This remains the case even if we abandon [EA] and maintain only [FC].

Libet-type experiments have been interpreted as showing on type ( $\alpha$ ) evidence that intentions do not cause bodily movements. In the original and most influential Libet experiments participants were instructed to move their hand at a moment of their choice within a period (e.g. 20 seconds). They watch a very fast clock, and report subsequently the moment at which they first had the 'intention' to move the hand. They reported the 'intention' to move the hand as (on average) occurring 200 msec before the onset of muscle activity initiating the hand movement. However electrodes placed on their scalp recorded (on each occasion of hand moving) a build up of 'readiness potential' (RP), which was evidence of a particular kind of brain event (which I'll call  $B_1$ ) occurring an average 550 msec before the muscle activity. Experiments of other kinds led Libet to hold that subjects misjudge the time of all conscious events by 50 msec, and so he concluded that the 'intention' first appeared 150 msec before the muscle activation, and long after  $B_1$ . So, many have argued, Libet experiments showed that  $B_1$  caused the hand movement, and that the 'intention' was a mere epiphenomenon. But the most natural explanation of the experiments is to suppose that  $B_1$  causes the 'conscious intention' ( $M_2$ ), and that the intention causes the brain event ( $B_3$ ) which directly causes the movement. We could only have type ( $\alpha$ ) evidence that these experiments show that intentions do not in these particular circumstances cause bodily movements by assuming that normally they do cause them.

It might however seem that someone could have a justified belief in epiphenomenalism because of a justified belief in some general theory about the world of which epiphenomenalism was a consequence – for example a deterministic theory that every brain event has as an immediate necessary and sufficient causal condition some other brain (or other physical) event. That theory would entail epiphenomenalism. The only predictions we would need to establish such a theory would be evidence about which brain events occur when (relative to other brain or other physical events), which I will call evidence of type ( $\beta$ ). But in order to have a justified belief that such a deterministic physical theory gave true predictions about relations between brain (or other physical) events, we would need not merely a justified belief that certain relations between brain events occurred, but also a justified belief that these relations were predicted by that deterministic theory. But for that anyone would need evidence provided by the apparent testimony of scientists to have calculated this, or their own apparent memory of having calculated this; and relying on these sources requires assuming that epiphenomenalism is false. Only if a scientist could hold in his mind at one time all his calculations from which it apparently followed that the deterministic theory predicted certain events, could he have a justified belief that that theory made successful predictions, and so a justified belief in epiphenomenalism - which is a most unlikely event.