Chapter 13
Philosophical Reflections on Brain–Computer Interfaces

Guglielmo Tamburrini

13.1   BCI as an Exemplary Case Study in the Philosophy of ICT

A Brain–Computer Interface (BCI) processes brain activity online and identifies patterns in this activity that can be used for communication, control, and, more recently, monitoring purposes (Wolpaw and Wolpaw 2012). BCI systems are prime examples of the actual and potential changes that novel information and communication technologies (ICT) are impressing on human–machine interactions, on public debate about the promotion and regulation of technological innovation, and on rational and irrational attitudes towards technological development. This contribution examines the impact that BCI systems are having on these aspects of human life from distinctive philosophical perspectives.

To illustrate, consider the question "Are there good reasons to believe that my BCI system will do the right thing in its operational environment?" The special epistemological interest of this question depends on the *adaptive* character of both BCI systems and the operational environment they are immersed in. Indeed, BCI architectures crucially involve a *learning* component, that is, a computer program which is trained to adapt to and identify activity patterns in the central nervous system (CNS). Since an adaptive BCI interacts with a CNS, in other words with the most complex adaptive system we are acquainted with, their sustained interactions give rise to both theoretical and practical impediments in the way of predicting exactly what a BCI system will do in its intended operational environment. These epistemological issues are addressed in Sect. 13.2.

Epistemological reflections on the expected behavior of BCI systems bear on a variety of issues in applied ethics. Retrospective responsibility and liability

G. Tamburrini (✉)
Dipartimento di Ingegneria Elettrica e delle Tecnologie dell'Informazione, University of Naples, Naples, Italy
e-mail: tamburrini@unina.it

problems are significant cases in point. The epistemic predicament affecting pro-gramming engineers, manufacturers, and users of BCI systems alike looms large on such questions as "Who is morally responsible for unpredicted damage caused by a BCI-actuated device?" and "How are liabilities and compensations for that damage properly distributed?" The connections between epistemology and ethics in BCI contexts are addressed in Sects. 13.3 and 13.4.

BCI research is paving the way for unprecedented forms of human–machine interaction. Notably, BCI systems enable one to establish communication channels with the external world *without requiring any voluntary muscular movement* on the part of their users. However, current limitations of BCI communication channels in the way of bit-rate transfer suggest that only special categories of healthy users may benefit in the near future from BCI systems. It has been proposed, for example, that astronauts might be one of these categories, insofar as mental teleoperation enables one to govern external semi-automatic manipulators in hampering microgravity work conditions (Millán et al. 2006).

In the light of current BCI strengths and weaknesses, the wider groups of users who are likely to take advantage of BCI systems in the near future are formed by people affected by severe motor impairments. Notably, BCI systems may mitigate excruciating communication and action barriers experienced by persons affected by Locked-In Syndrome (LIS) and other clinical conditions bringing about very limited or no residual capabilities for voluntary movement. By the same token, however, BCI experimentation and trials may induce unrealistic expectations in these remarkably vulnerable groups of persons. The ethical implications of this state of affairs are examined in Sect. 13.5 in the context of the ethical principles of respect for persons and medical beneficence.

Groundbreaking experimental trials showed that some persons diagnosed to be in a vegetative state were capable of answering correctly yes/no autobiographical questions by producing the required brain activity patterns (Owen et al. 2006; Monti et al. 2010). Similar brain activity patterns can be used to operate a BCI system. Accordingly, it was suggested that BCI systems may enable one to test whether conscious states occur in behaviorally unresponsive persons and to open, in the case of positive test outcomes, a unique communication channel with them. These envisaged uses of BCI systems give a new twist to the epistemological problem of assessing the soundness of inferences from behavioral observations to the attribution of conscious states to other human beings (or, more generally, to other entities). Here, inferences to consciousness are based on empirical premises involving BCI-detected patterns of brain activity rather than overt verbal behaviors or bodily movements (Owen et al. 2009). It transpires that basic distinctions drawn in the philosophy of mind about different aspects or varieties of consciousness help one to clarify empirical and conceptual problems arising in connection with these inferences, the correctness of their conclusions, and their ethical implications (Tamburrini and Mattia 2011). These various issues are examined in Sect. 13.6.

The new forms of human–machine interactions that are enabled by BCI systems give rise to special expectations about ICT technologies and their envisaged benefits for human life. In making public statements about their goals and

achievements, scientists cannot count on the shared background of tacit knowledge which shapes communication styles within their own scientific communities. Accordingly, the ethically praiseworthy goal of transferring correct and accessible information to the general public raises the formidable challenge of adapting communication styles to the needs of much broader audiences of non-specialists. This issue is examined in Sect. 13.7 in connection with different sorts of psychological attitudes, both rational and irrational, that BCI systems may give rise to.

On the whole, these various observations suggest that one may expect substantial rewards from an examination of BCI systems from the distinctively philosophical perspectives of epistemology and scientific method, ethics, philosophy of mind, and philosophical psychology. It is to this task that we now turn, starting from the vantage point offered by epistemology and scientific methodology.

13.2 Epistemological Reflections on Adaptive BCI Systems

Interaction between a BCI system and its user starts from the user performing a *mental task*. The user's brain activity is *recorded* during task performance, and processed online in order to *recognize* the presence of features that one *translates* into control signals for an external device. Finally, the user obtains (usually perceptual or linguistic) *feedback* about the outcomes of this repeatable interaction cycle.

This closed-loop functional scheme can be multiply realized in ways that are conditional on available kinds of BCI hardware and software; on the invasive or non-invasive character of brain signal recording devices; on the use of synchronous or asynchronous interaction protocols; and on the variety of brain signals that are acquired and processed. Additional taxonomies of BCI systems arise on the basis of a variety of informative distinctions, which include those between dependent and independent, active and passive, and hybrid and non-hybrid BCI systems (Wolpaw et al. 2002; Wolpaw and Wolpaw 2012).

In synchronous communication protocols, the mental effort required of the human user is locked to the presentation of some perceptual stimuli. For example, letters of the alphabet are consecutively flashed on a screen, and the user must concentrate on the letter he wants to select and write by means of a BCI-actuated word processor. In asynchronous communication protocols, users voluntarily initiate and pace their mental efforts. For example, one may choose to imagine some specific body movement or to carry out an arithmetic operation or another mental task taken from a fixed list in which each task is associated with the performance of an action. Imagining a specific body movement (as opposed to executing an arithmetic operation) is meant to transmit, through the BCI identification of some of its characteristic neural correlates, the command that action A (as opposed to action B) must be planned and carried out by an ICT or robotic device.

A BCI system relies on some suitable classification rules in order to distinguish between the neural correlates of different mental states. This classification rule is

usually generated by means of a supervised machine learning process (Müller et al. 2008), which involves the use of a training set formed by examples of recorded brain activity and their correct classification.

How reliable are learned classification rules in their identification of BCI user intents? This epistemological problem is addressed by extensive testing of learned rule performance or theoretical assessments of learning outcome reliability. Both approaches involve distinctive background assumptions about the significance of training data and the stability of the stochastic phenomenon one is dealing with. Indeed, reliability estimates obtained by testing the performance of learned rules are contingent on the assumption that both training and testing data are representative of a stochastically stable phenomenon. Similarly, probabilistic bounds on error frequency that one establishes within the more abstract mathematical framework of statistical learning theory (Vapnik 2000) are contingent on the background assumption that training inputs be independently drawn from a fixed probability distribution.

The background assumptions that are involved in both empirical testing and theoretical assessments of learned rule reliability are difficult to buttress when the classification of neural correlates of cognitive processing is at stake. Indeed, a variety of contextual factors jeopardize the stability of recorded brain signals. These factors notably include increased familiarity with the mental task, mental fatigue, variable attention level, and the reproducibility of initial conditions by means of technical set-up procedures that are carried out at the beginning of each BCI session. Briefly, both empirical and theoretical evaluations of the reliability of learned classification rules are based on the fulfillment of boundary conditions on brain processing and signal recording that are difficult to isolate, control, and reproduce (Tamburrini 2009). In turn, this epistemic predicament gives rise to ethical tensions between autonomy promotion and protection in practical applications of BCI systems, and especially so in connection with groups of users affected by severe motor impairments.

13.3   Protecting BCI User Autonomy

There are persons who can hardly turn *any* intention to act into appropriate sequences of bodily movements. This condition is dramatically witnessed by people affected by LIS, who preserve a basically intact mind trapped in an almost or completely paralyzed body – a butterfly in a diving bell, to use the words with which Jean-Dominique Bauby graphically conveyed his dramatic condition (Bauby 1997). In this condition of generalized motor impairment, human agency *as such* is mostly or completely compromised.

There is a direct conceptual connection between the protection of human agency as such and the protection of human dignity. The concept of agent – that is, the concept of an entity that is capable of performing a repertoire of actions guided by desires, intentions, and beliefs about the world – plays a central role in both Kantian

and neo-Aristotelian accounts of human dignity. According to Kantian approaches, the intrinsic worth of human beings is ultimately grounded in their capability to act as *homo noumenon*, by endorsing rationally and conforming their behaviors to the maxims of practical reason (Rothaar 2010). And according to neo-Aristotelian accounts based on actualizations of human capabilities, the exercise of practical deliberation, control over one's own environment, and engagement in social activities is conducive to realizing a dignified human life (Nussbaum 2006, 77–78 and 161). Clearly, both conceptions of human dignity afford ethical motivations for protecting a generalized capability to perform actions, whose vulnerability is dramatically underscored by the sweeping action impairments occurring in LIS and other pervasive motor disabilities.

By learning how to use BCI technologies, people affected by LIS once again acquire a *general-purpose* capability to act. Indeed, BCI-actuated devices presently include robotic manipulators, virtual computer keyboards, robotic wheelchairs, internet surfing systems, photo browsing, and virtual drawing and painting systems (Millán et al. 2011; Wolpaw and Wolpaw 2012). Thus, by restoring generalized capabilities for action, BCI technologies for functional substitution are instrumental to human dignity protection through the intermediary of human agency protection.

BCI communication and control protocols require that human users surrender to a computational system both user intent identification and low-level control of action. Therefore, the protection of disabled user autonomy *by means of* BCI systems requires that users rely on a machine for their intent identification and fulfilment. In particular, this sort of intent identification is needed to manifest concretely one's own legal capability by, say, engaging in e-banking transactions on the internet, expressing informed consent, and writing a testament. One should be careful to note, however, that BCI-recovered autonomy engenders the problem of protecting user autonomy *from* errors affecting the behaviors of BCI systems. For example, one may appeal to the sources of BCI misclassifications examined above in order to question the identity between the action intended by the BCI user and the action actually performed by a BCI-actuated device. Therefore, the protection of BCI user autonomy provides ethical motivation for BCI research to develop suitable intent-corroboration procedures, especially in the case of legally binding contracts and transactions.

Additional issues concerning the protection of autonomous action *from* behavioral errors of BCI systems arise in connection with shared control of action in BCI-actuated robotic systems (Tamburrini 2009; Santoro et al. 2008). The limited capacity of BCI communication channels confines direct BCI user control to high-level commands only. As a consequence, BCI-actuated robots (such as robotic wheelchairs and robotic arms for grasping and manipulating) are autonomous in their control of both low-level actions and ancillary tasks such as the avoidance of unforeseen obstacles (Millan et al. 2004; Galán et al. 2008). Autonomous robotic action may give rise to discrepancies between user intent and actual trajectories of robotic systems, in view of perceptual and planning errors, sensitivity to small perturbations of initial conditions, and sensor noise piling up in series of sensory readings (Nehmzow 2006).

13.4   Ascribing Responsibilities and Liabilities

A BCI-actuated robotic wheelchair may roll down a staircase on account of user intent misinterpretation or inaccurate sensory readings. And a robotic arm responding to a request to fetch a glass of water may fail to perceive and bump into another person standing in the room. How are responsibilities and liabilities sensibly distributed in view of the fact that programmers, manufacturers, and users are not in the position to predict exactly and certify what BCI-actuated robots will actually do in their intended operational environments? (Clausen 2008, 2009; Tamburrini 2009; Grübler 2011; Holm and Teck 2011; see also Chap. 14) Clearly, those who failed to predict damaging events arising from BCI operation cannot be held morally blameworthy, provided that they properly attended in their different capacities to every reasonable design, implementation, testing, and operational issue. Nevertheless, even in the absence of moral responsibilities deriving from negligence or malevolent intentions, caused damage and corresponding compensation claims call for a proper ascription of liabilities (also known as objective responsibilities).

In developing liability policies for BCI-actuated robots, one may note that some predictive failures arise there from learning, reasoning, and action planning capabilities that BCI systems share with human beings and other biological systems. In the light of this positive analogy between BCI systems and biological systems, one may suggest an extension to BCI-actuated robots of liability ascription criteria one adopts for damages deriving from unpredictable behaviors of biological systems. This suggested extension puts the inability of BCI users to predict exactly and control the behavior of brain-actuated robots on a par with the inability of dog owners to curb their pets in every possible circumstance; with the inability of employers to predict exactly and control the behavior of their employees; and even with the inability of parents to predict and control the behavior of their children. Parents are held to be vicariously liable for many kinds of damage caused by their children, just as pet owners are liable for damage caused by their pets, and employers are liable for certain types of damage caused by their employees. Judging by this yardstick, users should be held liable for damaging events resulting from hardly predictable behaviors of their BCI systems.

The suggestion of holding users liable for BCI-engendered damage is vulnerable to ethically motivated criticism, insofar as this criterion may lead to discrimination in assistive technology access between those who can and those who cannot afford insurance and compensation costs. As an alternative, one might shift the burden of economic compensation onto BCI manufacturers. Indeed, in view of their expected profits, producers of goods are often held liable for damaging events that are difficult to predict and control. This liability ascription policy is aptly summarized in the Roman juridical tradition by the formula *ubi commoda ibi incommoda*.

The suggestion of ascribing liability to BCI producers is exposed to ethically motivated criticism too. Indeed, the risk of high compensation costs may discourage investments in research and development (R&D) towards marketable BCI

systems, with the effect of diverting resources that are badly needed to launch a pioneering BCI industry. As a consequence, tensions are likely to arise between liability policies transferring compensation costs to BCI producers, the demands of beneficence in bioethics, and consequentialist evaluations of broader societal benefits that are expected to flow from BCI technological innovation.

These various observations suggest the opportunity of developing a more complex governance framework for BCI-engendered retrospective liabilities. Since BCI technological risk comes with beneficial opportunities for groups of disabled people and broader societal benefits in the way of technological innovation, one might allow for the socialization of risks associated with BCI systems, distributing insurance and compensation costs across a variety of stakeholder groups and governmental agencies.

13.5 Informed Consent and Respect for Persons

Along with healthy participants, research trials on the BCI-enabled replacement of motor functions may enroll people who have lost most of their abilities to communicate and act. These disabled participants may come to view a BCI research trial as a unique and last resort to overcome their communication and action barriers. These expectations might be so compelling in the dramatic human condition of severely paralyzed persons that they prevent a proper appreciation of the facts that one must know before participating in BCI experimentation (Haselager et al. 2009; Vlek et al. 2012). Accordingly, specific information aimed at anticipating and mitigating similar psychological attitudes must be included in setting up informed consent questionnaires and protocols for BCI experimentation. In particular, one should properly emphasize and carefully illustrate the phenomenon of BCI illiteracy, that is, the incapability to operate a BCI, which is estimated to affect 15–30 % of potential users, and for which effective remedies are still to be found (Vidaurre and Blankertz 2010). In addition to this, one has to provide more standard information about the likely absence of personal advantages flowing from participation in research trials, about psychological risks of depression which may derive from retracting BCI use at the end of time-limited studies, and about foreseeable discomfort in the care of disabled participants which may derive from prolonged operation and maintenance of BCI systems (Schneider et al. 2012).

It was pointed out above that machine-to-human adaptations deriving from computational learning are sources of distinctive risk and ethically motivated concern. Potentially deleterious changes in the CNS resulting from *human-to-machine adaptations* are even more important sources of risk and ethical concern about the physical and mental integrity of persons.

Psychological rewards and punishments deriving from the observation of BCI interaction outcomes are known to affect CNS activity patterns. As a matter of fact, an operant conditioning process usually intervenes to change brain activity if the user obtains negative feedback information, that is, information about occurring

discrepancies between expected and actual outcomes of his interaction with the machine. As a result of operant conditioning, brain activity patterns change in ways that have been found to facilitate ensuing machine classifications. Accordingly, as the overall functional implications of these adaptations are not fully understood and difficult to predict, one cannot rule out generic risks of detrimental effects on states of mind and behaviors of intensive BCI users (Schneider et al. 2012). Thus, every BCI candidate user must be properly informed of potentially detrimental effects of BCI-induced brain adaptations and plasticity (Dobkin 2007).

Unlike BCI systems for communication and control, some experimental BCI rehabilitation therapies for improving motor functions *target directly* brain areas, with the principal aim of modifying their structure and functions (Shih et al. 2012). One strategy for post-stroke motor rehabilitation involves a BCI system monitoring damaged brain areas which normally control movements that are impaired after brain injury. The BCI system provides appropriate feedback according to whether activation patterns in the targeted brain areas come closer to normal or not (Grosse-Wentrup et al. 2011; Pichiorri et al. 2011; Daly and Sitaram 2012; Mattia et al. 2012; Várkuti et al. 2013). Similarly, BMCI systems (with the letter "M" in the acronym standing for *muscle*) combine the analysis of brain signals and electromyographic signals to stimulate increasingly correct activity patterns in both brain and muscles (Bermudez et al. 2013). Ethical concerns about BCI systems directly fostering brain plasticity arise, at a general level, from awareness of limited etiological understanding of deleterious effects, if any, of these experimental therapies.

Additional ethical issues arise in connection with the use of *invasive* versus *non-invasive* BCI systems. Operation of invasive BCI systems requires the implant of electrodes in the cortex and the deployment of apparatus for electrocorticography (ECoG) or intracranial electroencephalography (iEEG) on the exposed brain surface (Moran 2010). In general, invasive BCIs enable one to achieve, in contrast to non-invasive ones, better signal spatial resolution and signal-to-noise ratio leading to improved control of peripheral devices. Among the non-invasive systems, those relying on electroencephalography (EEG) as a recording method afford better performances in terms of signal temporal resolution, cost, and practicality of use. Potential users of BCI systems must be informed of comparative advantages and disadvantages of invasive and non-invasive systems, including relevant facts about characteristic risks of invasive systems in connection with implant stability, reversibility, and infection. Interestingly, empirical data from interviews administered to people affected by Amyotrophic Lateral Sclerosis (ALS) suggest a definite preference for non-invasive systems, notwithstanding the functional advantages of invasive BCI systems which derive from better spatial resolution and signal-to-noise ratios, and corresponding disadvantages of non-invasive systems in the way of slower operation and more error-prone control (Birbaumer 2006a, b).

13.6   The Protection of Agentivity and Consciousness

Some explanations of the inability to learn how to operate BCI systems suggest that teaching people affected by LIS and other severe motor disabilities how to use a BCI system before they completely lose muscle control may contribute to protecting their agentivity and possibly to preventing their purposeful thinking from waning. Experimental studies show that among patients affected by Amyotrophic Lateral Sclerosis (ALS) and trained with a non-invasive BCI, none of those who were trained after entering a complete locked-in state (CLIS) were able to acquire stable communication abilities (2006a). Birbaumer (2006b) advanced two competing explanations for this observation. According to the first explanation, the onset of CLIS is accompanied by a generalized decline in perception, thinking, and attention abilities. It is this decline which prevents CLIS patients from learning to use a BCI, and therefore learning how to use a BCI cannot counteract this progressive and generalized decline.

Birbaumer's second explanation hinges on the hypothesis that the development and sustained preservation of purposive thinking crucially involves a reinforcement stage, concerning the verification of intended consequences of actions. This hypothesis is advanced in the framework of so-called motor theories of thinking, according to which thinking develops as a means for – and is sustained by – effective animal motion: "As early as the 19th century, the 'motor theory of thinking' hypothesized that thinking and imagery cannot be sustained if the contingency between an intention and its external consequence is completely interrupted for a long time period" (Birbaumer 2006b, 481). The reinforcement stage required by motor theories of thinking is hardly ever accessed in a CLIS subject. The sequence intention-action-consequence-verification cannot be enacted autonomously; it is occasionally completed through the intermediary of caretakers who happen to fulfill the patient's current desire. Therefore, thinking and imagery are no longer sustained, and the related ability to learn and operate a BCI fades away in a CLIS patient.

A third explanation of the inability of disabled people with no remaining muscle control to learn and operate BCI systems hinges on the theory of learned helplessness (Seligman 1975). A human may learn to behave helplessly in the face of adverse conditions that he cannot modify, and will fail to alter this response even though a new opportunity subsequently arises to help oneself and obtain positive rewards.

If the second explanation is correct, then learning how to use a BCI before the onset of CLIS may prevent the extinction of thinking and imagery, insofar as the sequence intention-action-consequence-verification is preserved through BCI operation. And if the third explanation is correct, then learning how to use a BCI before the onset of CLIS may prevent the insurgence of the condition of learned helplessness as an overwhelming obstacle towards BCI use at a later time. In either case, one should teach BCI operation to persons who may subsequently enter the CLIS state so as to preserve their status of agents. Thus, medical beneficence provides

ethical motivations for probing the effectiveness of these therapeutic interventions and for testing their theoretical premises in motor thinking and learned helplessness hypotheses.

Each one of the above applications of BCI systems presupposes the possession of a wide variety of mental capabilities on the part of their prospective users. These mental requirements are usually satisfied by people affected by LIS. In contrast to this, it is far from obvious that persons affected by disorders of consciousness possess the mental capabilities that are needed to operate a BCI. However, ground-breaking experiments (Owen et al. 2006; Monti et al. 2010) involving groups of persons who were diagnosed to be in a vegetative state (VS) or a minimally conscious state (MCS) unexpectedly suggested the possibility of using BCIs to communicate with people affected by disorders of consciousness.

In these experiments, VS and MCS patients were verbally instructed to perform a motor imagery task (playing tennis) or a spatial imagery task (visiting rooms in their home) in order to convey their "yes" or "no" answer, respectively, to questions posed by experimenters. From an analysis of functional magnetic resonance imaging (fMRI) scans taken upon administering autobiographical questions (e.g., "Is your father's name Thomas?" and "Do you have any sisters?"), regional brain activations were reliably and repeatedly found to correspond to correct imagery-conveyed answers in a small proportion of those patients.

Several extensions of this communication protocol have been proposed and discussed. These extensions concern a variety of dialogical purposes, ranging from subjective symptom reporting to informed consent, and even to continued medical care decision-making. It was suggested, for example, that "patients could be asked if they are feeling any pain, and this information could be useful in determining whether analgesic agents should be administered" (Monti et al. 2010). Moreover, it was claimed that "the first and obvious use of mental signaling by means of fMRI could be to preserve the patient's autonomy by querying his or her wishes regarding continued medical care" (Ropper 2010). Understanding which aspects of consciousness must be present for these interactions to take place, and how to detect them in behaviorally unresponsive patients who satisfy the medical criteria for being diagnosed as VS or MCS is a formidable scientific problem (Nachev and Husain 2007). A limited contribution that philosophy can provide towards the conceptual clarification of this problem is based on distinctions between aspects and varieties of consciousness that are routinely made in the framework of contemporary philosophy of mind. In particular, *phenomenal*, *access*, and *narrative* forms of consciousness appear to be selectively involved as preconditions for genuine communication about pain reporting, informed consent, and continued medical care (Tamburrini and Mattia 2011; Tamburrini 2013).

Roughly speaking, *access consciousness* is identified with the ability to introspect one's own mental states and to make the introspected mental states available for modification by reflection (Block 1995). The *introspective* and *reflective components* of access consciousness do not necessarily come together in conscious mental life, insofar as one may introspect, say, one's own desire without being able

to modify it by reflection: irresistible desires and incorrigible perceptual illusions that are introspectively accessible are significant cases in point.

The subjective feel or experiential dimension of one's own mental states is not captured by the notion of access consciousness. The capability to experience what it is like to be in a certain mental state is identified with another variety of consciousness. This is *phenomenal consciousness* (Nagel 1974), whose manifestations comprise perceptual experiences (like tasting something) and bodily experiences (like pain).

Let us finally observe that mental states can be recalled and unified as a narrative of episodes, which are experienced and accessed from the unitary perspective of the individual self by taking into account their causal and semantic connections. This ability to organize, access, and experience from a unitary subjective perspective a series of mental states is often referred to as *narrative consciousness* (Ricoeur 1990; Merkel et al. 2007).

Let us now bring these distinctions to bear on envisaged BCI communication with persons affected by disorders of consciousness. According to the definition advanced by the International Association for the Study of Pain (IASP), the term 'pain' denotes "an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage." Thus, being in pain involves having an *experience* of the qualitative feel of pain or, in other words, to possess phenomenal consciousness in connection with the bodily experiences that one calls pain. Accordingly, in order to deal competently with pain questions one must be endowed with some form of phenomenal consciousness.

Let us now turn to consider informed consent about medical treatment or participation in research trials. This kind of communication is usually motivated on autonomy protection grounds, and the individual autonomy one is presupposing there involves the capability to act on one's own desires. Since, however, agents who are driven by a strong desire or impulse to act fail to qualify as autonomous in the etymologically grounded sense of being *self-ruled agents*, individual autonomy additionally requires the ability to endorse, modify, or reject one's own desires to act. Briefly, autonomous decision-making involves full access consciousness ascriptions, insofar as an autonomous agent must be able to introspect and perform reflective interventions of various sorts on his own mental states.

Even more comprehensive consciousness ascriptions appear to be presupposed in communication concerning continued medical care. An expressed preference to discontinue medical care must be pondered over, remain open to revision for some time ahead, and be typically reinforced in iterated interviews. Accordingly, one is presupposing that the recipient of continued medical care questions is endowed with memory continuity of the sort that allows one to recall previously expressed preferences, if any, their motivations, and their evolving patterns from one interview to the next. Briefly, by raising questions about continued medical care, one is pragmatically assuming that their recipient is capable of putting together a narrative of reflective and deliberative episodes from a unitary first-person perspective. These abilities are typically associated with the *narrative* variety of consciousness.

The empirically elusive and formidable problem of detecting whether a person diagnosed to be in a VS or MCS still possesses the required varieties of consciousness calls for a cautious attitude in the process of evaluating the purported significance, if any, of the results in Monti et al. (2010) for BCI dialogical communication about pain reporting, informed consent, and continued medical care. And a cautious attitude is similarly required to communicate BCI research to stakeholders and to deal sensibly with both rational and irrational expectations that BCI systems may give rise to.

13.7 Communicating BCI Research Programs and Achievements

BCI systems prompt a special sense of wonderment which is related to the fact that no speech, no gesture, and no voluntary muscular movement are required to actuate user intents.

> I have mixed attitudes towards technology. I use it and take it for granted. I enjoy it and occasionally am frustrated by it. And I am vaguely suspicious of what it is doing to our lives. But I am also caught up by a wonderment at technology, a wonderment about what we humans have created. Recently researchers at the University of Pittsburgh developed technology that allows a monkey with tiny electrodes implanted in its brain to control a mechanical arm. The monkey does this not by twitching or blinking or making a slight movement, but by using its thoughts alone. The technology has obvious promise for impaired people. But that is not what causes me wonder. I wonder that we can put together circuits and mechanical linkages – in the end pieces of silicon and copper wire, strips of metal, and small gears – so that machinery responds to thought and to thought alone (Brian 2009, 9).

Along with reactions of wonderment, psychological responses to BCI systems present an intriguing combination of fantasies, worries, and rational and irrational expectations. Indeed, the openings of several popular science and media reports leverage on the "magic" allure of BCI technologies and their alleged ability to respond to the force of thought only. A reflection on these various mental attitudes is crucial to appreciate the symbolic roles of novel technologies in contemporary society, and their influence on public debate and decision-making about technological research, development, and dissemination. In particular, an understanding of generative mechanisms underlying unrealistic expectations about BCI systems is crucial to develop effective BCI research communication strategies, and to build mutual trust between scientists, users, and other groups of stakeholders. Let us consider, from this perspective, psychoanalytic mechanisms that may contribute to explaining the "magic" appeal of BCI systems (Scalzone and Tamburrini 2012).

In his *Thoughts for the Times on War and Death*, Sigmund Freud remarked: "It is an inevitable result of all this that we should seek in the world of fiction, in literature

and in the theatre compensation for what has been lost in life [. . .] For it is really too sad that in life it should be as it is in chess, where one false move may force us to resign the game, but with the difference that we can start no second game, no return-match. In the realm of fiction we find the plurality of lives which we need." (Freud 1964, vol. 14, 291).

Psychologically compensating scenarios once explored in literary work only are now in the purview of technological research programs promising substantive extensions and enhancements of human capabilities. These scenarios impinge on the general public through their dissemination in the media and popular science reports. But what is the psychological basis of compensation-seeking attitudes towards technology? Freud's theory of narcissism makes an explanatory basis available for understanding the compensating psychological role of technologies in general, and BCI presentations which strike the chord of the "force of thought" in particular. According to Freud, children in a normal stage of their development entertain primitive beliefs characterizing animistic conceptions of the world. These animistic beliefs, notably concerning magic wish-fulfilling and the omnipotence of one's own thoughts, lead one to overestimate the capability to bend external reality to one's own desires by the force of thought only. Adults give up these narcissistic beliefs in their conscious life, coming to terms with the reality principle and acknowledging death as inevitable. But these repressed beliefs persist and operate unconsciously in adult life. In particular, one unconsciously seeks compensations for the conscious acceptance of the reality principle and the attending psychological blows for naïve self-love. Thus, in particular, descriptions of BCI technologies may provide some such compensation – an opportunity for what Freud called a fictitious "return match" or "second game" – taking the form of illusory enhanced control of the external world by magic wish-fulfilling.

This interpretation of psychological responses towards BCI systems suggests that public statements of researchers about their research goals and activities may inadvertently become a powerful source of irrational attitudes towards techno-logical progress. These irrational attitudes are more likely to emerge as a response to popular expositions emphasizing the promises of scientific research programs, but neglecting to emphasize information which is needed to gauge the distance between promises and the actual results of research activity. The ambitious long-term goals of research programs play significant motivating roles within communities of scientists and may suggest fruitful lines of inquiry, even though their promises cannot be attained without making substantial and unforeseeable progress with respect to currently available models and techniques. However, the dividing line between concrete achievements and long-term or visionary goals of a research program is not invariably clear to the non-specialist. Therefore, the ethically praiseworthy goal of furnishing correct and accessible information to the general public raises the formidable challenge of adapting the communication styles that scientists use within their communities so as to anticipate and avoid the insurgence in the general public of disproportionate and irrational expectations towards tech-nological development.

Ethical motivations for carefully drawing the distinction between envisaged long-term goals of a research program and its tangible results have emerged in connection with several developments of BCI research. Thus, in communicating the research goal of making a variety of BCI technologies and systems available to people affected by LIS, one ought to emphasize properly problems of BCI system reliability, cognitive decline in LIS patients, the incidence of BCI illiteracy, and the generalizability to clinical contexts of results obtained in research trials involving healthy subjects only. Moreover, one must specify BCI costs and benefits with respect to alternative communication methods for persons who retain some communication capabilities, say by moving their eyes or eye-lids.

The need for responsible communication strategies emerges even more evidently in connection with the suggestion of using BCI for communicating with and promoting the autonomy of people affected by disorders of consciousness. Assessing whether persons affected by disorders of consciousness still possess the required varieties and degrees of consciousness is a scientifically formidable and empirically elusive problem. Accordingly, one ought to adopt an extremely cautious attitude in communicating to psychologically vulnerable families of people affected by disorders of consciousness the aims of these studies and their envisaged implications in therapy and quality of life improvement. There, the development of an effective and responsible communication strategy may take advantage of the above philosophical analyses of consciousness with their clarifying distinctions that are closer to common sense conceptualizations of the phenomenon of consciousness.

References

Bauby, J.-D. 1997. *Le scaphandre et le papillon*. Paris: Éditions Robert Laffont.

Bermudez, I., S. Badia, A. Garcia Morgade, H. Samaha, and P. Verschure. 2013. Using a hybrid brain computer interface and virtual reality system to monitor and promote cortical reorganization through motor activity and motor imagery training. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 21(2): 174–181.

Birbaumer, N. 2006a. Breaking the silence: Brain-computer interfaces for communication and motor control. *Psychophysiology* 43: 517–532.

Birbaumer, N. 2006b. Brain-computer interface research: Coming of age. *Clinical Neurophysiology* 117: 479–483.

Block, N. 1995. On a confusion about the function of consciousness. *Behavioral and Brain Sciences* 18: 227–247.

Brian, A.W. 2009. *The nature of technology*. New York: Free Press.

Clausen, J. 2008. Moving minds: Ethical aspects of neural motor prostheses. *Biotechnology Journal* 3(12): 1493–1501. doi:10.1002/biot.200800244.

Clausen, J. 2009. Man, machine, and in between. *Nature* 457(7233): 1080–1081.

Daly, J.J., and R. Sitaram. 2012. BCI therapeutic applications for improving brain function. In *Brain-computer interfaces: Principles and practice*, ed. J.R. Wolpaw and E.W. Wolpaw, 351–362. Oxford: Oxford University Press.

Dobkin, B.H. 2007. Brain-computer interface technology as a tool to augment plasticity and outcomes for neurological rehabilitation. *Journal of Physiology* 579: 637–642.

Freud, S. 1964. *The standard edition of the complete psychological works of Sigmund Freud*, ed. James Strachey. London: Hogarth Press.

Galán, F., M. Nuttin, E. Lew, P.W. Ferrez, G. Vanacker, J. Philips, and J.del.R. Millán. 2008. A brain-actuated wheelchair: Asynchronous and non-invasive brain-computer interfaces for continuous control of robots. *Clinical Neurophysiology* 119: 2159–2169.

Grosse-Wentrup, M., D. Mattia, and K. Oweiss. 2011. Using brain-computer interfaces to induce neural plasticity and restore function. *Journal of Neural Engineering* 8: 025004.

Grübler, G. 2011. Beyond the responsibility gap. Discussion note on responsibility and liability in the use of brain-computer interfaces. *AI and Society* 26: 377–382.

Haselager, P., R. Vlek, J. Hill, and F. Nijboer. 2009. A note on ethical aspects of BCI. *Neural Networks* 22: 1352–1357.

Holm, S., and C.V. Teck. 2011. Brain-machine interfaces and personal responsibility for action – Maybe not as complicated after all. *Studies in Ethics, Law, and Technology* 4(3), Article 7.

IASP. 1986. Pain terms: A list with definitions and notes on pain. *Pain* 3(Suppl): 216–221.

Mattia, D., F. Pichiorri, M. Molinari, and R. Rupp. 2012. Brain computer interface for hand motor function restoration and rehabilitation. In *Towards practical brain-computer interfaces*, ed. - B.Z. Allison et al., 131–153. Berlin: Springer.

Merkel, R., G. Boer, J. Fegert, T. Galert, D. Hartmann, B. Nuttin, and S. Rosahl. 2007. *Intervening in the brain. Changing psyche and society*. Berlin: Springer.

Millán, J.del.R., F. Renkens, J. Mouriño, and W. Gerstner. 2004. Brain-actuated interaction. *Artificial Intelligence* 159: 241–259.

Millán, J.del.R., P.W. Ferrez, and A. Buttfield. 2006. Non-invasive brain-computer interfaces. *ESA, European Space Agency Report*: ID 05/6402.

Millán, J.del.R., R. Rupp, G.R. Müller-Putz, R. Murray-Smith, C. Giugliemma, M. Tangermann, C. Vidaurre, F. Cincotti, A. Kübler, R. Leeb, C. Neuper, K.-R. Müller, and D. Mattia. 2011. Combining BCI and assistive technologies: State-of-art and challenges. *Frontiers in Neuroscience* 4: 161–168.

Monti, M.M., A. Vanhaudenhuyse, M.R. Coleman, M. Boly, J.D. Pickard, L. Tshibanda, A.M. Owen, and S. Laureys. 2010. Wilful modulation of brain activity in disorders of consciousness. *New England Journal of Medicine* 362: 579–589.

Moran, D. 2010. Evolution of brain-computer interface: Action potentials, local field potentials and electrocorticograms. *Current Opinion in Neurobiology* 20: 741–745.

Müller, K.R., M. Tangermann, G. Dornhege, M. Krauledat, G. Curio, and B. Blankertz. 2008. Machine learning for real-time single-trial EEG-analysis: From brain-computer interfacing to mental state monitoring. *Journal of Neuroscience Methods* 167: 82–90.

Nachev, P., and M. Husain. 2007. Comment on "Detecting awareness in the vegetative state". *Science* 315: 1221.

Nagel, Th. 1974. What is it like to be a bat? *Philosophical Review* 83: 435–456.

Nehmzow, U. 2006. *Scientific methods in mobile robotics*. Berlin: Springer.

Nussbaum, M.C. 2006. *Frontiers of justice. Disability, nationality, species membership*. Cambridge: Harvard University Press.

Owen, A.M., M.R. Coleman, M. Boly, M.H. Davis, S. Laureys, and J.D. Pickard. 2006. Detecting awareness in the vegetative state. *Science* 313: 1402.

Owen, A.M., N.D. Schiff, and S. Laureys. 2009. The assessment of conscious awareness in the vegetative state. In *The neurology of consciousness*, ed. S. Laureys and G. Tononi, 163–172. Amsterdam: Elsevier.

Pichiorri, F., F. De Vico Fallani, F. Cincotti, F. Babiloni, M. Molinari, S.C. Kleih, C. Neuper, A. Kübler, and D. Mattia. 2011. Sensorimotor rhythm-based brain-computer interface training: The impact on motor cortical responsiveness. *Journal of Neural Engineering* 8: 025020. doi:10.1088/1741-2560/8/2/025020.

Ricoeur, P. 1990. *Soi-même comme un autre*. Paris: Edition du Seuil.

Ropper, A.H. 2010. Cogito ergo sum by MRI. *New England Journal of Medicine* 362: 648–649.

Rothaar, M. 2010. Human dignity and human rights in bioethics: The Kantian approach. *Medicine, Health Care, and Philosophy* 13: 251–257.

Santoro, M., D. Marino, and G. Tamburrini. 2008. Robots interacting with humans. From epistemic risk to responsibility. *AI and Society* 22: 301–314.

Scalzone, F., and G. Tamburrini. 2012. Human-robot interaction and psychoanalysis. *AI and Society* 27: 297–307. doi:10.1007/s00146-012-0413-3.

Schneider, M.-J., J.J. Fins, and J.R. Wolpaw. 2012. Ethical issues in BCI research. In *Brain-computer interfaces: Principles and practice*, ed. J.R. Wolpaw and E.W. Wolpaw, 373–383. Oxford: Oxford University Press.

Seligman, M.E.P. 1975. *Helplessness: On depression, development, and death*. San Francisco: W. H. Freeman.

Shih, J.J., D.J. Krusienski, and J.R. Wolpaw. 2012. Brain-computer interfaces in medicine. *Mayo Clinic Proceedings* 87: 268–279.

Tamburrini, G. 2009. Brain to computer communication: Ethical perspectives on interaction models. *Neuroethics* 2: 137–149.

Tamburrini, G. 2013. Menschenwürde und Gehirnkommunikation und -steuerung. In *Menschenwürde und Medizin. Ein interdisziplinäres Handbuch*, ed. J.C. Joerden, E. Hilgendorf, and F. Thiele, 895–904. Berlin: Duncker & Humboldt.

Tamburrini, G., and D. Mattia. 2011. Disorders of consciousness and communication: Ethical motivations and communication-enabling attributes of consciousness. *Functional Neurology* 21: 51–54.

Vapnik, V. 2000. *The nature of statistical learning theory*, 2nd ed. New York: Springer.

Várkuti, B., C. Guan, Y. Pan, K.S. Phua, K.K. Ang, C.W. Kuah, K. Chua, B.T. Ang, N. Birbaumer, and R. Sitaram. 2013. Resting state changes in functional connectivity correlate with movement recovery for BCI and robot-assisted upper-extremity training after stroke. *Neurorehabilitation and Neural Repair* 27: 53–62.

Vidaurre, C., and B. Blankertz. 2010. Towards a cure for BCI illiteracy. *Brain Topography* 23: 194–198.

Vlek, R.J., D. Steines, D. Szibbo, A. Kübler, M.-J. Schneider, P. Haselager, and F. Nijboer. 2012. Ethical issues in brain-computer interface research, development, and dissemination. *Journal of Neurologic Physical Therapy* 36: 94–99.

Wolpaw, J.R., and E.W. Wolpaw. 2012. *Brain-computer interfaces: Principles and practice*. Oxford: Oxford University Press.

Wolpaw, J.R., N. Birbaumer, D.J. McFarland, G. Purtscheller, and T.M. Vaughan. 2002. Brain-computer interfaces for communication and control. *Clinical Neurophysiology* 113: 767–791.