ACCSCIENCE PUBLISHING

Beytulhikme An International Journal of Philosophy

REVIEW ARTICLE

The causal theory of action and brain–computer interfaces

Kübra Gizem Ganioğlu*

Independent Researcher, Ankara, Türkiye

Abstract

Brain–computer interface (BCI) technology was initially confined to healthcare applications but has gradually expanded into fields such as entertainment, digital gaming, and the defence industry. This expansion has led to growing academic and interdisciplinary interest in the technology. The expanding range of applications of BCI technology makes it necessary to consider not only its current uses but also potential future forms of use and the associated risks, such as threats to cognitive privacy and cyberattacks. The fundamental significance of BCI technology in this study lies in the fact that it presents a form of action that differs from the traditional understanding of action. Users communicate with the outside world and act upon their environment via robotic limbs, prostheses, and other digital systems. In this process, actions are executed through computer systems directed by brain signals. Direct bodily execution through limb movement is therefore not necessary for action in the traditional sense. This study examines the functions and applications of BCI technology and discusses its relationship with the concept of action within the framework of four fundamental problems. Firstly, it addresses what the indirect realisation of intention means from the perspective of the causal theory of action. Secondly, it examines whether BCI-mediated actions should be classified as basic or non-basic actions. Thirdly, the study discusses how legal responsibility should be attributed when actions occur without direct involvement of the body and muscular system. Finally, it considers who bears responsibility for actions that arise in scenarios involving external intervention, such as neurohacking.

Keywords: Causal theory of action; Brain–computer interface technology; Responsibility; Neurohacker; Free will

***Corresponding author:**
Kübra Gizem Ganioğlu
(gizemganioglu@hotmail.com)

Citation: Ganioğlu, K. G. (2026). The causal theory of action and brain–computer interfaces. *Beytulhikme An International Journal of Philosophy*. https://doi.org/10.36922/BIJP89571

Received: January 29, 2026

Revised: September 18, 2026

Accepted: September 23, 2026

Published online: October 15, 2026

Copyright: © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons AttributionNon-Commercial 4.0 International (CC BY-NC 4.0), which permits all non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Publisher's Note: AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

Scientific and technological advances are increasingly fostering integration between humans and machines. Technologies incorporated into the body, such as pacemakers, cochlear implants, deep-brain stimulation devices, and bionic prostheses, have moved beyond merely assisting bodily functions and have become part of those functions themselves. Among these technologies, the brain–computer interface (BCI) holds a unique position because it enables direct integration between humans and machines. By establishing a direct connection between the brain and external devices, the BCI allows users to control robotic limbs, computers, and digital tools through mental processes. The resulting movement is produced neither by muscular contraction nor

by reflex. Rather, it is a mediated movement generated by transmitting neural activity to an external system through an interface.

This technical innovation raises a philosophical problem that is deeper than it may initially appear. The classical definition of the concept of action establishes a close link between intention and bodily execution. BCI, however, loosens this link precisely. The user's intention remains intact, while the bodily execution through which that intention would ordinarily be realised in the external world is bypassed. For this reason, BCI deserves to be regarded not merely as a new application of existing theories of action, but as a case that tests the limits of these theories.

According to the causal theory of action, one of the most influential approaches to defining action, a bodily movement counts as an action only if it is causally related to the agent's beliefs and desires (Davidson, 2001). For example, a goal scored as a result of a reflex is not an action in this sense. A conscious and intentional strike, on the other hand, is accepted as an action (Davidson, 2001). At first glance, BCI systems appear to satisfy this condition. When the user intends to drink water, the prosthetic arm moves towards the glass in accordance with this intention. However, the real difficulty lies not in the existence of the causal relationship, but in its nature. In classical motor control, intention directly produces bodily movement via the relevant neurophysiological mechanisms. However, in BCI systems, the same relationship takes on a mediated structure through the decoding of brain signals and their reproduction via an external technological interface. The starting point of this study is therefore whether the mediated structure of BCI action satisfies the requirement of the right kind of causation imposed by the causal theory of action.

The central thesis of this study is that BCI-mediated actions can neither be identified with traditional bodily actions nor reduced to technical outputs independent of the agent. What matters for attributing such actions to the agent is not simply that the interface detects the user's intention and converts it into action, but the degree of control the user retains over the action process. This thesis will be developed through four related problems.

The first problem concerns what the indirect realisation of an intention means for the causal theory of action. The second problem is whether actions realised via BCI should be regarded as basic actions or as non-basic actions. In the classical literature, basic actions are defined as movements performed directly, without being realised through another action. Walking, for example, is commonly treated as a basic action because it is not achieved by first performing a more basic action. In BCI systems, however, movement is produced through intermediate processes such as brain-signal generation, attention, and signal decoding, together with technological mediation. This raises the question of whether BCI-mediated actions retain the status of basic actions or require the concept itself to be reconsidered.

The third problem concerns whether an action should be regarded as mental or bodily and consequently, how legal responsibility should be established. BCI technologies are often described as translating thought into action, and they complicate this distinction. Classical approaches tend to associate action primarily with bodily behaviour and to ground responsibility in bodily performance. The fourth problem concerns neurohacking scenarios in which third parties interfere with the system: to whom should the resulting action be attributed, and to what extent can cognitive privacy be protected?

The study proceeds in five steps. Section 2 introduces the functioning, areas of application, and main functions of BCI technology. Section 3 examines the tension between the central claims of the causal theory of action and the functioning of BCI systems. Section 4 discusses the distinction between basic and non-basic actions and the debate surrounding the sense of agency. Section 5 addresses the question of whether an action is mental or bodily in the context of responsibility. Section 6 then examines the distribution of responsibility when neurohacking interventions occur.

2. Brain–computer interface technologies

To discuss the philosophical status of BCI-mediated action, it is first necessary to clarify what this technology is and what functions it performs. Technology is also one of the most important themes in the science fiction genre. While scientific and technological developments transform the structure of societies and human life, science fiction explores these transformations through possible future scenarios (Doğan, 2023). For example, one of the most popular works relating to BCI technology and the science fiction genre is the *Star Trek* series. In the *Star Trek* series, which was first broadcast in 1966, Captain Pike is left completely paralysed after an accident and communicates with those around him through a wheelchair controlled by brain signals. This scenario is noteworthy because it anticipates the experience of individuals who retain their cognitive abilities while losing muscular control and offers an early imaginative representation of possible BCI applications. The later realisation of similar scenarios appears to have been enabled by advances in the technologies required to implement such ideas and by increased research on the brain (Graimann *et al.*, 2009).

The idea that signals recorded from the brain could be used for communication and control has long been a subject of scientific interest. In 1924, Hans Berger discovered that the electrical signals produced by the brain could be recorded without the need for invasive methods. His suggestion that people might act through brain signals without relying on ordinary bodily movement captured the imagination of the scientific community (Wolpaw & Wolpaw, 2012). BCI technology thus differs from the conventional operation of the central nervous system by enabling individuals to communicate with the external world and meet their needs without relying on bodily or muscular movement.

Brain–computer interface systems are capable of controlling devices such as computers, robotic arms, and wheelchairs by detecting brain and nerve activity. Actions performed without any physical or muscular movement can realise an individual's intentions in a manner similar to the natural actions they would perform through their own body and muscular activity (Kögel *et al.*, 2019). BCI systems now attract growing interest in both academic and applied research and have become an important field in which substantial advances are regularly reported.

2.1. Applications and recent developments in brain–computer interfaces

Although BCI technology is currently used most often in healthcare, it also has applications in remote communication, education, manufacturing and marketing, security, gaming, and entertainment, and its use is expected to extend into further domains (Abdulkader *et al.*, 2015). By using the neurofeedback functions of BCI technology, a person's level of concentration and motor behaviour can be analysed during activities that require sustained attention, such as driving. In this way, potential lapses in attention and delays in reaction may therefore be detected before they lead to an accident. BCI technology may also identify fatigue, drowsiness, and workload-related cognitive fatigue at an early stage, allowing preventive measures to be taken (Mudgal *et al.*, 2020).

While there are various examples in the literature on this subject, one of the most widely discussed applications is the Neuralink initiative. Neuralink has attracted attention for its aim of developing scalable, high-bandwidth BCI systems (Musk & Neuralink, 2019). Neuralink's official website states that BCI technology is being developed with two primary objectives in mind. The first is to increase the capacity of people with neurological conditions to interact with the external world and act independently. The second is to enable applications intended to enhance human cognitive and physical capacities. This latter aim positions BCI technology not only as a therapeutic tool but also as a means of human enhancement and, in this respect, connects it with transhumanist approaches (Neuralink, 2024).

Companies working on BCI technology generally aim to enable individuals with conditions such as locked-in syndrome to communicate with the external world. Locked-in syndrome is a serious neurological condition in which the individual has largely lost their ability to move and is unable to communicate verbally. Although it is a rare condition, the promising results achieved by BCI technologies in enhancing the communication capacity of individuals with locked-in syndrome have significantly increased interest in this field (Fenton & Alpert, 2008). The very name of the syndrome reflects how the condition is experienced by both the individuals themselves and those who assist in their care. People with the syndrome remain conscious but face severe difficulties in expressing themselves to the outside world. Some authors therefore treat the condition as a case that calls for renewed consideration of the relationship between self and body (Fenton & Alpert, 2008). On this view, locked-in syndrome can create the impression that a person is trapped within their own body, reinforcing the idea of a distance between self and body. As Fenton and Alpert put it: "The body is presented as a vessel in which the self is housed and through which the self must express itself, if it can at all." (Fenton & Alpert, 2008, p. 120). BCI technology may therefore affect not only a patient's ability to communicate with the external world and act upon their environment, but also the formation and experience of selfhood.

2.2. Functions of brain–computer interface systems

According to Jonathan R. Wolpaw and Elizabeth Winter Wolpaw (2012) BCI systems have five fundamental functions. Firstly, the replacement function allows a BCI to substitute for a natural output that an individual can no longer produce. A person who has lost the ability to speak, for example, may use a BCI to generate written output and, with the aid of artificial intelligence-supported systems, render that output as speech. Secondly, the restoration function aims to restore a lost natural bodily function by using the BCI to stimulate the affected muscles or nerves and thereby enable the person to perform the function again. Thirdly, the enhancement function supports better or more accurate performance of an existing output. In tasks that require sustained attention, such as driving, a BCI may detect distraction and issue a warning, thereby helping the user maintain attention. Fourthly, the supplementation function provides additional artificial capabilities that extend a person's natural abilities without replacing or restoring an existing bodily function. Fifthly,

the improvement function aims to improve a person's natural bodily or neural function through repeated BCI use, potentially promoting neural plasticity and more normal functional performance (Wolpaw & Wolpaw, 2012).

The enhancement, supplementation, and improvement functions can be connected to the goal of augmenting human cognitive and physical capacities, which is among the long-term aims of initiatives such as Neuralink. In this respect, BCI technologies parallel transhumanist approaches that seek to extend existing biological limits through neurotechnological interfaces. Transhumanism is associated with aims such as extending human lifespan, increasing cognitive capacity, and eliminating disease (Dağ, 2018). In this context, it can be understood as a philosophical framework that advocates the enhancement of human capacities through genetic engineering, nanotechnology, artificial intelligence, and particularly BCI systems.

3. Brain–computer interfaces in the context of the causal theory of action

According to the causal theory of action, an individual's bodily actions must be causally grounded in beliefs and desires (Davidson, 2001). The central idea of the theory is that an agent acts only when an appropriate internal state causes a particular outcome in the appropriate way (Davis, 2010). In other words, for a behaviour to be considered an action, it must be appropriately caused by certain types of mental states (Clarke, 2010).

Suppose, for example, that a woman wants to drink coffee. She raises her arm, reaches for the cup, and drinks. Considered separately, these movements may appear independent of one another, but when organised around the aim of drinking coffee they form a coherent sequence of action. The sequence can therefore be understood as meaningful behaviour causally guided by the person's beliefs and desires. A robot may perform similar functions, but if it lacks mental states, its behaviour is treated as programmed output rather than intentional action.

However, it is argued that there is a structural incompatibility between the causal theory of action and the way BCI systems operate. To illustrate this incompatibility, consider the following example. Suppose a neural monitoring device has been implanted in Ichiro's brain to detect whether he has formed an intention to act. Unknown to Ichiro, his motor areas and efferent pathways do not produce bodily movement despite his intention. Instead, when a neuroscientist detects the intention, the neuroscientist externally stimulates Ichiro's motor system and thereby produces the bodily movement. Although the movement is causally connected to Ichiro's mental states, the immediate mechanism that produces it is not Ichiro's own motor-control system but an intervention by an external agent. The behaviour therefore results from a causal process mediated by external intervention rather than from an ordinary direct response to Ichiro's intention (Aguilar & Buckareff, 2010; Buller, 2021b). "To put the point differently, Ichiro's case can be seen as an instance of deviant causation since it has the wrong type of causal process." (Buller, 2021b, p. 157). The example shows that the mere existence of a causal connection between intention and bodily movement does not by itself settle whether the movement is attributable to the agent.

According to Tom Buller (2021b), this mediation may not, on its own, eliminate agency. In his view, what is decisive is whether the causal process is of the right kind, as the causal theory of action requires. In this context, the system's sensitivity to the user's intentions and its reliability can be seen as fundamental criteria in assessing whether the resulting behaviour constitutes an action (Buller, 2021b). Consequently, the issue is whether intentional states are actually conveyed to the external world in a way that constitutes an action.

On the other hand, another factor that disrupts the chain of causality in BCI-mediated actions is the technical quality of the recorded brain signals. Today, BCI technology is capable of recording the brain's real-time signals using both invasive and non-invasive methods (Coin & Dubljević, 2023). However, the inability to obtain simultaneous, high-resolution data from a large number of neurons remains a significant limitation of this technology. Non-invasive methods can record aggregate activity from millions of neurons through the skull, but the resulting data are generally low-resolution and vulnerable to signal distortion. Invasive electrodes placed on the cortical surface provide clearer signals but remain limited in their ability to access deeper brain regions (Musk & Neuralink, 2019). Because these signals function as causal intermediaries that carry information about the user's desires and intentions, noise or technical deficiencies may cause the interface to misinterpret an intention and produce an unintended action. This raises the question of whether such an action should be attributed to the interface, the user, or the algorithm. The next section develops this issue through the distinction between basic and non-basic actions.

4. The debate on basic and non-basic actions

According to Arthur C. Danto, saying that person M causes stone S to move may appear straightforward, but it does not by itself clarify the nature of the causal relation between the

action and its outcome (Danto, 1965). From an external perspective, we cannot tell whether the stone moved as the direct result of something M did or as the indirect outcome of a sequence of actions. To clarify this distinction, Danto argues that actions occur at two levels. Some actions are brought about indirectly by performing another action, whereas basic actions are performed directly and do not require the performance of a further action (Danto, 1965). An action attributable to M is therefore either a basic action performed directly by M or the result of a causal chain that begins with such a basic action (Danto, 1965).

In everyday life, many actions based on bodily movements are generally regarded as basic actions. In this context, movements such as moving a finger, raising a hand, turning the head to the right or left, and moving a foot forward are among the basic actions that the agent performs directly, without mediating another action. In contrast, non-basic actions result from the performance of other actions. In this case, the individual achieves the intended action indirectly (Buller, 2021a). For example, situations such as text appearing on the screen as a result of pressing keys on the keyboard, or a light coming on as a result of pressing a button, are the product of a broad chain of causal and technical processes. Such cases therefore count as non-basic actions.

From the perspective of action theory, the external movements produced through a BCI are generally classified as non-basic actions. The central philosophical issue is the multiple layers of mediation involved in their production. A person with ordinary motor control can raise an arm without first performing another action. A BCI user, by contrast, must generate the relevant mental command, after which the intended movement is produced through stages such as signal decoding, device commands, and motor output. The mediated and indirect structure of BCI systems has therefore been argued to distinguish these actions from ordinary bodily actions and to generate a distinctive set of legal problems (Bublitz *et al*., 2019). At the same time, the distinction between basic and non-basic action may become less clear as users grow accustomed to BCI systems and begin to perform BCI-mediated actions as naturally as ordinary bodily actions. If users in active mode internalise control over the action and acquire the relevant skill over time, BCI-mediated actions may come to approximate the status of basic actions (Drosselmeier & Sellmaier, 2021).

The skills acquired through continued use of a BCI may lead users to internalise control over the system and reduce the felt significance of external mediation. This can strengthen the basis for attributing the resulting action to the user. The sense of agency is the experience of being the source of one's own actions, the subjective sense expressed in the judgement "I did that." When a person initiates an action, the action plan, bodily sensory feedback, and the outcome in the external world are integrated. When these elements are sufficiently congruent, the person experiences the action as their own.

Experiments examining changes in the sense of agency in BCI contexts show that participants can experience a sense of disembodied agency even in the absence of bodily feedback (Caspar *et al*., 2021). These findings suggest that the key determinant in the formation of the sense of agency is not physical feedback, but rather the individual's belief that they are in control of the action. When a BCI user feels that they are personally operating a robotic hand and that this action is taking place in accordance with their own will, they attribute the action to themselves and can arrive at the judgement "I performed this action" (Caspar *et al*., 2021).

The ability to use a BCI is widely recognised as a competence acquired over time. The key factors enabling the learning process are the brain's neuroplastic structure and the neural adaptations that result from training. Through such adaptation, users may eventually experience the BCI system as a natural extension of their own action (Drosselmeier & Sellmaier, 2021). The process can be compared with learning to ride a bicycle. At first, maintaining balance, stopping, and coordinating movement require deliberate effort, but with practice these abilities improve, and cycling comes to feel natural (Wander *et al*., 2013). Similarly, as users adapt to a BCI, control processes that are initially deliberate and cognitively demanding can become more efficient, automatic, and natural over time (Wander *et al*., 2013).

Despite the substantial experimental literature on the sense of agency, BCI systems must address problems of feedback if they are to help resolve questions about the attribution of action (Steinert *et al*., 2019). BCI systems typically involve a temporal delay between the brain signal and the resulting output. One study, for example, found that increasing the delay between a mental command and the movement of an on-screen cursor reduced participants' sense of ownership over the action. When this neuro-visual delay exceeded approximately one second, the perceived degree of control weakened substantially (Evans *et al*., 2015). Some scholars describe this lack of control in BCI systems as follows:

> It was noted that the possibility of full control of BCI devices will be limited because of gaps in knowledge concerning the neural story of intentional action (versus mere imaginings, or mental planning, for example). This basic design

> constraint on the possibility of control has knock-on effects for the use of specific devices. Control will only ever be partial. This stands as an umbrella concern for neuro-controlled technologies. There will always be a cap on the amount of control available to a user, even if they are maximally proficient in its use (Rainey *et al.*, 2020, p. 56).

If the motor movements of an individual using the BCI system do not meet these conditions, one might argue that the BCI system, rather than the individual, is the more decisive source of the resulting behaviour. It then becomes unclear whether such actions should be classified as non-basic actions or placed in a separate category. It is also uncertain to whom their moral consequences should be attributed and how their capacity to cause benefit or harm should be understood. The debate over basic action therefore leads to a more fundamental question: is an action performed through a BCI a mental action or a bodily action? The next section considers this question in the context of responsibility.

5. Responsibility in brain–computer interface systems

Brain–computer interface users are generally understood to initiate actions through mental processes, yet the execution of those actions also depends on a physical device. Whether BCI-mediated actions should therefore be classified as mental or bodily remains contested. Because these actions differ from traditional actions mediated through the muscular and nervous systems, they raise distinctive ethical and legal questions. Consider a BCI user who wants to drink water. The user imagines moving a prosthetic left arm with the intention and desire to drink, and the prosthesis moves towards the glass in response to this mental orientation. The key question is whether the action realised when the prosthetic arm grasps the glass should be understood as a bodily action or a mental action. This matters because theories of action are closely connected to assessments of moral and legal responsibility.

Some scholars, for example, distinguish in legal contexts between bodily actions and mental acts such as imagining, calculating, and thinking, and restrict legal assessment to bodily conduct (Bublitz *et al.*, 2019). As the brain signals controlling BCI systems are based on mental actions rather than muscle movements, BCI-mediated behaviour does not fit neatly within the law's traditional conception of bodily action. How such behaviour should be treated under existing legal concepts of action is therefore open to debate (Bublitz *et al.*, 2019).

Let us devise a thought experiment in the context of BCI technology. At a meeting, one colleague keeps talking non-stop. Another employee becomes angry and throws the glass of water in front of them at the colleague's face. The employee has never previously behaved disrespectfully towards colleagues. Although they might later say, "It was as if my arm was not under my control," they are held responsible because the action is taken to have been conscious and under their control (Rathkopf, 2024). Now modify the case by supposing that the employee's arm has been replaced by a robotic arm controlled through a BCI. The employee merely imagines throwing the water. This mental representation generates a particular pattern of neural activity in the motor cortex. For the employee to be able to say that the arm was not under their control, the system would have to misinterpret that neural pattern (Rathkopf, 2024). If no such malfunction occurs and the movement accurately realises the user's intention, the question becomes how responsibility for the resulting action should be attributed.

A further question is whether every mental image or fantasy should be translated into action. Suppose that, in a moment of anger, the employee merely imagines throwing the glass but does not actually intend to do so (Rathkopf, 2024). If BCI technology interprets this mental scenario as a concrete command to act and causes the glass to be thrown, the question of whether the employee bears responsibility for the action creates a philosophical conundrum. Can a person be held responsible for uncontrolled thoughts that cross their mind against their will, or for thoughts that they would not normally act upon?

There are two possible responses to this thought experiment: that BCI systems create a knowledge gap in terms of attributing responsibility (Bublitz *et al.*, 2019) and that the real issue lies in the decoding of intentions (Rathkopf, 2024). Firstly, from the perspective of third parties, it may be impossible to determine with certainty whether an action performed through a BCI genuinely expresses the user's will or instead results from system error or external intervention. Third parties have no direct access to the user's mental states, neural activity, or the BCI system's data-processing procedures (Bublitz *et al.*, 2019). Both the user's mind–brain system and the BCI may therefore function as "black boxes" from an external perspective (Bublitz *et al.*, 2019). It is also debatable whether BCI systems accurately interpret the user's intentions (Rathkopf, 2024). The claim that the system must be able to correctly identify intentions and desires demonstrates that the system's effectiveness depends on processes of adaptation and learning. Indeed, according to some authors, the success of BCI systems relies not so much on the direct reading of mental contents as on the

mutual adaptation that develops over time between the user and the system (Rathkopf, 2024). Accordingly, when a BCI user harms another person, responsibility should be assessed by considering whether the person acted intentionally or negligently, together with the complexity and predictability of the action and the extent to which the system was controllable, much as one would evaluate a person driving a car or performing robotic surgery (Rathkopf, 2024). This criterion also supports the thesis stated in the introduction.

6. Neurohacking interventions and the distribution of responsibility

The problem of the distribution of responsibility becomes most acute once the possibility of external interference with a BCI system is taken into account. To appreciate the significance of this possibility, it is useful first to consider the current scope of the technology. BCI systems may be regarded as one of the technologies underlying visions of a future "Humanity 2.0." In transhumanist accounts of human enhancement, the interaction between humans and machines is expected to become increasingly intensive, and BCI systems provide one of its most concrete forms. Humans integrated with machines to an unprecedented degree may eventually extend existing biological limitations through possibilities ranging from brain-to-brain interaction resembling telepathic communication to the optimisation of memory.

Some of the most prominent developments in BCI technology are in the defence sector. The Defense Advanced Research Projects Agency (DARPA), established in 1958, is described as seeking to maintain United States military technological superiority and to prevent technological surprises from rival actors (Miranda *et al*., 2015). DARPA has developed various programmes and conducted research within the scope of BCI technologies to enhance the physical and cognitive capabilities of soldiers. For example, some BCI systems developed for military training aim to significantly increase the rate at which new skills are acquired. These systems, which classify the user's learning level as novice, intermediate, or expert based on their brain activity, adapt the training process according to individual performance. In training programmes utilising this technology, it has been reported that new recruits acquire rifle marksmanship skills two to three times faster than with traditional training methods (Miranda *et al*., 2015).

Regardless of whether BCI technology is approached with technological optimism or scepticism, its potential ethical and social consequences cannot be ignored. Among the most important are privacy and cognitive autonomy, issues that are examined in the philosophy and ethics of computing. Within information ethics, privacy concerns the prevention of the unauthorised collection, use, processing, or sharing of personal data and information (Kantar, 2024). A major concern surrounding BCI technology is whether implants and related devices are secure against privacy breaches and unauthorised third-party access. Modern medical devices are increasingly connected to external digital systems. Implantable devices such as cardiac implants, insulin pumps, and deep-brain neurostimulation systems, for example, can be monitored and, to some extent, controlled through smartphones or computer-based applications using wireless technologies such as the internet and Bluetooth (Hassija *et al*., 2021). Medical implants provide clinical support for diagnostic, treatment, and surgical decision-making processes by enabling the monitoring of patients' physiological and clinical data. Yet unauthorised access to these data can produce serious privacy and security risks, particularly because some implantable devices contain location information in addition to health data (Hassija *et al*., 2021).

Recent findings regarding BCI systems reveal that these technologies harbour potential security vulnerabilities to cyberattacks. This possibility brings neurocrime, also described as neurohacking or brain hacking, to the fore. The concept of neurocrime extends cybercrime to neural interfaces and neurotechnological devices (Ienca & Haselager, 2016). Brain implants developed for the purposes of memory enhancement and neurological rehabilitation may threaten an individual's cognitive integrity and privacy as a result of cyberattacks. Such interference could enable the manipulation of memories, disclosure of personal information or, if a device were disabled, memory loss. BCI systems may therefore fail to satisfy the principles of confidentiality, integrity, and availability, creating significant security risks (Denning *et al*., 2009). Former United States Vice President Dick Cheney, "a longtime cardiac patient with an implanted cardioverter defibrillator, revealed in an interview that he had his device's Bluetooth capabilities disabled in 2007 over concerns about potential assassination attempts via hacking." (Browning & Tuma, 2016, pp. 644–645). In another example, at a 2011 conference, McAfee security expert Barnaby Jack demonstrated security vulnerabilities in insulin pumps: "Jack demonstrated how the same tiny radio transmitter that permits doctors to adjust the pump's functions could be turned against it, allowing the cyber attacker to disable the alert feature and dispense a potentially lethal dose of insulin." (Browning & Tuma, 2016, p. 647).

The significance of the misuse or abuse of BCI systems stems from the fact that these systems are directly or indirectly linked to the organs of the human body. The brain is vital not only for sustaining basic processes such as nutrition and respiration but also for cognitive capacities including perception, thought, judgement, memory, and language. Using BCI technologies as instruments of cybercrime could therefore have effects extending from an individual's physical safety to their behavioural processes (Ienca & Haselager, 2016).

It would be appropriate to revisit the water-throwing thought experiment devised as part of this study in this context. Let us suppose that a BCI user throws the glass in front of them, despite having no intention of doing so, because a hacker has infiltrated their system. From an external perspective, the action may appear to follow an ordinary chain of intention, desire, and belief, yet it is in fact disconnected from the agent's free will. Free will may be understood as the capacity to make choices and decisions independently. An outcome whose goals have been predetermined by algorithms and software cannot, on this account, be characterised as an action grounded in free will (Doğan, 2021). The BCI user may be aware of throwing the glass and may even perform the movement voluntarily, while the relevant intention itself has arisen through external intervention rather than the user's own free will. At this point, questions such as "Who bears responsibility?" or "Can responsibility for indirectly carried out actions be determined?" take on vital importance.

Perhaps the most critical issue in debates on the causal theory of action is whether, at the end of the causal chain linking the agent's mental states to the resulting physical outcome, an agent can be identified who bears responsibility for the action. However, when BCI systems are involved, the indirect nature of the physical action leads to the inclusion of a technical interface as one of the links in the causal chain, and this situation raises the problem of to what or to whom responsibility should be attributed. Although the interface link is conceptualised as a physical component integrated with the individual, the possibility that it may produce outcomes independent of the user's supervision and control in certain critical situations, such as hacking, must not be overlooked. This situation, given the "black box" nature of the system in BCI-mediated actions, renders the establishment of a legal framework and the application of the principle of individual criminal responsibility uncertain. The widespread adoption of BCI technologies across a broad spectrum, from the entertainment sector to the defence industry, transforms the debate on responsibility from an academic curiosity into a critical security issue. In military contexts, manipulation of a soldier's BCI system could lead to irreversible large-scale harm or unlawful conduct in war. The future development of BCI systems therefore appears to require greater transparency, algorithms capable of distinguishing the user's intentions from external intervention, and robust cyber-neurosecurity protocols designed to protect interfaces against third-party interference.

7. Conclusion

Brain–computer interface technologies have not only brought about a medical revolution by enabling direct integration between the human brain and external devices, but have also reopened the debate on the problem of action. In the philosophical literature, the causal theory of action explains action in terms of mental states such as beliefs and intentions producing behaviour through the right kind of causal relation. BCI technologies require the scope of this classical explanatory model to be reconsidered. In BCI-mediated action, the causal relation between mental states and action or output is not realised through direct bodily execution. It is realised through a multi-stage system involving intermediate processes such as signal decoding, classification, and device control. The central issue is therefore whether, and to what extent, mental states stand in the right kind of causal relation to the resulting behaviour.

The debate on basic and non-basic actions addressed in this study raises the question of whether actions performed via BCI can be considered to have the same status as traditional bodily actions. On the one hand, the fact that users acquire skill in using the system over time and develop a strong sense of agency supports the attribution of these actions to the user. On the other hand, the mediated nature of BCI systems, issues with feedback, and technical limitations on control make it difficult to accept these actions as basic actions in the classical sense. BCI-mediated actions therefore cannot simply be equated with traditional bodily actions, yet neither can they be fully accommodated within existing theories of action without further qualification.

The issue of the ownership and control of BCI-mediated actions raises the question of to whom responsibility should be attributed and demonstrates the extent to which these technologies challenge the traditional understanding of action in assessments of moral responsibility. As the user's intention, the system's interpretation processes and the functioning of the technological infrastructure are intertwined in BCI-mediated actions, it is not always clear to what extent the resulting outcomes should be attributed

to the user. In particular, system errors, misclassifications, or incompatibilities between the user and the system make it problematic to attribute responsibility to the user alone. However, this difficulty does not mean that the assessment of responsibility is entirely eliminated. On the contrary, the thesis put forward at the beginning of this study finds its justification here. The decisive factor in BCI actions is the user's level of control over the action. As BCI technologies become more widespread, new theoretical frameworks will therefore be needed to explain how moral and legal responsibility should be distributed among users, designers, and manufacturers.

The final section addressed neurohacking, an issue that is receiving increasing attention in the literature. The water-throwing thought experiment was expanded in this context to open up a discussion on the responsibility issues that may arise if a BCI user's action is altered by external intervention. Legal systems generally treat individuals as responsible for actions attributable to them, and this principle presupposes an agent to whom the action can properly be imputed. Because BCI systems generate action through a brain-interface relation, however, the interface may in principle be open to third-party intervention. In such cases, the question of whose action it is becomes substantially more complex, and legal principles such as individual criminal responsibility and mental privacy become harder to apply. By raising the possibility that an action may be produced independently of the agent's will, neurohacking challenges the limits of classical approaches to responsibility and calls for a reassessment of legal responsibility regimes in the context of BCI technology.

Acknowledgments

None.

Funding

None.

Conflict of interest

The author declares no competing interests

Author contributions

This is a single-authored article.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data

Not applicable.

References

Abdulkader, S. N., Atia, A., & Mostafa, M.-S. M. (2015). Brain computer interfacing: Applications and challenges. *Egyptian Informatics Journal, 16*(2), 213–230.

https://doi.org/10.1016/j.eij.2015.06.002

Aguilar, J. H., & Buckareff, A. A. (2010). The causal theory of action: Origins and issues. In *Causing human actions* (pp. 1–26). Cambridge, MA: The MIT Press.

https://doi.org/10.7551/mitpress/8614.003.0002

Browning, J. G., & Tuma, S. (2016). If your heart skips a beat, it may have been hacked: Cybersecurity concerns with implanted medical devices. *South Carolina Law Review, 67*(3), 638–677.

Bublitz, C., Wolkenstein, A., Jox, R. J., & Friedrich, O. (2019). Legal liabilities of BCI-users: Responsibility gaps at the intersection of mind and machine? *International Journal of Law and Psychiatry, 65*, 101399.

https://doi.org/10.1016/j.ijlp.2018.10.002

Buller, T. (2021a). Actions, agents, and interfaces. In *Clinical neurotechnology meets artificial intelligence* (pp. 11–23). Cham, Switzerland: Springer International Publishing.

https://doi.org/10.1007/978-3-030-64590-8_2

Buller, T. (2021b). Brain-computer interfaces and the translation of thought into action. *Neuroethics, 14*(2), 155–165.

https://doi.org/10.1007/s12152-020-09433-9

Caspar, E. A., De Beir, A., Lauwers, G., Cleeremans, A., & Vanderborght, B. (2021). How using brain-machine interfaces influences the human sense of agency. *PLoS ONE, 16*(1), e0245191.

https://doi.org/10.1371/journal.pone.0245191

Clarke, R. (2010). Skilled activity and the causal theory of action. *Philosophy and Phenomenological Research, 80*(3), 523–550.

https://doi.org/10.1111/j.1933-1592.2010.00339.x

Coin, A., & Dubljević, V. (2023). An introduction to policy, identity, and neurotechnology: The neuroethics of brain–computer interfaces. In *Policy, identity, and neurotechnology* (pp. 1–7). Cham, Switzerland: Springer International Publishing.

https://doi.org/10.1007/978-3-031-26801-4_1

Dağ, A. (2018). *Transhümanizm: İnsanın ve dünyanın dönüşümü* [Transhumanism: The transformation of humanity and the world] (M. Vural, Ed.). Ankara, Türkiye: Elis Yayınları. [In Turkish]

Danto, A. C. (1965). Basic actions. *American Philosophical Quarterly, 2*(2), 141–148.

Davidson, D. (2001). *Essays on actions and events* (2nd ed.). Oxford, United Kingdom: Clarendon Press; Oxford University Press.

Davis, W. A. (2010). The causal theory of action. In *A companion to the philosophy of action* (1st ed., pp. 32–39). Hoboken, NJ: Wiley.

https://doi.org/10.1002/9781444323528.ch5

Denning, T., Matsuoka, Y., & Kohno, T. (2009). Neurosecurity: Security and privacy for neural devices. *Neurosurgical Focus, 27*(1), E7.

https://doi.org/10.3171/2009.4.FOCUS0985

Doğan, M. (2021). Yapay zekâ ve özgür irade: Yapay özgür iradenin imkânı [Artificial intelligence and free will: The possibility of artificial free will]. *TRT Akademi, 6*(13), 788–811. [In Turkish]

https://doi.org/10.37679/trta.969912

Doğan, M. (2023). Sinemada yapay zekâ: Robotlarda bilinç, duygular ve etik [Artificial intelligence in cinema: Consciousness, emotions, and ethics in robots]. *Kültür ve İletişim, 26*(2)(52), 318–343. [In Turkish]

https://doi.org/10.18691/kulturveiletisim.1316168

Drosselmeier, S., & Sellmaier, S. (2021). Skilled action and the ethics of brain-computer interfaces. In *Clinical neurotechnology meets artificial intelligence* (pp. 25–37). Cham, Switzerland: Springer International Publishing.

https://doi.org/10.1007/978-3-030-64590-8_3

Evans, N., Gale, S., Schurger, A., & Blanke, O. (2015). Visual feedback dominates the sense of agency for brain-machine actions. *PLoS ONE, 10*(6), e0130019.

https://doi.org/10.1371/journal.pone.0130019

Fenton, A., & Alpert, S. (2008). Extending our view on using BCIs for locked-in syndrome. *Neuroethics, 1*(2), 119–132.

https://doi.org/10.1007/s12152-008-9014-8

Graimann, B., Allison, B., & Pfurtscheller, G. (2009). Brain–computer interfaces: A gentle introduction. In *Brain-computer interfaces* (pp. 1–27). Berlin, Germany: Springer Berlin Heidelberg.

https://doi.org/10.1007/978-3-642-02091-9_1

Hassija, V., Chamola, V., Bajpai, B. C., Naren, & Zeadally, S. (2021). Security issues in implantable medical devices: Fact or fiction? *Sustainable Cities and Society, 66*, 102552.

https://doi.org/10.1016/j.scs.2020.102552

Ienca, M., & Haselager, P. (2016). Hacking the brain: Brain–computer interfacing technology and the ethics of neurosecurity. *Ethics and Information Technology, 18*(2), 117–129.

https://doi.org/10.1007/s10676-016-9398-9

Kantar, N. (2024). Bilişim felsefesi bağlamında bilişim etiği ve güncel tartışmalar [Information ethics and contemporary debates in the context of the philosophy of informatics]. *Iğdır Üniversitesi Sosyal Bilimler Dergisi*, (36), 112–133. [In Turkish]

https://doi.org/10.54600/igdirsosbilder.1438547

Kögel, J., Schmid, J. R., Jox, R. J., & Friedrich, O. (2019). Using brain-computer interfaces: A scoping review of studies employing social research methods. *BMC Medical Ethics, 20*(1), 18.

https://doi.org/10.1186/s12910-019-0354-1

Miranda, R. A., Casebeer, W. D., Hein, A. M., Judy, J. W., Krotkov, E. P., Laabs, T. L., Manzo, J. E., Pankratz, K. G., Pratt, G. A., Sanchez, J. C., Weber, D. J., Wheeler, T. L., & Ling, G. S. F. (2015). DARPA-funded efforts in the development of novel brain–computer interface technologies. *Journal of Neuroscience Methods, 244*, 52–67.

https://doi.org/10.1016/j.jneumeth.2014.07.019

Mudgal, S. K., Sharma, S. K., Chaturvedi, J., & Sharma, A. (2020). Brain computer interface advancement in neurosciences: Applications and issues. *Interdisciplinary Neurosurgery, 20*, 100694.

https://doi.org/10.1016/j.inat.2020.100694

Musk, E., & Neuralink. (2019). An integrated brain-machine interface platform with thousands of channels. *Journal of Medical Internet Research, 21*(10), e16194.

https://doi.org/10.2196/16194

Neuralink. (2024). *Neuralink — Pioneering brain computer interfaces*. Accessed December 10, 2025. https://neuralink.com/

Rainey, S., Maslen, H., & Savulescu, J. (2020). When thinking is doing: Responsibility for BCI-mediated action. *AJOB Neuroscience, 11*(1), 46–58.

https://doi.org/10.1080/21507740.2019.1704918

Rathkopf, C. (2024). Culpability, control, and brain-computer interfaces. In *Neuro-ProsthEthics* (Vol. 9, pp. 89–102). Berlin, Germany: Springer Berlin Heidelberg.

https://doi.org/10.1007/978-3-662-68362-0_6

Steinert, S., Bublitz, C., Jox, R., & Friedrich, O. (2019). Doing things with thoughts: Brain-computer interfaces and disembodied agency. *Philosophy & Technology, 32*(3), 457–482.

https://doi.org/10.1007/s13347-018-0308-4

Wander, J. D., Blakely, T., Miller, K. J., Weaver, K. E., Johnson, L. A., Olson, J. D., Fetz, E. E., Rao, R. P. N., & Ojemann, J. G. (2013). Distributed cortical adaptation during learning of a brain–computer interface task. *Proceedings of the National Academy of Sciences, 110*(26), 10818–10823.

https://doi.org/10.1073/pnas.1221127110

Wolpaw, J., & Wolpaw, E. W. (2012). Brain-computer interfaces: Something new under the sun. In *Brain–computer interfaces: Principles and practice* (pp. 3–12). Oxford, United Kingdom: Oxford University Press.

https://doi.org/10.1093/acprof:oso/9780195388855.003.0001