Why the Assessment of Causality in Brain–Behavior Relations Requires Brain Stimulation

Juha Silvanto[1] and Alvaro Pascual-Leone[2]

Abstract

■ A central aim in cognitive neuroscience is to explain how neural activity gives rise to perception and behavior; the causal link of paramount interest is thus from brain to behavior. Functional neuroimaging studies, however, tend to provide information in the opposite direction by informing us how manipulation of behavior may affect neural activity. Although this may provide valuable insights into neuronal properties, one cannot use such evidence to make inferences about the behavioral significance of the observed activations; if A causes B, it does not necessarily follow that B causes A. In contrast, brain stimulation techniques enable us to directly modulate brain activity as the source of behavior and thus establish causal links. ■

Imagine a heater, which controls the temperature of a house. This heater has an LED on–off indicator on its side. As the temperature of the room varies, so does the light on the LED indicator. Can we deduce from this covariance that the LED indicator controls the room temperature? The obvious answer to the question is "no"; it is clearly the heater that controls the room temperature, with the light on the LED indicator a mere by-product of this process. But how can we be sure of this? The only option is to intervene. For example, we could disable either the LED light or the heater and then measure the room temperature. Only the latter will have an impact on the room temperature, and thus we can confirm a causal link between the heater and the room temperature.

Now, replace the room temperature in this analogy (originally described by Sarter, Berntson, & Cacioppo, 1996) with behavior and the activity of the LED light and the heater with brain activity. The inherent problem in fMRI interpretation becomes apparent: Neural activity, which covaries with behavior, does not necessarily contribute to cause that behavior. An observed BOLD signal change may be an indirect consequence of the experimental manipulation and play no causal role in the cognitive function (see D'Esposito, in press; Sarter et al., 1996, for more detailed discussions). How can we determine which of the observed activations are causal to behavior and which are mere by-products?

The only option is to intervene. We need to manipulate the activity in these regions (akin to disabling the LED light) independently of the behavior. It might be possible to do this using brain imaging methods, for example, real-time fMRI feedback (e.g., deCharms et al., 2005): Patients appear capable of modifying BOLD activity in a given brain region using fMRI-based biofeedback, and then, we can inquire whether the behavior is secondarily modified. This is akin, in the room temperature analogy, to increasing the light in the LED indicator by means independent of the temperature and then checking to see whether the temperature in the room has gone up. However, presently, the most reliable way to intervene in such a way and thus establish causal brain–behavior relations is with brain stimulation techniques. In such a setting, fMRI data are of value because they provide experimenters with clues as to which regions to target, and the causality of these regions can then be tested, for example, with TMS. The correlational nature of fMRI does not render the utility of fMRI-guided TMS as inexplicable, quite the contrary: It provides a list of candidate regions that may be causally involved in the cognitive function under investigation. fMRI-guided TMS is thus a good example of how converging methods are indispensable in cognitive neuroscience. In addition, TMS enables the assessment of the causal relation in terms of the timing of the engagement of a brain area in a given behavior because of its excellent temporal resolution. We can thus explore the "chronometric causality" of brain activations identified by functional imaging.

A fundamental principle of cognitive neuroscience is that neural activity causes perception and behavior. It is precisely this direction of causality, from brain to behavior, that makes brain stimulation a unique methodology by allowing a manipulation of the source of behavior. Manipulating patients' perception and measuring brain activity at different levels of that manipulation to assess the influence of behavior on brain activity provides information in the

[1]Aalto University School of Science and Technology, Espoo, Finland, [2]Harvard Medical School, Boston, MA

opposite direction. This can certainly be of great interest, but it is not directly relevant for understanding how neural processes give rise to behavior. In other words, although behavioral changes may produce changes in brain activity, this is not the direction of causality that is generally meant when discussing causal links between brain and behavior. Furthermore, one cannot use such evidence to make inferences about the behavioral significance of the observed activations: If A causes B, it does not necessarily follow that B causes A.

A further aspect to consider is that, because brain activity causes behavior, modulation of patients' behaviors by different kinds of intervention (e.g., different visual stimuli) must be mediated by a change in brain activity. In other words, it may not be the percept or behavior that is changing brain activity. Instead, changes in sensory stimulation induce a change in brain activity that secondarily modifies perception. So, perhaps, such instances might be best conceptualized as "indirect noninvasive brain stimulation" (see Figure 1), one that psychophysicists and experimental psychologists have particularly longstanding expertise with—indeed, visual adaptation is often referred to as "psychologist's microelectrode." The problem, as outlined above, is to determine which neural effects of the change in stimulation play a causal role in the subsequent perceptual change. Even when visual stimulation remains constant and the percept changes, as is the case in many paradigms assessing visual awareness (Frith, Perry, & Lumer, 1999), one cannot infer which of the observed activations is causally related to the actual percept.

Of course, all techniques have potential limitations, and TMS is no different. As Weber and Thompson-Schill (2010) note, a TMS-induced behavioral impairment does not necessary imply that the computation of interest is processed in the region from which the impairment is induced. This is indeed the case in conventional "virtual lesion" TMS paradigms in which a behavioral impairment is expected if the stimulated region in any way contributes to the perceptual or cognitive function under investigation. For instance, TMS applied over the primary visual cortex (V1) disrupts face detection. V1 is unlikely to play a central role in face perception but does provide input to face-selective regions, and distributed transynaptic effects of TMS likely account for such effects. In TMS studies, it is thus imperative to tease apart the different contributions of various cortical regions. Once again, the combination of TMS and functional imaging can be quite powerful in such an endeavor. However, state-dependent TMS (see Silvanto & Pascual-Leone, 2008) provides a particularly powerful paradigm to further examine issues of causality. The strength of this approach is that, by systematically

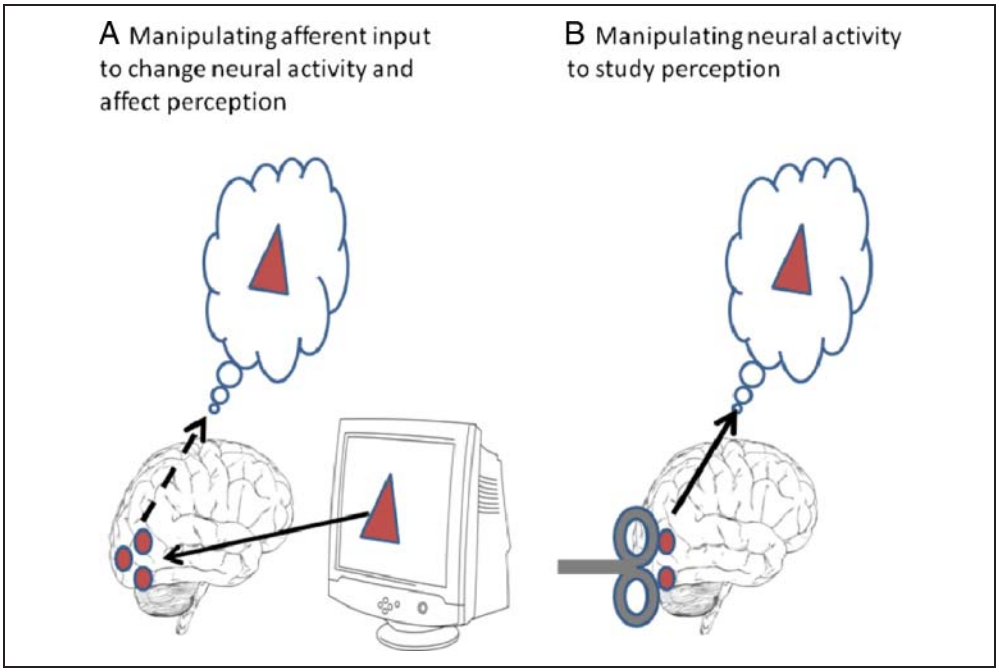

Figure 1. A central aim in cognitive neuroscience is to explain how neural activity gives rise to perception and behavior. Functional neuroimaging studies, however, tend to provide information in the opposite direction (A) by providing information on how changes in external factors such as sensory stimulation change brain activity (black line). The induced change in neuronal activity may secondarily modify perception (dashed black line)—one could thus conceptualize this approach as "indirect noninvasive brain stimulation." It is however difficult to determine which neural effects of sensory stimulation play a causal role in subsequent perception. In other words, one cannot deduce which locus of activation (depicted by the three red circles) gives rise to perception. In contrast, brain stimulation techniques enable us to directly modulate the source of perception and behavior—neural activity (B)—and thus make causal links between brain and behavior.

manipulating neural activation states before application of TMS, one can differentially target functionally distinct neuronal populations. If, for example, one were to adapt or prime subjects to faces before application of TMS, a state-dependent TMS effect would only be observed in regions that contain face-selective neurons, and therefore, V1 effects, as described above, would not be observed (Silvanto & Pascual-Leone, 2008). This state-dependent approach has so far been used in studies on numerical cognition (Cohen Kadosh, Muggleton, Silvanto, & Walsh, in press), action observation (Cattaneo, Sandrini, & Schwarzbach, in press), and visual imagery and STM (Cattaneo, Vecchi, Pascual-Leone, & Silvanto, 2009).

Brain stimulation techniques are important because they enable us to directly modulate the source of perception and behavior: neural activity. In contrast, functional neuroimaging measures neural activity associated with behavioral or perceptual changes (although real-time fMRI feedback may offer an alternative). Both approaches provide valuable information, but when the direction of causality that is of most interest is from brain to behavior, manipulation of brain activity, rather than behavior, is of greatest value. That is what brain stimulation can deliver, safely and reliably, if appropriate guidelines are followed (Rossi, Hallett, Rossini, Pascual-Leone, & The Safety of TMS Consensus Group, 2009).

Reprint requests should be sent to Alvaro Pascual-Leone, Berenson-Allen Center for Noninvasive Brain Stimulation, Beth Israel Deaconess Medical Center, Harvard Medical School, 25 Shattuck Street, Boston, MA 02115, or via e-mail: apleone@bidmc.harvard.edu.

REFERENCES

Cattaneo, L., Sandrini, M., & Schwarzbach, J. (in press). State-dependent TMS reveals a hierarchical representation of observed acts in the temporal, parietal, and premotor cortices. *Cerebral Cortex*.

Cattaneo, Z., Vecchi, T., Pascual-Leone, A., & Silvanto, J. (2009). Contrasting early visual cortical activation states causally involved in visual imagery and short-term memory. *European Journal of Neuroscience, 30,* 1393–1400.

Cohen Kadosh, R., Muggleton, N., Silvanto, J., & Walsh, V. (in press). Double dissociation of format-dependent and number-specific neurons in human parietal cortex. *Cerebral Cortex*.

deCharms, R. C., Maeda, F., Glover, G. H., Ludlow, D., Pauly, J. M., Soneji, D., et al. (2005). Control over brain activation and pain learned by using real-time functional MRI. *Proceedings of the National Academy of Sciences, U.S.A., 102,* 18626–18631.

D'Esposito, M. (in press). Why methods matter in the study of the biological basis of the mind: A behavioral neurologist's perspective. In P. A. Reuter-Lorenz, K. Baynes, G. R. Mangun, & E. A. Phelps (Eds.), *The cognitive neuroscience of mind: A tribute to Michael S. Gazzaniga*.

Frith, C., Perry, R., & Lumer, E. (1999). The neural correlates of conscious experience: An experimental framework. *Trends in Cognitive Sciences, 3,* 105–114.

Rossi, S., Hallett, M., Rossini, P. M., Pascual-Leone, A., & The Safety of TMS Consensus Group. (2009). Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. *Clinical Neurophysiology, 120,* 2008–2039.

Sarter, M., Berntson, G. G., & Cacioppo, J. T. (1996). Brain imaging and cognitive neuroscience. Toward strong inference in attributing structure to function. *American Psychologist, 51,* 13–21.

Silvanto, J., & Pascual-Leone, A. (2008). State-dependency of transcranial magnetic stimulation. *Brain Topography, 21,* 1–10.

Weber, M. J., & Thompson-Schill, S. L. (2010). Functional neuroimaging can support causal claims about brain function. *Journal of Cognitive Neuroscience, 22,* 2415–2416.