Contested Visions of Cyborgs: Sociotechnical Futures in the Field of Brain-Computer Interfaces

Roberto Favalli

Corresponding author

Roberto Favalli Department of Philosophy, Sociology, Education and Applied Psychology Via Cesarotti 10/12, 35123, Padova, Italy Sorberto.favalli@phd.unipd.it

Submitted: April 26, 2022 Accepted: March 23, 2023

Abstract

Brain-computer interfaces (BCIs) are a cutting-edge neurotechnology that allows direct interaction between the neural system and external electronic systems. Over the past decade, the field of BCI has witnessed far-reaching promises and expectations, especially regarding the merging of humans and machines. Drawing from the STS literature analysing the relationship between technological innovation and representations of the future, the paper looks at the field of BCIs and at the related expert-scientific debate as a discursive arena where particular visions of cyborg are enacted. Through the notion of visions, representations of the future and discourses of BCIs are explored and analysed within the scholarly literature. Whilst the future of these technologies is usually approached through the question: "Are we all going to become cyborgs?", the question addressed in this contribution is the following: "Which kind of cyborgs are enacted here?". The findings reveal two visions of cyborg; the first considers the subject's intentionality as a key element in steering the human-machine interaction, while the second looks at human-machine entanglement as a machine adaptation that takes place beyond the subject's intentionality. A detailed analysis also shows how these visions are enacted as assemblages of different discursive repertoires and associations mediating and shaping specific cyborg visions.

Keywords

cyborg; future visions; sociotechnical futures; science and technology studies; human-machine entanglement; brain-computer interfaces.

1. Introduction

In 2003, renowned engineer Kevin Warwick stated that "the era of the cyborg is now upon us" (2003, 131). In 2005, the technologist Raymond Kurzweil argued that "by the time of the singularity, there won't be a distinction between humans and technology"¹ (2005, 69). These are just two illustrative expressions of a broader contemporary myth that views humanity as marching towards an intimate amalgamation of humans with technologies. These narratives, often presented as "post-human" or "trans-human", outline a future in which humans will no longer be just strictly biological bodies, but will turn into human-machine hybrids, or

Favalli

cyborgs (Coenen 2007; Jasanoff 2016). Against these highly speculative future discourses and as an effort to problematise current conceptualisations of cyborgs, this paper maintains the necessity to look closer at the actual material-discursive settings where visions of cyborgs are enacted. More precisely, the paper looks at the field of *brain-computer interfaces* (BCIs), also called *brain-machine interfaces*, as one the most cutting-edge emerging technoscientific fields where novel visions of cyborgs are envisioned.

Since the 2000s, an ecosystem of scientific and industrial actors around BCIs has emerged. In 2013, the dedicated scientific journal *Brain-Computer Interfaces* was founded. Thus, in 2015 the BCI Society was established with the aim of "connecting BCI-related organisations and individuals"². In more recent years, the entrance of big-tech companies into the field, such as Meta, Microsoft, and Elon Musk's Neuralink, has expanded public interest and media coverage of BCIs.

BCIs can be defined as neurotechnological devices that connect a biological brain to a computer in real-time. BCIs operate by tracking the user's electrophysiological brain activity and translating it into signals to interact with external devices (such as a personal computer or prosthetic arm) without activating muscles or peripheral nerves. Even though BCIs are still mostly confined to the laboratory, within the expert-scientific debate the technological artefact has been envisioned for several different actionable applications. Initially, BCIs were considered for the medical-clinical area to provide alternative forms of communication and control of the external environment for subjects with disabilities, such as moving a cursor, steering a wheelchair, or operating a speech synthesiser. More recently, the envisioned applications have also included contexts such as work environments, wellness, entertainment, art, and virtual reality. Regardless of the particular application envisioned, John Donoghue, one of the leading experts in BCIs, writes: "Nearly all in the field will agree that one major goal of BCI research is to create a bridge from the brain to the outside world" (2008, 512). Similarly, the neurotechnologist Gerwin Schalk's (2008) overview of the potentials of the field emphasises the "brain-computer symbiosis" that BCI technologies will allow.

This paper relies on the science and technology studies (STS) literature about the relation between technological innovation and the future (Konrad et al. 2017; Lösch et al. 2019; Crabu and Magaudda 2022). By mobilizing such theoretical perspective, it looks at the field of BCIs, and at the related scientific debate, as a discursive-representational arena where particular cyborg visions are enacted, outlined, and contested. In this way, cyborg visions are considered as performative instances that envisions specific modes of human-machine entanglement or hybridisation (Heffernan 2019).

More in details, the paper investigates *which kind of human-machine entanglement representations emerge from the discourses that circulate within the technoscientific field of brain-computer interfaces*; or in other words, *which kinds of cyborgs are enacted here?* In addressing this question, the paper will focus on the expectations and future-oriented visions outlined within scientific publications on BCIs. Furthermore, to fully understand the specific modes of human-machine entanglement that circulate within the BCI technoscientific field, particular attention will be given to the socio-technical dimension and the configurations of hybrid agency enacted by the discourses through which these cyborg visions are articulated.

The next two sections present both the conceptual scaffold and the methodological framework behind this study. Then, after a brief overview of the historical development of BCIs, the paper will discuss and analyse the different visions of the cyborgs identified in the scientific debate.

2. Analytical Framework: Visions and Cyborgs

Since the late 1990s, social sciences have re-engaged in the analysis of the social, cultural, and political aspects of the future (Beckert and Suckert 2021). Consistent with this increasing attention, STS scholars have refined a wide set of perspectives to analyse how futures play a fundamental role in the perception and imagination of emerging technologies, in the material structuration of a technoscientific field, and in "the 'doing' of innovation, from the laboratory to funding and policy agencies" (Konrad et al. 2013, 5).

In the STS domain, representations and discourses about the future are not considered as merely speculative claims, but as historically and culturally contingent discourses that may play a key role in current innovation processes. From this perspective, futures are defined as performative, namely "expectation statements are not only representations of something that does not (yet) exist, but they also do something: advising, showing direction, creating obligations" (van Lente 1993, 191).

In this regard, the concept of "prospective structure" has been developed (van Lente and Rip 1998) to emphasise how a particular vision of the future can become dominant in a technological field, thus guiding (and constraining) the innovation processes. Furthermore, the shaping of future-oriented visions may become a matter of controversy between different contested futures promoted by competing constellation of actors (Brown et al. 2000). Hence, the notion of *arena* is usually adopted to refer to different social settings where expectations and promises are launched, transformed, contested, and affirmed (Bakker et al. 2011).

To investigate which kind of visions of cyborgs circulate within the BCIs academic debate, the paper mostly refers to STS works that have developed the notion of "vision" (Hedgecoe 2003; Lösch et al. 2019). A vision can be defined as "a framework within which the future shape and application of a technology are constructed" (Hedgecoe 2003, 355). Visions are shared by a range of actors and articulate socio-technical futures in which techno-scientific potentials are coupled with the anticipation of particular social changes.

Additionally, in line with recent theoretical insights (Alvial-Palavicino 2016; Schneider and Lösch 2018), the concept of vision is here adopted to take distance from a perspective centred on actors and their strategic mobilisation of expectations, and instead to understand visions as precariously emerging from the ongoing interactions between the heterogeneous elements that constitute the innovation process. Accordingly, in this paper the concept of vision is used to consider specific statements and scenarios as representational elements of a broader assembling process that can be called a "visionary assemblage". The notion of *visionary assemblage* relies on the work by Law (2004), who mobilises the notion of assemblage to analytically grasp:

a process of bundling, of assembling, or better of recursive self-assembling in which the elements put together are not fixed in shape, do not belong to a larger pre-given list but are constructed at least in part as they are entangled together. (Law 2004, 42)

Consistent with this, the concept of visionary assemblage highlights the fact that visions can be seen not as static outcomes of previous social construction processes, but rather as processes in themselves. The visionary assemblage can then be defined as a continuously re-enacted system of associations in which heterogeneous entities (both technical and social, human and non-human) are discursively interwoven, performing a precarious ordering effort through the enactment of specific socio-technical futures. In other words, a vision only exists in terms of the heterogeneous and evolving system of semiotically drawn associations through which a particular future is continuously outlined.

In this way, the concept of visionary assemblage allows for a dialogue with actor-network theory (ANT), in particular by mobilising the concept of "actor-world" (Callon 1986; Rip 2009) to trace the semiotic work through which visionary assemblages are enacted. According to Callon (1986), if an actor-network can be defined as a collective of heterogeneous entities that comes to act as a whole through a chain of material-semiotic associations (Latour 2005), an actor-world is a projection of a future actor-network, of a future world (Rip 2009). The actor-world is part of the actor-network. It is a semiotic construction that, through its circulation, plays a role in connecting and holding together the different bits of an actor-network. In these terms, future-oriented visions can be conceptualised as actor-worlds enacted through the semiotic work performed by the actors comprising the innovation network. In this paper, these actors are mainly the researchers and developers of BCIs who voice statements and expectations. In any case, the enactment of an actor-world is to be understood as the emerging effect of different and intricate social and technological arrangements, rather than the construction of individual actors. Hence, the emergence and the structuration of a technoscientific field such as that of BCIs can be read as the "enactment of overlapping and contrasting actor-worlds" (Rip 2009, 407). Furthermore, the actor-world must be inscribed materially to increase its durability and circulation capacity.

In this paper, the production and circulation of scientific texts is treated as a fundamental practice for the enactment of socio-technical visions. Through scientific texts, certain actor-worlds are enacted and mobilised in the attempt to enrol heterogeneous entities and readers into a particular representation of the future that they will possibly contribute to extending and stabilising.

With respect to these visions as actor-worlds, the paper focuses on which modes of human-machine hybridisation (or cyborg visions) are shaped along with, and as part of, the networks of associations through which the visions themselves are semiotically enacted. Following this conceptual framework, this paper will provide an understanding of the visions of the cyborg that circulate within the "BCI community" – with respect to which the related scientific literature can be treated as a forum for practitioners (van Lente 1993, 97) where agendas are built, applications are envisioned, and authors try to capture attention and interest.

Given the different meanings attributed to the notion of cyborg within the social sciences, it is necessary to specify what is meant here by this term. Caronia defines the cyborg as "an imaginary figure that signals a real process, a change in the relationship between human beings and technology" (2020, 96). Consequently, the concept of "cyborg" is adopted to look at the culturally situated way of imagining human-machine entanglement or hybridisation.

The relevance in the social sciences of the theoretical and political re-appropriation of the term "cyborg" is acknowledged especially in the critical readings of Haraway (1991), where the term is adopted to stress, and overcome, the boundaries between humans and animals, organisms and machines, nature and culture. Moreover, within STS there is also a recognition of what can be referred to as epistemologies of hybridity (Lipp and Dickel 2022), which

promote a theoretical framework focused on the entanglement between humans and artefacts, and where the notion of cyborg is adopted to rethink the phenomenon of agency as the hybrid outcome of configurations of human and non-human entities.

At the same time, these epistemologies may risk lead to a "naturalisation of *cyborgisation*". In fact, assuming the "human-machine hybrid" as an overarching aspect of modernity (Haraway 1991) – if not of the entire human species (Clark 2003) – comes with the risk of overshadowing the specificities of, and the difference between, situated and emerging modes of human-machine hybridisation; or, to put it differently, of specific kinds of cyborg that are envisioned and enacted within specific technoscientific cultures. Thus, the concept of "cyborg" is adopted here to identify different (visions of) cyborgs based on the ways in which the roles of the human and the machine are defined and the interplay between them is characterised.

Furthermore, in the naturalisation of *cyborgisation* the human-machine hybridisation is treated as a "inherent" normal feature of the current societies. In this way, there is an inherent tendency to overlook the specificity of the existing relationship between the human-machine entanglement and its projection into the current representation of the not-yet. A relationship that appears so relevant in many contemporary narratives and imaginaries circulating in popular culture and mainstream media, where the notion of cyborg is often intertwined with the socio-technical imaginary of technological enhancement (Coenen 2007; Heffernan 2019). Lastly, it is important to stress how these considerations are relevant not only from an academic perspective, but also from a wider societal perspective. The development of a consistent conceptual framework to go beyond the naturalisation of cyborgisation appears increasingly necessary, especially considering the potentially disruptive trajectories of emerging technological innovations. Indeed, it seems urgent to understand the actual complexities and risks involved in the widespread imaginaries of "becoming cyborgs" and to promote a gaze that can critically address the political, ethical, as well as economic circumstances involved in their circulation.

3. Methodology

The analysis of the visions that circulate within the BCI field draws on a sample of scientific texts, mainly review articles. Review articles can be considered as a particular form of academic article that aspires to provide an overall systematisation of the scientific debate about a particular issue or technoscientific domain. As Hedgecoe (2003) and Weiner and Martin (2007) highlighted, these scientific outcomes can play an active role in the process of construction and circulation of socio-technical futures and visions. After the first round of analysis, other documents (6 papers, 4 handbooks, 2 roadmaps) referred to in the review articles were added to the corpus. The procedure that led to the construction of the corpus is outlined in Fig.1. The corpus was initially analysed following an exploratory approach using MAXQDA2022. After the first round of coding, subsequent rounds and in-depth analysis focused on the discourses that explicitly (e.g., scenarios, examples) and implicitly (e.g., technical definitions, comparisons with other technologies or applications, frames, legitimation strategies) contribute to the semiotic assembly of socio-technical visions.



Figure 1. Procedure for the construction of corpus.

4. Historical contextualization on BCIs

According to the historiography that circulates in academic publications, the science behind BCIs began about 100 years ago, with a German professor of psychiatry named Hans Berger. In 1929, in Berger's paper entitled "About the Human Electroencephalogram" (original title: "Uber das Elektrenkephalogramm des Menschen"), the term *electroencephalography* (EEG) was for the first time adopted in a scientific publication, along with speculations on the possibility of reading human thoughts from detected EEG brain waves (Borck 2018). Within the BCI community, the development of *brain-computer interfaces* is framed as a fulfilment of the possibilities initiated by Berger's work. This is nicely exemplified by an extract from a widely adopted handbook:

This possibility – that people could act through brain signals rather than muscles – has fascinated scientists and nonscientists alike for many years. Now, nearly a century after Berger's epochal discovery, possibility is becoming reality. (Wolpaw and Wolpaw 2012, 9) After Berger's EEG, the next and main "father" of the field was identified in Jacques J. Vidal, a Belgian researcher who worked at UCLA University. By the 1970s, Vidal led the "Brain-Computer Interface project" within a larger program funded by ARPA and the U.S. Department of Defense interested in evaluating the possibility of adopting biological signals to control computers, vehicles, weaponry, and other systems (Vidal 1999). In this context, the term "brain-computer interface" was used for the first time. In a 1973 paper named "Toward Direct Brain-Computer Communication", which still represents one of the most renowned publications in the BCI community, Vidal claims:

Can these observable electrical brain signals be put to work as carriers of information in man-computer communication or for the purpose of controlling such external apparatus as prosthetic devices or spaceships? Even on the sole basis of the present states of the art of computer science and neurophysiology, one may suggest that such a feat is potentially around the corner. (Vidal 1973, 157)

Nonetheless, the field remained slightly uncertain until the mid-1980s. Then, during the 90s a small group of researchers from the United States and Europe spearheaded the BCI field by introducing the first real-time and working brain-computer interfaces and developing approaches and techniques that are still used today. In 1999, the first BCI international meeting was held in Rensselaerville, New York. "Fifty scientists and engineers participated. They represented 22 different research groups from the United States, Canada, Great Britain, Germany, Austria, and Italy" (Wolpaw et al. 2000, 1). From 2000 to the present, the BCI researcher-expert community has experienced an exponential expansion in terms of both BCI peer-reviewed publications and attendees at BCI conferences and other related events (Nam et al. 2018). In the following sections, two different visions of BCIs will be presented, discussing how these visions are assembled in the interwining of other discourses and elements that intersect the field of BCIs, anticipating specific types of cyborg.

4.1 Active Vision: Communicating and Controlling

Especially during the 90s, BCI research gradually came to coincide with the promise of developing systems to support severely disabled patients suffering from conditions such as amyotrophic lateral sclerosis (ALS) or locked-in syndrome. In this historical phase, the demarcation process of the boundaries of the BCI field is particularly instantiated within the first conference of the concerned BCI researchers-experts community in the 1999. Indeed, with the first international meeting on BCIs, a vision – that can be referred to as "Active BCIs" – starts being formally articulated along with the envisioning of a particular type of cyborg. As stated in the review on the first BCI international meeting:

Brain-computer interfaces give their users communication and control channels that do not depend on the brain's normal output channels of peripheral nerves and muscles. Current interest in BCI development comes mainly from the hope that this technology could be a valuable new augmentative communication option for those with severe motor disabilities. (Wolpaw et al. 2000, 164) Hence, one of the key aspects of the active vision is to trace an association between the future of BCIs and the future of people with severe disabilities. In this way, the vision operates performatively by articulating an obligatory passage point (Callon and Law 1992), which structures the future landscape by prioritising certain future trajectories at the expense of alternatives. This prospective structuring involves effects at the level of enrolment dynamics. First, the enrolment of persons with disabilities as the main reference for BCIs' envisioned applications. Particularly those forms of disability (such as the "complete locked-in syndrome", i.e., CLIS) for which BCIs are depicted as the only potential solution to interact with the outside world, since every other voluntary muscle control on which conventional assistive technologies depend are precluded. Second, researchers interested in entering the BCI field must learn (or at least consider) that the field is explicitly oriented towards the development of solutions for people with severe disabilities, and this will affect their courses of action.

Furthermore, by linking the emerging technological artefact (BCIs) with its potential social impact (overcoming disabilities), the active vision helps to reconfigure the BCIs innovation process as a protected space (van Lente and Rip 1998; Konrad et al. 2017), namely an innovation niche where the development of a new technology is perceived in a positive light, the resources invested are deemed legitimate, and evaluation standards can be relaxed despite technical challenges. This performative effect of the vision on the innovation process is also supported by a number of national surveys suggesting that medical and assistive applications are the most well perceived use of BCIs among the general public (Sample et al. 2020).

The association of BCIs with persons with severe disabilities is explicitly evoked within most of the reviews by means of captivating titles, such as: "Brain-computer communication: Unlocking the locked-in" (Kübler et al. 2001) or "Breaking the silence: Brain-computer interfaces for communication and motor control" (Birbaumer 2006). This association not only plays a key role in the enactment of a particular future vision of BCIs, but also structures the discourses on the socio-technical future of BCIs along certain trajectories, allowing certain modes of human-machine hybridisation to be imagined and silencing alternative visions of cyborgs, e.g., those that consider applications outside the medical field because they are considered unethical (Nijboer et al. 2011).

Regarding the kind of cyborg outlined by the active vision, a first aspect to highlight is how, with respect to the human-machine interplay, a predominant role is given to the human actor. The 2002 review "Brain-computer interfaces for communication and control" – which at 5,828 citations³ represents the most cited BCI-related article – provides what is generally considered the first technical definition of BCI:

A BCI is a communication system in which messages or commands that an individual sends to the external world do not pass through the brain's normal output pathways of peripheral nerves and muscles. (Wolpaw et al. 2002, 769)

and:

BCI operation depends on the interaction of two adaptive controllers, the user, who must maintain a close correlation between his or her intent and these phenomena [variations in

electrophysiological signals], and the BCI, which must translate the phenomena into device commands that accomplish the user's intent. (ibid., 770)

This conceptualisation of hybrid-distributed agency stresses an idea of linearity, where signals move from the human actor to the external world and where the machine is represented as a channel that mediates the voluntary command of subjects over the environment.

The cyborg envisioned here is also heavily characterised by the emphasis on intentionality. For example: "Control should emerge from the voluntary intent to carry out an action" (Donoghue 2002, 4), or: "Successful operation of brain-computer interfaces depends significantly on the degree to which neural activity can be volitionally controlled" (Fetz 2007, 571). Furthermore, the issue of the intentional control of the user on the external devices is also usually associated with the necessity of learning. The human ability to control a BCI therefore is intended as a skill to be learned: "Individuals are extensively trained to intentionally control certain aspects of recorded brain activity" (Haynes and Rees 2006, 524).

The shape of this active vision can also be traced by reference to what is generally defined as "not a BCI": "Devices that only passively detect changes in brain activity that occur without any intent [...] are not BCIs" (Graimann et al. 2010, 3), or: "Brain-computer interfaces do not read minds in the sense of extracting information from unsuspecting or unwilling users but enable users to act on the world by using brain signals rather than muscles" (Shih et al. 2012). The definition of boundaries with respect to "what is" and "what is not" a BCI is also part of the semiotic construction of a particular cyborg vision.

The active vision is therefore based upon the following dimensions: the enrolment of people with severe disabilities as end-users; the mobilisation of the disability as the main legitimising frame; the intentional control of external devices as envisioned applications; and the process of learning BCI as a necessary practice in current and future socio-technical worlds, as well as the efforts to specify how BCIs are not "mind-reading devices".

Finally, the mode of human-machine entanglement envisioned along with active vision is also mediated by descriptions of what BCIs might enable in the far future. Within the active vision, these depictions extend the use of BCIs from individuals with severe disabilities – who nevertheless remain the principal end-users – to less severe disabilities or healthy individuals. Usually, the argument for this extension is that if BCIs can help people to regain movement and sensation today, imagine what can be done in the years ahead.

This glimpse into the more distant future comes about as an extension within the same prospective structure, hence without questioning the "communication and control" trajectory. This extension is also enacted through the enrolment of a quantitative parameter within the visionary assemblage, the *information transfer rate* (ITR), through which specific future developments of BCIs are anticipated. While for severe disabilities "even the modest rates of communication that will initially be achieved should dramatically improve quality of life" (Schalk 2008, 10), "the future value of BCI technology will depend substantially on how much information transfer rate can be increased" (Wolpaw et al. 2002, 779). Here, future projections for BCIs are associated with quantifiable parameters, narrowing the future land-scapes to expected technical developments, and simultaneously marginalising the envisioning of qualitatively alternative applications or scenarios.

The enrolment of "healthy users" in the envisioned future world related to the active vision also entails an extension from the medical-clinical scenario to that of technological human enhancement. It is through the augmentation of human capabilities that both disabled and "healthy" individuals overcome the limitations of their bodies. The difference between therapy and empowerment itself appears blurred and relies on quantitative criteria (e.g., ITR, number of electrodes) rather than qualitative differences. The same trajectory through which BCIs would enable people with disabilities to interact with the outside world is envisioned as adoptable in the more distant future, and along with an increase in *information transfer rate*, by healthy people to enhance their *communication and control* abilities.

4.2 Passive Vision: Monitoring and Adapting

Around the second half of the 2000s, an alternative visionary assemblage for BCIs began to be discussed in the expert-scientific discourses, in conjunction with which a different kind of cyborg is envisioned. This vision will be referred to as "passive vision". The 2011 review by Zander and Kothe entitled "Towards passive brain-computer interfaces" is considered by the BCI community as the main intermediary for early passive BCI articulations. The authors write:

A passive BCI is one that derives its outputs from brain activity arising without the purpose of voluntary control, for enriching a human-machine interaction with implicit information on the actual user state. (2011, 3)

and:

it can be seen as modifying the general approach of BCI and substituting the usually voluntary and directed command with passively conveyed implicit information. [...] The resulting approach of passive BCI opens up the field of applications based on BCI technology to a broader context, especially for using it also for healthy users. (ibid., 2)

Here, it is worth noting that the enactment of an alternative visionary assemblage in the field of BCI entails not only a new perspective, but also a redefinition of the already existing visions. To put it another way, the actor-world projected with passive vision is made up of a variety of interconnected elements, one of which is a specific translation-redefinition of active vision. Proponents of the passive vision frequently present it as an alternative to the dominant active vision, highlighting the limitations of the latter.

Compared to the active vision, in the passive vision there is a fundamental shift from patients to users, i.e., from people with disabilities to "healthy users". Firstly, the passive vision underlines the limitations of communication and control applications, loosening the enrolment of healthy users within the active vision:

it should be taken into consideration that BCI for healthy users aims at partially different applications than BCIs for disabled users. In particular, direct input primarily for communication and control seems not to be the most promising BCI-related application for healthy users, due to the still low reliability and bandwidth of current BCI systems compared to standard communication channels. (Zander and Kothe 2011, 2)

Secondly, a new legitimising frame that associates the technology with benefits for the broader society is mobilised:

Mental state monitoring is of particular interest in safety-critical applications where human performance is often the least controllable factor. For example, consider that fatal car accidents are one of the leading causes of death in the United States and the leading cause among children (9-18 years) worldwide. (Blankertz et al. 2010, 7)

Indeed, despite abandoning the promise to address cases of disability, the passive vision seeks to maintain the "protected space" by still signifying BCIs as a future technological solution to pressing social issues – for example, in critical situations (such as driving a car or a surgical operation) where they could reduce the impact of human errors and support human decision-making processes.

Another key aspect of passive vision is that of *unintentionality*, which plays a significant role in the assembly of a different kind of cyborg. In passive vision, it is the technical system itself that identifies the user's spontaneous brain activity, rather than the user voluntarily controlling the technical device. In the scientific literature, this feature is associated with and supported by two types of discursive repertoires, which will be referred to as *unobtrusiveness* and *smartness*. These discourses, mobilised as argumentative logics in favour of passive BCI, can be analysed to reconstruct the visionary assemblage and reasoning around the type of cyborg envisioned here.

Unobtrusiveness maintains that detecting spontaneous (and thus non-voluntary) brain activity allows researchers to bypass the influence of the subjects themselves on the signal, and thus allows for more objective measures (Blankertz 2016, 9). Additionally, structuring a protected space for emerging technology is often coupled with redefining other previous and competing technologies in terms of their limitations and lacks (Brown et al. 2000). While in the active vision BCIs are compared primarily with other assistive technologies, in the passive vision they are compared with other measurement tools. For example:

Traditional methods for capturing mental states and user ratings are questionnaires, video surveillance of the task, or the analysis of errors made by the operator. However, questionnaires are of limited use for precisely assessing the information of interest, as the reported answers are often distorted by subjectiveness. Questionnaires cannot determine the quantities of interest in real-time but only in retrospect; moreover, they are intrusive because they interfere with the task. (Blankertz 2010, 7).

Thus, even this change in the application landscape – from "medical" to "general measurement and monitoring" – plays a role in shaping the passive vision. *Smartness* refers to the argument that:

the use of modern machine learning and signal processing methods allowed to relocate the burden of training from a learning subject toward statistical learning machines and thereby achieve BCI communication for a naïve user already in the first session. (ibid., 1)

In other terms, the *smartness* discourse of the passive vision is discursively performed as a shift from the "BCI as a skill to be learned" to the motto "let the machine learn" (Blankertz et al. 2006, 583).

In this way, the passive vision associates BCIs with the socio-technical imaginary of *auto-mation*, mediating a BCI future trajectory oriented toward reducing human agency – and "human error" – within human-machine interaction. This change is consistent with the broader visionary assemblage, considering that in the passive vision the envisioned end-users are healthy users. As opposed to the case where end-users were people with disabilities, here the expectation shifts towards reducing the users' training burden. This aspect is seen by the promoters as critical to technology adoption by a broader audience. It is also interesting that the proposed notion of passive BCI is defined from the user's perspective, since "passive" in fact refers to the position of the human with respect to the machine. If we look at the different weights attributed to the main entities (human and machine) involved in the cyborg-hybrid agency, it is in fact the machine that takes on a more active role.

By looking at representations of the far future, the passive vision anticipates future applications where the subjects of the enhancement are not humans (at least not directly) but rather the "machine":

BCI technology is used for detecting the state of the user in a given human-machine system and for augmenting the information space available to the system with context information about the user. (Zander 2011, 2)

This future is extensively represented through far-future scenarios mobilised within the scientific literature:

It seems worthwhile to employ BCIs to infer implicit information during software usage and to use that information to augment the explicit interaction. In other words, to make the computer better at understanding the human user on the basis of soft skills. (Blankertz 2016, 10)

or:

a system sensing a user getting verbally overloaded could attempt to turn down the music, since musical lyrics get subconsciously processed and consume valuable verbal resources. Or perhaps the cell phone could alert the remote speaker and pause the phone call if the driver has to suddenly focus on the road. (Tan and Nijholt 2010, 15)

The kind of cyborg envisioned within the passive vision is quite distant from the "classic cyborg" conceived as a human subject that merges with technological devices to enhance its capabilities. Instead, the hybrid agency is configured as the capacity of a distributed computer system to monitor and adapt in real-time to changes in the human part, beyond the intentionality of the involved subjects. The human part is here marginalised and disconnected from the

attribute of intentionality and autonomy, thus becoming in a way part of the environment with respect to which the computer-machine part acquires the information necessary to refine its ability to adapt autonomously. Therefore, the main object of enhancement within the passive vision is not the human component, but the technical system. As Zander writes:

Neuroadaptive systems can be said to be systems with an agenda, having a goal of their own. By autonomously initiating each interaction cycle using a specifically selected probe stimulus, they would be in a position to "guide" the interaction such that specific information can be gathered, and to change the interactive experience based on that or other information. (2016, 5)

Here, the machine-artefact is not a prosthesis that mediates and extends human action in the external environment, but rather stands as an adaptive and autonomous interface between human actors and their worlds.

5. Discussion: Back to Cyborg Visions and Beyond

The previous paragraphs highlighted how along with each of the considered visions that circulate in the BCI field (the active and passive one) a particular mode of human-machine entanglement, or a vision of the cyborg, is also envisioned. Following the concepts of visionary assemblage and actor-world, these visions were reconstructed and examined by tracing the associations with different discursive repertoires and entities through which they are enacted. It was thereby shown that these cyborg visions are shaped in the semiotic interweaving of different elements, such as expectations, interests, artefacts, technical aspects, legitimation strategies, and different anticipated applications and publics.

For example, within the active vision, the primary role attributed to individuals with severe disabilities as end-users, together with "communication and control" as the near-exclusively envisioned applications, heavily contribute to the enactment of a cyborg-hybrid agency configured as the human capacity to voluntarily control technological devices. In contrast, in the context of passive vision, BCIs are envisioned for applications that directly affect the broader society, from monitoring mental states in occupational and risky environments to developing neuroadaptive technologies (Zander 2016). This latter vision invokes a future-world at the centre of which there are no longer patients with disabilities, but healthy subjects represented as unwilling to endure long learning times or apply large amounts of effort. Therefore, the criterion of intentional control of external devices is replaced by the ability of the machines themselves to adapt to human subjects, thus anticipating a different human-machine entanglement. Here, the concept of visionary assemblage suggests that both the type of cyborg envisioned and the networks of heterogenous associations articulated with the enactment of the concerned vision are assembled as aspects of the same movement.

Furthermore, the concept of visionary assemblage emphasises that socio-technical visions are not only assembled, namely semiotically enacted and mobilised, but that they are also assembling, which indicates that through their circulation they actively participate in the configuration of the technoscientific field to which they are linked. In fact, visions participate in

the structuring of the field along particular innovation trajectories by attracting resources, legitimising investments, defining shared agendas, downplaying alternative trajectories, and interesting other actors who will contribute to the expansion and articulation of the vision itself.

For instance, the active vision, whereby BCIs are primarily configured as medical and assistive technologies, operates as an *interessement* device (Akrich et al. 2002) that enrols, in addition to the patients themselves, also the investments of economic actors operating in the medical-clinical field, the interests of institutions promoting the issue of disability support (e.g., BCIs have in fact been included among the *future and emerging technologies* [FETs] by the European Research Council) but also – redefining the boundaries of the field itself – experts from other fields, such as neurosurgeons, rehabilitative physicians, and assistive technology engineers. Consistently, within the active vision, the passive vision is marginalised as unethical since it prioritises the general consumer before patients.

In the case of passive vision, the *interessement* dynamics primarily involve general consumers, especially innovative device enthusiasts, but also companies looking for investment opportunities in the user experience/user interface design (e.g., the recent involvement of Meta and Microsoft) or in the gaming sector (e.g., the interest of Valve⁴ in BCIs). Concerning experts from other fields, the passive vision extends the field not toward the medical area but by involving different figures such as human-computer interaction experts, dry electrode producers, designers, and entrepreneurs especially in the realm of wellness, gaming, and wearable technologies. Conversely, the active vision is often silenced as difficult to implement technically, as the reliability and bandwidth are too low to actually "control objects with the mind". It is interesting to note that technical limitations in the extraction and transmission of the brain signal, and thus in the fulfilment of active vision, also seem to play a role in the shaping and unfolding of passive vision.

Thus, simultaneously, a change in the semiotically traced networks leads to a change in the content of the vision, just as a change in the actor-world projected by the vision leads to changes at the level of the current system of associations and enrolments. Each vision articulated in the field of BCIs is collectively enacted by actors in the field and simultaneously participates in redefining the field towards different trajectories of innovation. Indeed, around each of these emerging trajectories, visions seem to play a pivotal role in holding together the different assemblages of actors, artefacts, discourses and imaginaries through which the innovation trajectories are materialised.

Interestingly, on a strictly technical level the differences between active and passive BCIs are quite blurred. The two modalities share the same building blocks (signal extraction, processing, and translation) and similar protocols and approaches are adopted in research and development contexts. Even the difference between intentionality and unintentionality – so crucial in the discursive articulation of the two visions – does not hold much weight when considered from a strictly "technical" perspective. In fact, from a neuroscientific gaze, the very concept of intentionality takes on hazy and insubstantial traits (Pickersgill 2011). The "active control" of active vision is far from being an established assumption among physiologists and neuroengineers. The same applies to the "passivity" of the healthy subject within the passive vision. As Wolpaw puts it:

passive and active are subjective terms that lack clear neuroscientific definitions. Furthermore, continued use of a passive BCI might well induce CNS adaptations that improve its performance, so that the term passive becomes no longer applicable. (Wolpaw and Wolpaw 2012, 6)

Despite this, these two visions continue to play a fundamental role in organising dynamics and discourses within the field. Research streams, projects, and devices are still discussed and promoted following the two visions articulated in the scientific literature and in other arenas such as scientific conferences and technological events. In fact, looking at the main BCI-related conferences, passive and active BCIs are usually discussed not only in different panels, but - especially recently - also in different conferences where distinct networks of actors, expertise, imaginaries, and issues are intertwined, such as novel medical applications, surgical procedures, and human enhancement for the active vision, and neuromarketing, affective computing, and IoT for the passive vision. Therefore, visions are enacted not only through scientific publications but also in other arenas where they are continuously articulated, shared, inscribed in texts, and promoted while seeking funds, attention, and recognition. Socio-technical visions seem to work as a bridge that connects the same technological object (the BCI, when considered strictly from a technical perspective) to different assemblages of actors, discourses, and imaginaries, thereby enacting different cyborg futures. Furthermore, in addition to an organisational function, the persistence of the visions can also be attributed to a rhetorical adoption. For example, promoters may define their products or their research under the notion of "BCIs" to refer to devices - such as headbands for monitoring brain activity - that under the more traditional definition of BCIs would not have been defined as such, and in this way, besides participating in the dissemination of the passive vision, benefit from the advantages in terms of attention and resources derived from the hype over BCIs in general. Otherwise, especially in outreach settings, BCIs are typically promoted by focusing on the excitement of controlling objects with the mind and on the technology's potential to address disabilities, avoiding a popular depiction of BCIs as "mind-reading devices". This portrayal of BCIs tend to overshadow the passive vision and shields its potentially more problematic acceptance from public and policy discussions. The possibility of this rhetorical exploitation of futures, consistently with the concept of the visionary assemblage, again shows how visions, far from being an imposing structure steering the actions of actors, can be better understood as part of an assembling process in which visions and actors are both involved in the continuous re-ordering of the technoscientific field and its trajectories of innovation.

6. Conclusion

With the intention of problematising contemporary narratives predicting an inevitable future in which humans will merge with technologies and become cyborgs, the field of *brain-computer interfaces* was examined more closely as an arena where potential ways of relating humans and machines are envisioned and rearranged – that is, an arena where potentially novel visions of cyborgs that inform current innovation processes are articulated. Drawing on an understanding of cyborgs as specific modes of human-machine hybridisation, two different

Favalli

visions of cyborgs were identified: on the one hand the active vision, where the cyborg is configured in terms of the human's ability to intentionally control a machine solely through his or her mind. This involves the use of artificial prostheses connected to the human's brain, allowing human subjects to compensate for their limitations or enhance their abilities while still maintaining control over the machine. On the other hand, there is a passive vision whereby a machine assumes a more central position and the human-machine entanglement relies on the machine-system's ability to learn and adapt autonomously to the monitored brain states of the human actor beyond its intentionality. Both visions discussed in this article are in continuity with imaginaries circulating in popular culture and the mainstream media, which describe a future in which humans and machines will merge. However, through the concept of visionary assemblage, it was shown how distinct future-oriented visions are enacted in the interweaving of these imaginaries with different discursive repertoires, entities, and different intended audiences and applications. Simultaneously, it has been shown how different visions circulating in an emergent field can participate in the reconfiguration and expansion of the field along different trajectories, through the enrolment of different range of actors, the mobilisation of different discursive repertoires, and the adoption of different imaginaries. Furthermore, it was highlighted how organisational and promotional dynamics contribute to the permanence and continuous re-enactment of the visions, even though, from a technical point of view, the differences between the visions are weak and blurred. Finally, the analysis presented in this article aimed to consider the impact of these visions on society at large by broadening the definitions of cyborgs used in popular and academic conceptualisations. This is especially necessary in view of the need for a critical examination of the imaginaries of the cyborgisation process, the "normalisation" of which risks overlooking the different types of human-machine entanglement articulated in it as well as the political, economic, and promotional logics involved.

Notes

¹Kurzweil refers to the concept of the "technological singularity", a term used in the fields of computer science and science fiction to refer to a hypothetical future point in which technological advancement will lead to a radical, unforeseeable, and irreversible change in human civilisation. In his book *The Singularity is Near* (2006), Kurzweil defines the *singularity* as the transcending of the biological limitations of human beings through merging with artificial intelligence.

² <u>https://bcisociety.org/</u>.

³ On scopus.com on 16/12/2022.

⁴ Valve is a leading provider of gaming software and hardware based in the United States.

References

Akrich, Michel, Callon, Michel, Latour, Bruno and Monaghan, Andrew (2002) The Key to Success in Innovation Part I: The Art of Interessement, in "International Journal of Innovation Management", 6(2), pp. 187-206.

- Alvial-Palavicino, Carla (2016) The future as practice, in "TECNOSCIENZA: Italian Journal of Science & Technology Studies", 6(2), pp. 135-172.
- Bakker, Sjoerd, van Lente, Harro and Meeus, Marius (2011) Arenas of Expectations for Hydrogen Technologies, in "Technological Forecasting and Social Change", 78(1), pp. 152-162.
- Beckert, Jens and Suckert, Lisa (2021) The future as a social fact: The analysis of perceptions of the future in sociology, in "Poetics", 84, pp. 101-499.
- Birbaumer, Niels P. (2006) Breaking the silence: Brain-computer interfaces (BCI) for communication and motor control, in "Psychophysiology", 43(6), pp. 517-532.
- Blankertz, Benjamin, Acqualagna, Laura, D\u00e4hne, Sven, Haufe, Stefan, Schultze-Kraft, Matthias, Sturm, Irene and M\u00fcller, Klaus-Robert (2016) The Berlin Brain-Computer Interface: Progress Beyond Communication and Control, in "Frontiers in Neuroscience", 10.
- Blankertz, Benjamin, Tangermann, Michael, Vidaurre, Carmen, Fazli, Siamac, Sannelli, Claudia, Haufe, Stefan and Müller, Klaus-Robert (2010) *The Berlin brain-computer interface: Non-medical* uses of BCI technology, in "Frontiers in Neuroscience", 4.
- Blankertz, Benjamin, Dornhege, Guido, Krauledat, Matthias and Müller, Klaus-Robert (2006) The Berlin Brain-Computer Interface: Machine Learning Based Detection of User Specific Brain States, in "Journal of Universal Computer Science", 12(6), pp. 581-607.

Borck, Cornelius (2018) Brainwaves: A Cultural History of Electroencephalography, New York, Taylor & Francis.

- Brown, Nik, Rappert, Brian and Webster, Andrew (2000) Contested Futures: A Sociology of Prospective Techno-Science, London, Routledge.
- Callon, Michel (1986) *The Sociology of an Actor-Network: The Case of the Electric Vehicle*, in Michel Callon, John Law and Arie Rip (eds.), *Mapping the Dynamics of Science and Technology*, London, Palgrave Macmillan, pp. 19-34.
- Callon, Michel and Law, John (1992) *The Life and Death of an Aircraft: A Network Analysis of Technical Change*, in Wiebe E. Bijker and John Law (eds.), *Shaping Technology/Building Society: Studies in Sociotechnical Change*, Cambridge, Massachusetts, The MIT Press, pp. 21-52.
- Caronia, Antonio (2020) Dal cyborg al postumano: Biopolitica del corpo artificiale, Milano, Meltemi.
- Clark, Andy J. (2003) Natural-Born Cyborgs: Minds, Technologies, and the Future of Human Intelligence, New York, Oxford University Press.
- Coenen, Christopher (2007) Utopian Aspects of the Debate on Converging Technologies, in Gerhard Banse et al. (eds.), Assessing Societal Implications of Converging Technological Development, Berlin, Edition Sigma, pp. 141-174.
- Crabu, Stefano and Magaudda, Paolo (2022) *Disentangling Futures from a Science and Technology Studies Perspective*, in "TECNOSCIENZA: Italian Journal of Science & Technology Studies", 13(2), pp. 5-12.
- Donoghue, John P. (2008) Bridging the Brain to the World: A Perspective on Neural Interface Systems, in "Neuron", 60(3), pp. 511-521.
- Fetz, Eberhard E. (2007) Volitional control of neural activity: Implications for brain-computer interfaces, in "The Journal of Physiology", 579(3), pp. 571-579.
- Graimann, Bernhard, Allison, Brendan Z. and Pfurtscheller, Gert (eds.) (2010) *Brain-Computer Interfaces: Revolutionizing Human-Computer Interaction*, Springer Science & Business Media.
- Haraway, Donna (ed.) (1991) A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century, in Simians, Cyborgs, and Women: The Reinvention of Nature, New York, Routledge, pp. 149-182.

- Haynes, John-Dylan and Rees, Geraint (2006) *Decoding mental states from brain activity in humans*, in "Nature Reviews Neuroscience", 7(7), pp. 523-534.
- Hedgecoe, Adam M. (2003) Terminology and the Construction of Scientific Disciplines: The Case of Pharmacogenomics, in "Science, Technology, & Human Values", 28(4), pp. 513-537.
- Heffernan, Teresa (ed.) (2019) Cyborg Futures: Cross-Disciplinary Perspectives on Artificial Intelligence and Robotics, Springer Nature.
- Jasanoff, Sheila (2016) Perfecting the Human: Posthuman Imaginaries and Technologies of Reason, in Benjamin J. Hurlbut and Hava Tirosh-Samuelson (eds.), Perfecting Human Futures: Transhuman Visions and Technological Imaginations, Wiesbaden, Springer Fachmedien Wiesbaden, pp. 73-95.
- Konrad, Kornelia E., Coenen, Christopher, Dijkstra, Albert Meijer, Milburn, Caroline and van Lente, Harro (eds.) (2013) Shaping Emerging Technologies: Governance, Innovation, Discourse, Amsterdam, IOS Press.
- Konrad, Kornelia E., van Lente, Harro, Groves, Christopher and Selin, Cynthia (2016) Performing and Governing the Future in Science and Technology, in Clark A. Miller, Ulrike Felt, Rayvon Fouché and Laurel Smith-Doerr (eds.), The Handbook of Science and Technology Studies, 4th Edition, Cambridge, MIT Press.
- Kübler, Andrea, Kotchoubey, Boris, Kaiser, Jochen, Wolpaw, Jonathan R. and Birbaumer, Niels P. (2001) *Brain-computer communication: Unlocking the locked in*, in "Psychological Bulletin", 127(3), pp. 358-375.
- Kurzweil, Ray (2005) The Singularity is Near: When Humans Transcend Biology, London, Penguin.
- Law, John (2004) After Method: Mess in Social Science Research, New York, Routledge.
- Lipp, Benjamin and Dickel, Sascha (2022) Interfacing the human/machine, in "Distinktion Journal of Social Theory", pp. 1-19.
- Lösch, Andreas, Grunwald, Armin, Meister, Martin and Schulz-Schaeffer, Ingo (2019) Socio-Technical Futures Shaping the Present: Empirical Examples and Analytical Challenges, Wiesbaden, Springer Nature.
- Millán, José del R., Rupp, Rüdiger, Müller-Putz, Gernot R. and Mattia, Donatella (2010) *Combining* brain-computer interfaces and assistive technologies: State-of-the-art and challenges, in "Frontiers in Neuroscience", 161.
- Nam, Chang S., Nijholt, Anton and Lotte, Fabien (eds.) (2018) Brain-Computer Interfaces Handbook: Technological and Theoretical Advances, Florida, CRC Press.
- Nijboer, Femke, Clausen, Jens, Allison, Brendan Z. and Haselager, Pim (2011) *The Asilomar Survey: Stakeholders' Opinions on Ethical Issues Related to Brain-Computer Interfacing*, in "Neuroethics", 6(3), pp. 541-578.
- Pfurtscheller, Gert, Allison, Brendan Z., Brunner, Clemens, Bauernfeind, Gunther, Solis-Escalante, Teodoro, Scherer, Reinhold and Birbaumer, Niels P. (2010) *The hybrid BCI*, in "Frontiers in Neuroscience", 4, pp. 3-11.
- Pickersgill, Martyn (2011) Connecting neuroscience and law: Anticipatory discourse and the role of sociotechnical imaginaries, in "New Genetics and Society", 30(1), pp. 27-40.
- Rip, Arie (2009) Technology as prospective ontology, in "Synthese", 168, pp. 405-422.
- Sample, Matthew, Sattler, Sebastian, Blain-Moraes, Stefanie, Rodríguez-Arias, David and Racine, Eric (2020) Do Publics Share Experts' Concerns about Brain-Computer Interfaces? A Trinational Survey on the Ethics of Neural Technology, in "Science, Technology, & Human Values", 45(6), pp. 1242-1270.

Schalk, Gerwin (2008) Brain-Computer Symbiosis, in "Journal of Neural Engineering", 5(1), pp. 1-15.

Schneider, Christoph and Lösch, Andreas (2018) Visions in assemblages: Future-making and governance in FabLabs, in "Futures", 109, pp. 203-212.

- Shih, Jerry J., Krusienski, Dean J. and Wolpaw, Jonathan R. (2012) Brain-computer interfaces in medicine, in "Mayo Clinic Proceedings", 87(3), pp. 268-279.
- Tan, Desney S. and Nijholt, Anton (eds.) (2010) Brain-Computer Interfaces and Human-Computer Interaction, in Brain-Computer Interfaces, London, Springer, pp. 3-19.
- van Lente, Harro (1993) Promising Technology: The Dynamics of Expectations in Technological Developments, PhD Thesis, Delft, Eburon.
- van Lente, Harro and Rip, Arie (1998b) Expectations in Technological Developments: An Example of Prospective Structures to be Filled in by Agency, in Cornelis Disco and Barend van der Meulen (eds.), Getting New Technologies Together: Studies in Making Sociotechnical Order, Berlin, De Gruyter, pp. 203-230.
- Vidal, Jacques J. (1973) Toward Direct Brain-Computer Communication, in "Annual Review of Biophysics and Bioengineering", 2(1), pp. 157-180.
- Vidal, Jacques J. (1999) Cyberspace Bionics, in Jonathan P. Marsh, Barbara Gorayska and Jacob L. Mey (eds.), Human Factors in Information Technology, 13, pp. 203-218.
- Warwick, Kevin (2003) Cyborg morals, cyborg values, cyborg ethics, in "Ethics and Information Technology", 5(3), pp. 131-137.
- Weiner, Kate and Martin, Paul (2008) A genetic future for coronary heart disease?, in "Sociology of Health & Illness", 30(3), pp. 380-395.
- Wolpaw, Jonathan R. and Wolpaw, Elizabeth W. (eds.) (2012) Brain-Computer Interfaces: Something New under the Sun, in Brain-Computer Interfaces: Principles and Practice, Oxford University Press, pp. 3-12.
- Wolpaw, Jonathan R., Birbaumer, Niels P., Heetderks, William J., McFarland, Dennis J., Peckham, Paul H., Schalk, Gerwin and Vaughan, Theresa M. (2000) *Brain-computer interface technology: A review of the first international meeting*, in "IEEE Transactions on Rehabilitation Engineering", 8(2), pp. 164-173.
- Wolpaw, Jonathan R., Birbaumer, Niels P., McFarland, Dennis J., Pfurtscheller, Gert and Vaughan, Theresa M. (2002) *Brain-computer interfaces for communication and control*, in "Clinical Neurophysiology", 113(6), pp. 767-791.
- Zander, Thorsten O. and Kothe, Christian (2011) Towards passive brain-computer interfaces: Applying brain-computer interface technology to human-machine systems in general, in "Journal of Neural Engineering", 8(2).
- Zander, Thorsten O., Krol, Laurens R., Birbaumer, Niels P. and Gramann, Klaus (2016) Neuroadaptive technology enables implicit cursor control based on medial prefrontal cortex activity, in "Proceedings of the National Academy of Sciences", 113(52), pp. 14898-14903.