Do neural correlates of consciousness cause conscious states?

Juergen Fell*, Christian E. Elger, Martin Kurthen

Department of Epileptology, University of Bonn, Sigmund-Freud Str. 25, D-53105 Bonn, Germany

Received 20 November 2003; accepted 18 December 2003

Summary  Neural correlates of consciousness (NCCs) have been defined as neural systems and states of those systems, which are correlated with certain states of consciousness and which are minimally sufficient for these states of consciousness to occur. Here, it is discussed whether based on the identification of a NCC one may claim that this neural correlate causes the corresponding state of consciousness. In a strict sense, neural states and corresponding states of consciousness should occur simultaneously according to the definition of NCCs. However, two causally related events should occur consecutively in time according to the general view. Thus, within the NCC framework the question of a causal psychoneural relationship is only valid under the premise of a small and practically unobservable time shift between neural states and corresponding states of consciousness. The term causality is usually applied either in the sense of efficient causation or in the sense of explanatory causation. Efficient causality addresses the physical relation of two events and the criterion of an energy exchange between both events. Explanatory causality, on the other hand, refers to the law-like character of the conjoined occurrence of two events. Under a dualistic point of view, the interpretation that neural states effectively cause consciousness would not be justified, since effective causation, as understood in physical sciences, should be associated with energy transfer. But such an energy transfer between the domain of consciousness and the neural domain would violate presently accepted physical laws. Therefore, it is argued that one may state that neural states cause states of consciousness only in the broad sense of explanatory causality with regard to the observation of a law-like psychoneural relation. Within the viewpoint of identity theory, however, one would rather claim that certain states of consciousness basically are neural states.

© 2004 Elsevier Ltd. All rights reserved.

Neural correlates of consciousness (NCCs) have been defined as neural systems and properties of those systems, which are correlated with conscious mental states and which are minimally sufficient for the occurrence of those states of consciousness [1]. “Minimally” in this context implies that system constituents and system states are excluded if they are redundant for a state of consciousness to occur. The weaker criterion of sufficiency has been introduced by Chalmers instead of necessity, since it cannot be ruled out that two different neural systems are both associated with the same state of consciousness. Thus, they could not both be denoted to be necessary for the state of consciousness. In addition to minimal sufficiency another obvious requirement for a NCC is that neural states and corresponding states of consciousness should occur simultaneously. In mathematical notation, a NCC thus may be defined as the minimally sufficient neural system $S$ plus $\oplus$ a state of this system.

* Corresponding author. Tel.: +49-228-287-9347; fax: +49-228-287-6294.
E-mail address: juergen.fell@ukb.uni-bonn.de (J. Fell).
$N(t)$, which are correlated with a certain state of consciousness $C(t)$.

The identification of NCCs poses many problems, which will not be discussed in this article. But, assuming that it is possible to identify NCCs, does that imply that those neural correlates cause the corresponding conscious states? Psychoneural correlations are generally understood as cases of one-to-one correspondence (sometimes called isomorphism [2]), or in other words, as cases of a constant conjunction between certain neural states and certain states of consciousness. Constant conjunction is a term already used by Hume [3] to describe the relation between two events, which seem to have a causal relationship. According to the general view, two causally related events should occur consecutively in time. (Recently, it has been argued that causation may better be regarded as simultaneous [4]. But this argument relates only to a certain kind of propositions, such as "a force acting on a body causes a certain (simultaneous) acceleration of that body"). However, two correlated events should occur simultaneously according to the usage of the term conjunction underlying the definition of NCCs. Thus, by definition, one cannot infer a causal relationship between neural states and states of consciousness from the identification of NCCs. In practice, it is nevertheless impossible to exactly demonstrate simultaneousness of a neural state and a corresponding state of consciousness. In particular, it is experimentally extremely demanding, to determine the exact point of time (on a sub-second scale), at which a conscious experience occurs [5]. In any case, it seems to be impossible, to reach the same time resolution in the domain of consciousness as can be reached in the neural domain, because of the subjective fusion of experiences occurring within a certain time window [6].

This so-called fusion threshold was estimated, for instance, to be around 4 ms for auditory sensations and around 25 ms for visual sensations [7]. The issue of causal relationships is therefore still relevant in the context of the NCC program, under the premise that a small and practically unobservable time shift $\pm \Delta t$ may lie between neural states and corresponding states of consciousness.

Hume argued, that constant conjunction together with subsequence in time and contiguity in space and time, generates the idea of causality in the observer ("post hoc, ergo propter hoc"). More precisely, causality according to Hume is only an idea and has no implications beyond these features. According to Hume, there is no ontological difference between post hoc and propter hoc. Particularly in the natural sciences, this view has been questioned and it has been pointed out, that there is indeed a physical characteristic of causal interactions. This characteristic is the exchange of a conserved quantity, which in general is energy [8,9] (there are cases of causal interaction, where no exchange of energy, but of another conserved quantity takes place: for example, when two bowls of equal weight and velocity collide, an impulse exchange, but no energy exchange occurs).

Today, there are many different usages of the term causality [10,11]. The present analysis will focus on two frequently adopted meanings. Following the arguments introduced above, causality is understood as efficient causation addressing the physical relation of two events and the energy exchange criterion [12]. Secondly, causality is often used in the broad sense of an explanatory feature indicating the nomological character of the conjoined occurrence of two events [13]. With these two aspects in mind, what does the NCC program implicate in terms of causation? If a one-to-one correspondence between certain neural states and certain states of consciousness is observed, this correspondence may be integrated into a general psychophysiological law. With such a "bridge law" a certain state of consciousness may be deduced from a given neural state (the term "bridge law" was originally suggested as an empirical hypothesis, which establishes a connection between two scientific terms belonging to two different theories [14]). Thus, one may speak of a deductive-nomological explanation according to the covering-law model [15]. Under the assumption of an infinitesimally small time lag, $\pm \Delta t$, between neural states and corresponding states of consciousness, one may claim that a certain neural state $S \oplus N_S(t_n)$ causes a certain state of consciousness $C(t_n + \Delta t)$ in the sense of explanatory causality. However, it would also be justified to read the same bridge law backwards and postulate that a state of consciousness $C(t_n - \Delta t)$ may cause a certain neural state $S \oplus N_S(t_n)$. In other words, the direction of this kind of nomological relation is not obvious per se, unless there is a preceding commitment that consciousness is the explanandum.

With respect to the physics-based meaning of causality, efficient causality, it is valid to state that one neural state $S \oplus N_S(t_n)$ may cause a subsequent neural state $S \oplus N_S(t_{n+1})$. However, to claim, that a neural state $S \oplus N_S(t_n)$ causes a certain state of consciousness $C(t_n + \Delta t)$ would not be justified. The concept of causally efficient interactions between neural states and states of consciousness, in a way, implies a dualistic view of the mind-brain relation. Within this view, the problem arises that present scientific knowledge indicates that the physical domain is causally closed and that
the energy conservation law is strictly valid (with the yet only exception of space-time singularities (black holes)). It has to be remarked that interactions between the neural and phenomenal space were proposed to occur in the quantum-mechanical domain \[16,17\]. But, in spite of the long tradition of this argument \[18\], there is as yet no direct scientific evidence for such a causal interaction between conscious mind and quantum-physical objects. Moreover, since single quantum events are not significant for electric brain processes, this hypothesis is lacking an element connecting it to neural mass activity, and therefore to the most probable NCC candidates. It is not even conceivable, how an exchange of a quantity could occur between the physical domain and the domain of consciousness at all. Thus, it is not appropriate to state, that the concept of NCCs yields a causal explanation of consciousness, when causation is understood as it is in physical sciences (Wesley Salmon’s causal-statistical theory of explanation actually uses another, process based definition of causality \[19\], which, however, has been criticized to be insufficient from a physics-based point of view \[20\]). Of course, the problems associated with the dualistic approach are avoided by identity theory, which postulates that conscious and neural processes are essentially identical. Within this view, however, it would not be appropriate to state that certain neural states cause certain states of consciousness in the sense of efficient causality. One would rather claim that certain states of consciousness basically are neural states.

To summarize, the NCC program, in its present form, justifies the claim that “certain neural states cause certain states of consciousness” only in the broad sense of explanatory causality. But this kind of nomological relation in principle works both ways: states of consciousness may be caused by neural states and vice versa. In the more specific sense of efficient causality, the identification of NCCs can provide no causal explanation for the occurrence of certain states of consciousness.

Acknowledgements

This article is based on a chapter of a philosophical thesis advised by Prof. Thomas Metzinger at the University of Mainz (Germany). We are grateful to Thomas Metzinger, Laura Parkes and Bettina Walde for helpful comments and suggestions.

References