Articles

Brain–Computer Interfaces

Lessons to Be Learned from the Ethics of Algorithms

ANDREAS WOLKENSTEIN, RALF J. JOX, and ORSOLYA FRIEDRICH

Abstract: Brain–computer interfaces (BCIs) are driven essentially by algorithms; however, the ethical role of such algorithms has so far been neglected in the ethical assessment of BCIs. The goal of this article is therefore twofold: First, it aims to offer insights into whether (and how) the problems related to the ethics of BCIs (e.g., responsibility) can be better grasped with the help of already existing work on the ethics of algorithms. As a second goal, the article explores what kinds of solutions are available in that body of scholarship, and how these solutions relate to some of the ethical questions around BCIs. In short, the article asks what lessons can be learned about the ethics of BCIs from looking at the ethics of algorithms. To achieve these goals, the article proceeds as follows. First, a brief introduction into the algorithmic background of BCIs is given. Second, the debate about epistemic concerns and the ethics of algorithms is sketched. Finally, this debate is transferred to the ethics of BCIs.

Keywords: brain–computer interface; ethics of algorithms; responsibility

Introduction

During the last decade, not only has the technological development around brain–computer interfaces (BCIs) advanced, but so also has the ethical and philosophical preoccupation with the implications of connecting brains to machines.[1,2,3] Sasha Burwell, Matthew Sample, and Eric Racine identify issues of user safety and risk–benefit analysis, humanity and personhood, stigma, autonomy and self-determination, responsibility, research ethics and informed consent, privacy, and security as well as justice as the main concerns that ethicists have identified so far.[4] Of particular importance is the question that some authors have debated using the term "responsibility gap" to suggest that responsibility for harm caused by BCIs might be difficult to ascribe.[5,6,7,8,9,10,11]

The potential responsibility gap arises because control over the BCI-mediated action is shared between the user and a computer. This makes it difficult to discern whether a particular action can be traced back to a human or to the machine. Existing answers to the potential responsibility gap often point to other domains of technology, in which comparable issues have emerged, whereas other scholars propose modifications to the concept of responsibility itself, in order to face the gap. However, none of the previous discussions take a sufficiently close look at the technology itself; that is, at the *use of algorithms*. Algorithms are mathematical constructs that run a program or device to solve given tasks, and that are implementable in various technologies.[12,13] Some algorithms also employ machine-learning techniques; that is, they do not perform preprogrammed decisionmaking steps, but change their decisionmaking rules autonomously.

Nonetheless, engaging with the question of algorithms is essential for obtaining fundamental answers to the responsibility question and other ethical issues, because

the computer control in BCI activities is essentially based on algorithms. We will, therefore, follow up on the observation that the algorithm-based technology behind BCIs is connected to epistemic concerns that need further analysis, in order to gain new insights into the responsibility question, and also into ethical issues that have been the focus of BCI ethicists. By "epistemic concerns," we mean concerns that relate to the way algorithms produce actionable insights or knowledge. Because algorithms relate input from the human brain to an output device (see subsequent discussion), epistemic concerns deal with the information algorithmically generated about the brain activity, its algorithmic interpretation, and transmission to the output channel.

Therefore, the goal of this article is twofold: First, it aims to seek whether (and how) the problems related to the ethics of BCIs can be better grasped with the help of the ethics of algorithms. Its second goal is to see what kinds of solutions are available in that body of scholarship, and how these solutions relate to BCIs.[14] In short, we want to inquire what lessons can be learned about the ethics of BCIs from looking at the ethics of algorithms. This field of inquiry deals with the ethical issues that arise out of the development, design, and use of algorithms for a variety of purposes. It has recently emerged as an attempt to grasp the fact that human interaction in and with the world is increasingly mediated by algorithms that pose distinct ethical questions.

To achieve these goals, we proceed as follows. First, a brief introduction into the algorithmic background of BCIs is given in the next section. Second, the debate about epistemic concerns and the ethics of algorithms is sketched in the third section. Third and finally, this debate is transferred to the ethics of BCIs in the fourth section.

Algorithms in BCIs

A BCI is essentially a system that translates brain activity into actionable output commands for a technological device. BCIs consist of four elements.[15]

1) The generation of brain signals by the user. Typically, BCIs are divided into three categories:[16] active BCIs in which the user intentionally produces certain brain states (e.g., motor imagery) that the BCIs learn to connect with the intended output; reactive BCIs in which the user is presented with certain (mostly visual or auditive) stimuli while the BCI measures a particular reactive brain signal; and passive BCIs in which the user's brain activity is monitored and action is taken as soon as a predefined state occurs.
2) The measurement of these signals, using various instruments such as electroencephalography (EEG), magnetoencephalography (MEG), or functional magnetic resonance imaging (fMRI)
3) The decoding of the measured brain signals; that is, the identification of the relevant aspects of the signal, and the classification according to the brain activity of interest
4) The output commands that direct a given external device (e.g., a wheelchair or a computer)

In the functioning of a BCI, algorithms can be found throughout several steps. *First*, the extraction and classification of brain signals makes use of (machine-learning) algorithms.[17] Only with effective algorithms can the preprocessed brain

data be classified to enable spelling for communication-disabled patients.[18,19] The challenge is, as some authors note,[20] to find algorithms that extract features in a reliable way.

Second, algorithms are again utilized to initiate the performance of certain actions autonomously, based on the data received.

Third, algorithms tend also to be used to initiate actions self-reliantly, in order to enhance BCI usage when the user is unable or unwilling to perform accurate brain-based controlling operations. Pim Haselager[21] uses the example of a BCI-directed wheelchair that autonomously drives around obstacles, and Rafael Yuste, Sara Goering, Blaise Agüera y Arcas, Guoqiang Bi, Jose M. Carmena, Adrian Carter et al.[22] reason about a possible "auto-complete" or "auto-correct' function." This use of algorithms is dependent on combining brain data with other information (e.g., about the external world and the larger context of action).

Given the centrality of algorithms in BCI use on the one hand, and the numerous ethical questions raised with regard to BCIs on the other hand, it appears natural to inquire what role algorithms play in the ethics of BCIs.

The Ethics of Algorithms: On How to Create Transparency

Brent Mittelstadt, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter and Luciano Floridi offer an overview of concerns with regard to the use of algorithms across a multitude of contexts (e.g., credit rating in banking, prediction of recidivism in judicial sentencing and predictive policing more generally).[23] According to these authors, an algorithm is used to provide evidence on which an action is taken, generically speaking. This evidence, however, can either be inconclusive, inscrutable, or misguided. In other words, algorithms suffer from epistemic problems concerning the nature of algorithmic evidence. Moreover, the use of algorithms for decisionmaking can lead to social discrimination and affects notions such as privacy and autonomy. And finally, the general issue of traceability is mentioned as a source of concerns with algorithms. For the present purpose it is useful to closely examine the impact of these epistemic concerns on the use of BCIs, because the epistemic dimension of BCIs are considered to contribute to the emergence of a responsibility gap in BCI use.

Assessing the Ethics of Algorithms

The first epistemic concern holds that the connection between the production of evidence and the resulting conclusion might be inscrutable.[24] Inscrutability or opacity can derive from *intentional hiding* (e.g., because of competition on the market), from people's *inability to technically understand* programming and algorithms, or from the *nature of algorithms* and their high dimensionality and complexity. Algorithms are often judged as black boxes because, especially in machine-learning algorithms, the working mechanisms are opaque, even to experts.

Second, the evidence generated by algorithms can be inconclusive. There is an inevitable amount of uncertainty attached to the conclusion drawn from the data an algorithm uses, because, in essence, algorithms typically detect correlations between variables, but rarely causation.[25] Imagine the case of correlating one's place of living and one's financial situations: credit rating algorithms take this correlation (which is not a causal relation) as an indicator of one's potential to

pay back. However, these kinds of correlations might be spurious, as many authors maintain.[26]

Third, epistemic worries arise because the quality of an algorithm's output depends on the data that serve as input ("misguided evidence"). If the data are misguided, so is the resulting conclusion, it seems. The term "misguided" in this context refers to biases on the part of those who design and implement these algorithms. Biases can enter an algorithm as the result of a *designer's or user's decision* (e.g., when automated surveillance incorporates the values and biases of law enforcement authorities), but they can also be *technologically driven* (e.g., when flaws in the data reoccur in the resulting conclusions), or represent an *emergent phenomenon* linked to new contexts of use (e.g., when clinical decision-support systems [CDSSs] can only recommend treatments that are within its database).

Where Are the Problems?

There a number of problems related to these issues. Opacity or lack of transparency in algorithms' data mining or decisionmaking, for example, decreases *trust* in algorithms and data processors. Because users of algorithms; such as banks, insurance companies, or courts, depend on their customers' and citizens' trust, making use of inscrutable algorithms whose decision-logic defies human comprehension has negative implications for cooperation.

Furthermore, opacity makes *oversight, comprehension, interpretation, and control* quite impossible. If an algorithm and its decisionmaking logic cannot be adequately interpreted and understood, the resulting decisions cannot be controlled, tested, and investigated. Both the concrete logic behind a decision and the wider implications of using a certain model in the algorithm potentially remain opaque and are therefore hidden from control and oversight.

Algorithmic opacity, hampering accessibility and comprehensibility, also stands in the way of *informational* privacy.[27,28] If the basic elements on which a decision against a person is based are inscrutable and defy accessibility and control, the person loses the capacity to control what is known and who knows it. Moreover, people can no longer control what is done with the information available if they does not know what information exists in the first place. This has implications for *autonomy* as well. People who do not have knowledge of, or control over. their data, namely its acquisition and processing, might be restricted, or even restrict themselves pre-emptively, in their choices, and, therefore, suffer a decrease of autonomy.

Inscrutability, but also inconclusiveness, are often related to questions of *discrimination*, because not knowing how an algorithm works makes it harder to detect or locate biases and, consequently, to find solutions to address these possible shortcomings. Finally, through algorithmic discrimination, existing inequalities might be reinforced, leading to less *autonomy* and *participation* on the part of the discriminated groups and persons. An example of the reinforcement of inequalities might be a case in which homeless people did not receive assistance because an algorithm made the assumption that someone who had been in prison previously cannot be homeless.[29]

Solutions

Several scholars have nonetheless pointed out a number of possible solutions to these problems. First, there is the relatively intuitive idea of creating *transparency*

as a means to overcome opacity and other shortcomings. However the usefulness of transparency in itself can also be questioned, with the argument that transparency is often neither necessary nor sufficient for accountability in algorithmic decisionmaking.[30] Transparency itself may not even be the most desirable goal. Additional technological instruments are suggested to provide accountability instead. *Transparency* can be created either by making explicit input–output relations such that the algorithm only produces results that are expected beforehand. This, however, would amount actually to *not using machine-learning algorithms* altogether because, almost by definition, machine learning defies the inclusion of fixed decisionmaking criteria.[31] The problem is that if algorithms are programmed in this manner, their effectiveness might be reduced.

Alternatively, algorithms can also be tested and licensed by third parties, also known as *auditing*.[32,33] The auditing approach might be useful in the case of effective machine- learning algorithms that remain opaque, sometimes even to experts.

Moreover, implementing an *explanation module* that exposes the decisionmaking of the algorithm "post-hoc" in humanly understandable terms might be another way to address opacity in machine-learning algorithms.[34,35] This solution has the problem of potentially leading to false reassurance or at least to an incomplete understanding if supposedly humanly comprehensible explanations are given. *Public education* in programming, code writing, and computational thinking in general represents a further path toward enhancing people's awareness and sensitivity to algorithms, and therefore toward creating oversight and control. The use of *open source software* might also contribute to this end.[36]

With the aim of avoiding discrimination, further *technological solutions* are discussed; for example, programming without discriminating factors.[37] Finally, other innovative approaches have been suggested to resolve the responsibility gap. These approaches tend to presuppose that not only humans can act morally or be responsible, and thus suggest a framework for programming moral competence in robots or of devising a means to implement moral reasoning in machines.[38,39,40,41]

The ethics of algorithms debate has identified a number of epistemic and resulting ethical concerns and has provided some possible solutions for these concerns. In the following, we will address the question of if and how these concerns and solutions can be used to shed light on the ethics of BCIs.

Algorithms and the Ethics of BCIs

Algorithms and Ethical Concerns Regarding BCIs

The ethics of algorithms debate shows how algorithmic opacity creates problems in BCIs as well. The evidence created by BCI algorithms can be inscrutable, because their working mechanisms are somewhat opaque. Moreover, the evidence can also be inconclusive. In the case of BCIs, the "evidence" in question is the identification of relevant brain signals (filtering out all artefacts) and their classification by the algorithm; for example, as a certain intention or affective state. Thus, algorithms create evidence for the conclusion that a certain state of mind in the user has occurred. In most cases, this evidence is directly transmitted to the output device; for example, an exoskeleton, a wheelchair, or a communication device. This kind of evidence is opaque to the extent that the particular decision-logic defies understanding on part of the user, and it is inconclusive to the extent that

there is a degree of uncertainty attached to the identification and classification of brain states.

It can be argued, however, that opacity in BCIs is not as relevant a question as it is in other algorithms. In BCIs, algorithms only use a relatively narrow set of data. The purpose of their use is, as stated previously, to identify relevant brain signals, classify them, and correlate them with a certain output. Compared with the huge amount of data acquired by, for example, a credit rating algorithm, the volume of BCI data is rather low. Thus, the range of possible outcomes that need to be understood is limited.

Moreover, contrary to algorithms in general, in BCI use, the user (i.e., the one who uses an algorithm to make decisions), and the entity mostly dealing with the effects of the algorithm, is one and the same person. In other contexts considered in the debate around algorithmic decisions, algorithms are used by one party (such as a bank, an employer, a court) to make decisions about another party (e.g., a person wishing to get credit, an applicant for a job, or an offender in a trial). If a person or passer-by sees a problem emerging with the algorithm in these cases, that person might complain directly to the respective user and ask for an explanation or, ultimately, the two parties might mutually agree that not using the algorithm is the preferred option in certain contexts. The consequences of algorithmic opacity seem thus to differ qualitatively in BCI use. A quadriplegic person who uses a BCI to direct a wheelchair is both the user of an algorithm and the recipient of its use, because it is that person who experiences the algorithm's outcomes. Moreover, in the case of BCIs, the algorithms are more closely linked to the action that the user wishes to perform, compared with other contexts of algorithm use (as in the case of a credit-rating algorithm). The problem of opacity seems, therefore, to be less far-reaching in BCIs. A BCI user cannot simply complain to another party because of the lack of transparency of an algorithm and ask or demand that the algorithm not be utilized in certain contexts. This picture and its related ethical consequences change again, if third parties are also involved through the BCI use, as might be the case if a wheelchair controlled by a BCI hits another person who happens to stand aside.

None of these thought experiments deny the opaque and inconclusive nature of the algorithms in BCIs as such. The algorithm's output, that is, a particular movement of the BCI-directed exoskeleton or wheelchair, might still be "surprising" or opaque for the user, precisely because the formation of that movement must be attributed to the algorithm whose decisionmaking process remains opaque. It is true that some amount of training is needed to enable a user to control an external device via the BCI, but presumably there remains a degree of uncertainty at least in some cases. This dynamic demonstrates that the inscrutable and inconclusive nature of algorithmic decisions in BCIs can be reframed as a safety problem. The BCI literature has addressed this problem as the trade-off between speed and accuracy: the longer an algorithm can do its work, the better the results of identification and classification.[42] Increasing BCI accuracy through training, however, is an often frustrating endeavor and can lead to a lack of trust in BCIs.[43] This would have serious implications for further studies in improving BCIs, with serious repercussions for those who stand to benefit from these technological advancements.

Algorithmic opacity is at the core of other problems that have been identified with BCIs, with one among these being privacy, for example. At first sight, the

issue of privacy as a right to control access to one's data,[44] is touched only to the extent that data are shared with third parties, because BCI algorithms are tailored to the individual brain. However, one might consider the case of a passive BCI that detects brain states unknown to the individual, and imagine that the BCI acts accordingly. For example, a BCI might detect a negative affective state during an interaction with another person. This might indeed become a privacy-related issue because it questions the individual person's capacity to control access to that person's core personality. In a sense, it is no longer the person who has privileged access, but the BCI. The BCI user shares effective control and access with the algorithms.

Such questions raised by the use of the algorithmically fixed BCIs have direct implications for the notion of autonomy. It rapidly becomes a questionable matter whether or not people are autonomous if they do not have privileged knowledge about the mental states that drive their motivations and actions. If algorithms are used to detect brain states and to inform an external device to perform a certain task, this might be seen as enhancing autonomy for those who cannot perform that task otherwise. But autonomy can also be reduced in BCI use if people have no insight into and control over the resulting effects of their activities with a BCI. Furthermore, if there is a constant mismatch between what users believe about their mental state, and what the BCI detects, users' self-concept might be confused to an extent that their ability for self-determination is ultimately diminished. Somewhat paradoxically, individually tailored algorithms could be designed to enhance autonomy by opening a path toward real self-knowledge, enabling people to find their "true" preferences, or deepest emotions. However, if users are not in the epistemic position to discern their real selves when the BCI gives outputs that are not clearly understood or expected, then autonomy may be significantly reduced.

A further aspect for ethical consideration is that the evidence generated by a BCI algorithm is often incomplete. Imagine that a person's BCI system detects that the person feels uneasy and therefore stops the person's wheelchair automatically. Whether the uneasiness was meant to be the action-guiding mental state is an open question. It seems that the question of which mental state should be considered as "action guiding," needs to be well defined prior to using a BCI, if a person wants to stay autonomous.

Because algorithmic opacity is said to contribute to biases and discrimination, the question arises whether opacity is observable in BCIs as well. Discrimination in BCIs appears to be not as clear-cut as in other cases. What could emerge as "discrimination" is a bias in the ability to use BCIs. As Wilhelm Bauer and Mathias Vukelic discuss,[45] elderly people exhibit differences in their assessment of negative emotions compared with younger people. This has an impact on older people's neural correlates of affectivity, which makes it more difficult to detect affective brain states that then are informing a passive BCI. If these results are confirmed in future studies, the problem of opacity for elderly BCI users is increased, and, more fundamentally, using a BCI will not be open to everybody on an equal level. To state that this is a problem requires evidence to the extent that people have a right to use such a technology, or have a vital interest in it. One could argue that should BCIs be used widely in society one day, the unequal ability to use a BCI constitutes a disadvantage. As Guglielmo Tamburrini[46] notes, unequal access to BCIs because of personal variations in BCI proficiency might raise questions of healthcare justice.

The workplace context is another area in which the age-based tendency to be unable to use BCIs could pose serious problems of discrimination. Note, however, that in this case, the ethical problem is not one that arises out of a working algorithm's inherent bias (wherever it comes from), but one of an algorithm being unusable for a given population. This adds a particular aspect to BCI ethics that goes beyond what the ethics of algorithms has offered so far.

Potential Solutions

Algorithmic *transparency* has been put forward as a solution to decrease the level of opacity; to increase trust, understanding and oversight, privacy, and autonomy; and to reduce bias and discrimination. One particular means to achieve transparency would be to refrain from using machine-learning algorithms altogether. This, however, would most probably lead to less effective BCIs.

A more promising approach would likely involve making the algorithm's model *interpretable post-hoc*; namely, by explaining the decision logic and working mechanism after a particular decision has been made. One can imagine that BCI users gradually learn about the algorithm and its working mechanism through behavioral, but also explanatory feedback, and can subsequently adapt to it. However, whether this is the case depends on the individual's ability to understand computational issues, and it makes room for potential discrimination because it limits the range of people who could understand and subsequently use BCIs with more insight. Post-hoc explanation might even lead to a false assurance and reliance on algorithms, potentially causing harm at a later point in time.[47] If a BCI repeatedly explains its decision, the user might increasingly trust the technology and lose the awareness that the technology could still be prone to making errors.

As a further option, *auditing* has been proposed as a solution to the problem of inscrutable evidence and opacity. Auditing does appear to be highly useful for BCIs. Trusted auditors such as researchers or external regulators could (be required to) analyze algorithms and either make them explainable or certify their acceptability. These auditors could examine cases of BCI uses and see how reliably the algorithms work (e.g., ex-post audit studies). They could run simulations or use algorithm internal reporting mechanisms. All these audits leave the black box model intact; that is, they do not attempt to understand the algorithm, but rather prove its reliability and verify correct functioning. However, more research is needed to find practical ways of conducting audits of BCIs, particularly because some of the scenarios envisaged here do not even exist yet.

Education in programming and coding has been put forward as yet another means to counter inscrutable evidence, and is also applicable in the case of BCIs. However, knowledge about programming is a rather generic solution that does not guarantee that a particular algorithm and its particular decisions are understood. And if some machine-learning algorithms are said to remain opaque even to experts, by extension this will hold true for educated BCI users. Using *open source* codes and software will not improve BCI transparency, because it is typically not the BCI user who elaborates on the BCI's algorithm.

Given the opacity of a BCI algorithm and the impact of this question of opacity on safety, privacy, and autonomy, it might be useful to implement measures that locate the final say as to what action the technological device performs with the

user. A *veto* right to intervene in the BCI-mediated action can be justified on these grounds, through implementing an *emergency stop*. This is, at root, a technical question that affects the design of the algorithm. The case appears to be that algorithms need to be created that do not only detect and classify brain signals, but also leave control over the action with the user. This distinguishes BCI algorithms from other, non-BCI instances of using algorithms. Here, users have control over the algorithms because they are the ones to decide whether or not to use an algorithm for a certain task; for example, to decide whether people who apply for credit can be granted their wish. There is no need to talk of a special veto right here, because the user would have the veto right in the first place (to use or not use an algorithm). One part of the explanation is that the action of interest here does not rely on algorithms. To decide whether an applicant should receive credit does not need to make use of algorithms. The decision can be made using various other means. In BCIs, however, the algorithm directly connects with the intention-generating and, therefore, action-initiating processes within a person's brain. A BCI is constitutive of an action and not merely a replaceable instrument. Therefore, a veto right appears to be necessary in BCIs, which complements what some authors have termed the "right to explanation" in the case of other algorithms.

With regard to the discriminatory elements in BCIs, available remedies do not directly target the algorithm and its design. Rather, these elements appeal to the way that those who put the BCI algorithm to use in the real world have to compensate for the resulting harms. It could include *mandatory provisions of alternative means* that those with less or no BCI proficiency can use. The workplace context is one area in which age-based tendency toward being unable to use BCIs could pose serious problems of discrimination. *Technological solutions* could be provided by taking into account the age-bias and program algorithms accordingly. *Political instruments* might be applicable here, too; for example, that using BCIs as automation tools in factories could require paying an extra tax.

On a more general level, the debate about the ethics of algorithms has also revealed various reinterpretations and reallocations of responsibility as a solution to the responsibility gap. One way is to maintain that there remains some portion of responsibility with designers and developers of algorithms to test their algorithms and to care for epistemic justification as well as nondiscrimination. It might be interesting and fruitful to approach this question from the perspective of corporate social responsibility[48] (CSR) or even political corporate social responsibility[49] (PSCR) so as to include a broader range of potential bearers of responsibility. This approach would introduce a debate that has a long-standing experience with ethically assessing market behavior and could shed new light on designing and selling medical devices.

As a final note, an entirely novel way to evaluate user responsibility might be necessary. This is reflected by attempts to conceptualize the notion of responsibility toward BCI-mediated actions as being analogous to the scenario of interaction with animals. To explain, similarly to the case of the owner of a domestic animal, the BCI user does not always know how the algorithm will eventually "act" or behave. Nonetheless, the owner has responsibility for the results of the interaction of the animal. Whether harmful or unwanted results occur depends on the knowledge, the virtues, or the morality of the user, or on other factors at least partially and sufficiently under the control of the user.

Andreas Wolkenstein, Ralf J. Jox, and Orsolya Friedrich

Conclusion

In this article, we have raised the central question of whether concerns raised in the general ethical debates about algorithms can be transferred to ethical discussions of BCIs. Our second goal was to see whether some of the solutions offered in the debate about the ethics of algorithms can be used to regulate BCIs as well.

The algorithms used in BCIs account for some of the problems identified with algorithms generally. For example, algorithmic opacity and inconclusiveness contribute to the safety problems associated with BCIs that do not work reliably. In addition, trust in BCI (and potentially in BCI research) might decrease. Moreover, to the extent that algorithms reliably identify and classify mental processes and states, this can lead to a problem concerning privacy and autonomy, because it is the algorithm and not the person that can be said to have privileged knowledge of the person's most personal states. Discrimination through BCIs is another issue that can be traced back to algorithms, but, importantly, this is caused by the algorithm not being usable for certain classes of users. This again raises important autonomy- and justice-related problems.

We have furthermore discussed various possible solutions to these concerns with algorithms. Restricting the use of algorithms or even of BCIs more generally in certain contexts and applications, and auditing, as well as a broader concept of responsibility, are part of the suggested solutions.

Future research needs to focus on particular ethical problems caused by algorithms and determine what solutions are available. For example, because some authors have argued that there is a responsibility gap caused by increased autonomous decisionmaking capacities of machines, the question arises as to what role algorithms play in this context. Some scholars have already noted the problem of traceability with regard to algorithms so that a deeper engagement with these ideas is called for. Moreover, some of the solutions to the problems with algorithms discussed previously need further investigation. For example, making use of CSR needs to be explored in a more detailed manner. This will allow a richer understanding of how ethical issues can be solved, and how responsibilities need to be allocated, respectively. It will thus enrich our understanding of the ethics of BCIs and its solutions.

Notes

1. Burwell S, Sample M, Racine E. Ethical aspects of brain computer interfaces: A scoping review. *BMC Medical Ethics*. 2017;18(60):1–11.
2. Grübler G, Hildt E, eds. *Brain–Computer Interfaces in Their Ethical, Social and Cultural Contexts*. Dordrecht: Springer; 2014.
3. Tamburrini G. Brain to computer communication: Ethical perspectives on interaction models. *Neuroethics* 2009;2(3):137–49.
4. See note 1, Burwell et al. 2017.
5. Clausen J. Man, machine and in between. *Nature*. 2009;457(7233):1080–1.
6. Grübler G. Beyond the responsibility gap. Discussion note on responsibility and liability in the use of brain–computer interfaces. *AI & Society*. 2011;26:377–82.
7. Grübler G. Shared control—shared responsibility? *International Journal of Bioelectromagnetism*. 2011;13(1):56–7.
8. Haselager P. Did I do that? Brain–computer interfacing and the sense of agency. *Minds & Machines*. 2013;23(3):405–18.
9. Matthias A. The responsibility gap: Ascribing responsibility for the actions of learning automata. *Ethics and Information Technology*. 2004;6(3):175–83.

10. Holm S, Voo TC. Brain–machine interfaces and personal responsibility for action – maybe not as complicated after all. *Studies in Ethics, Law, and Technology*. 2010;4(3)7.

11. Lucivero F, Tamburrini G. Ethical monitoring of brain–machine interfaces. *AI & Society*. 2008;22(3):449–60.

12. Mittelstadt BD, Allo P, Taddeo M, Wachter S, Floridi L. The ethics of algorithms: Mapping the debate. *Big Data & Society*. 2016;3(2):1–21.

13. Hill RK. What an algorithm is. *Philosophy & Technology*. 2016;29(1):35–59.

14. The article can thus be considered as realizing the idea that "[d]esigning imprecise regulation that treats decision-making algorithms, AI, and robotics separately is dangerous," and that "[c]oncerns about fairness, transparency, interpretability, and accountability are equivalent, have the same genesis, and must be addressed together, regardless of the mix of hardware, software, and data involved"; see note 12, Mittelstadt et al. 2016.

15. See note 8, Haselager 2013.

16. McFarland DJ, Wolpaw JR. EEG-based brain–computer interfaces. *Current Opinion in Biomedical Engineering*. 2017;4:194–200.

17. Lotte F, Congedo M, Lécuyer A, Lamarche F, Arnaldi B. A review of classification algorithms for EEG-based brain–computer interfaces. *Journal of Neural Engineering*. 2007;7(2):R1–R13.

18. Bauer W, Vukelic M. Forschungsprojekt EMOIO. In: Neugebauer R, ed. *Digitalisierung Schlüsseltechnologien für Wirtschaft & Gesellschaft*. Berlin & Heidelberg: Springer Verlag; 2018:135–51.

19. See note 16, McFarland, Wolpaw 2017.

20. See note 16, McFarland Wolpaw 2017; note 18, Bauer, Vukelic 2018.

21. See note 8, Haselager 2013.

22. Yuste R, Goering S, Agüera y Arcas B, Bi G, Carmena JM, Carter A, et al. Four ethical priorities for neurotechnologies and AI. *Nature*. 2017;551(7679):159–63.

23. See note 12, Mittelstadt et al. 2016.

24. See note 12, Mittelstadt et al. 2016.

25. See note 12, Mittelstadt et al. 2016.

26. O'Neil C. *Weapons of Math Destruction. How Big Data Increases Inequality nd Threatens Democracy*. New York: Crown; 2016.

27. See note 12, Mittelstadt et al. 2016.

28. van Wel L, Royakkers L. Ethical issues in web data mining. *Ethics and Information Technology*. 2004;6(2):129–40.

29. Eubanks V. *Automating Inequality. How High-Tech Tools Profile, Police, and Punish the Poor*. New York: St. Martin's Press; 2017.

30. Kroll J, Huey J, Barocas S, Felten EW, Reidenberg JW, Robinson DG, et al. Accountable algorithms. *University of Pennsylvania Law Review*. 2017;165:633–705.

31. de Laat PB. Algorithmic decision-making based on machine learning from Big Data: Can transparency restore accountability? *Philosophy & Technology* 2017.

32. Sandvig C, Hamilton K, Karahalios K, Langbort C. Auditing algorithms: Research methods for detecting discrimination on Internet platforms. Presented at Data and Discrimination: Converting Critical Concerns into Productive Inquiry, a preconference at the 64th Annual Meeting of the International Communication Association, May 22, 2014, Seattle, WA.

33. Tutt A. An FDA for algorithms. *Administrative Law Review* 2017;69(1):83–123.

34. Burrell J. How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*. 2016;3(1):1–12.

35. Lepri B, Oliver N, Letouze E, Pentland A, Vinck P. Fair, transparent, and accountable algorithmic decision-making processes. *Philosophy & Technology*. 2017 [epub ahead of print].

36. See note 35, Lepri et al. 2017.

37. Romei A, Ruggieri S. A multidisciplinary survey on discrimination analysis. *The Knowledge Engineering Review*. 2014;29(5):582–638.

38. See note 12, Mittelstadt et al. 2016.

39. Malle BF. Moral competence in robots? In: Seibt J, Hakli R, Norskov M, eds. *Sociable Robots and the Future of Social Relations: Proceedings of Robo-Philosophy 2014.* Amsterdam: IOS Press; 2014:189–98.

40. Malle BF. Integrating robot ethics and machine morality: The study and design of moral competence in robots. *Ethics and Information Technology*. 2016;18(4):243–56.

41. Anderson M, Anderson SL. Machine ethics: Creating an ethical intelligent agent. *AI Magazine*. 2007;28(4):15–26.

42. See note 8, Haselager 2013.

43. Glannon W. Ethical issues with brain–computer interfaces. *Frontiers in Neuroscience*. 2014;8:1–3.
44. See note 12, Mittelstadt et al. 2016.
45. See note 18, Bauer, Vukelic 2018.
46. See note 3, Tamburrini 2009.
47. See note 34, Burrell 2016.
48. Beckmann M, Pies I. The constitution of responsibility: Toward and ordonomic framework for interpreting (corporate social) responsibility in different social settings. In: Luetge C, Mukerji N, eds. *Order Ethics: An Ethical Framework for the Social Market Economy*. Dordrecht: Springer; 2016:221–50.
49. Scherer AG, Rasche A, Palazzo G, Spicer A. Managing for political corporate social responsibility: New challenges and directions for PCSR 2.0. *Journal of Management Studies*. 2016;53(3):273–98.