**Joanna K. Malinowska[1] & Tomasz Żuradzki**

**Non-Epistemological Values in Collaborative Research in Neuroscience: The Case of Alleged Differences Between Human Populations**

To cite this article:

Joanna K. Malinowska & Tomasz Żuradzki (2020) Non-Epistemological Values in Collaborative Research in Neuroscience: The Case of Alleged Differences Between Human Populations, AJOB Neuroscience, 11:3, 203-206, DOI: 10.1080/21507740.2020.1778126

To link to this article: https://doi.org/10.1080/21507740.2020.1778126

Published online: 27 Jul 2020.

The goals and tasks of neuroethics formulated by Farahany and Ramos (2020) link epistemological and methodological issues with ethical and social values. The authors refer simultaneously to the social significance and scientific reliability of the BRAIN Initiative. They openly argue that neuroethics should not only examine neuroscientific research in terms of "a rigorous, reproducible, and representative neuroscience research process" (148) as well as "explore the unique nature of the study of the human brain through accurate and representative models of its function and dysfunction" (148), but also its responsibilities or social consequences.

Let us take a closer look at the first neuroethics goal enlisted by Farahany and Ramos (2020, 149), which concerns the "necessary and beneficial collaborations for responsible discovery." The concept of responsible discovery or, in other words, responsible science, as such, raises many issues. As noted by Resnik and Elliott (2016), dilemmas related to this problem can be grouped into at least three basic categories: "(i) dilemmas related to problem selection, (ii) dilemmas related to publication and data sharing, and (iii) dilemmas related to engaging society" (Resnik and Elliott 2016, 31). We would like to concentrate on problem selection, which is shortly noticed by (Farahany and Ramos 2020, 10–11), and by BRAIN Initiative's Neuroethics Report itself (NIH 2019). In Chapter 6, the document raises an important issue related to problem selection, which is strengthening or perpetuating existing

[1] Joanna K. Malinowska, Adam Mickiewicz University, Faculty of Philosophy, ul. Szamarzewskiego 89c, 60-568 Poznań, Poland; malinowska@amu.edu.pl

prejudices and biases by choosing a research subject: "scientists are prompted to consider how the questions they choose to study in the laboratory might amplify existing biases." This leads to several further problems: what constitutes bias?; how biases may be embedded in the selection of research programs?; is it possible to conduct completely unbiased research?; who should be a gatekeeper in the case of research that may amplify biases? We try to notice possible answers to these questions in the context of the research on differences (e.g., cognitive, medical, behavioral) between human populations.

Choosing a research subject is a process influenced by numerous values accepted and shared by researchers themselves as well as their environment (Resnik and Elliott 2016). In order to study the process of subject selection and regulate it, it is indispensable to reconstruct these values and decide, which ones can be recognized in a democratic society, and which potentially disqualify a given topic from the research area. What values should we, therefore, consider wondering over whether the given research subject is worthwhile? We propose to specify four categories of values particularly relevant to the above question: social, ethical, epistemic, and epistemological.

The difference between the epistemic and epistemological values is that, while epistemic values (such as reliability, coherency, utility) regulate cognitive behavior of an individual subject in a "natural attitude": all processes related to perception, experience or common sense inference, i.e., colloquial knowledge; epistemological values apply at the level of scientific knowledge and practices. It is essential to acknowledge their role in the process of problem selection in neuroethics, especially in the face of the development of intergroup studies. The group differences in cognition are analyzed from many different perspectives—the studied categories may be representatives of different cultures, populations, professions, or admirers of various types of music. This area of research raises doubts concerning not only ethical, social, or epistemological values but also epistemic ones. For instance, should we evaluate different epistemic styles even though this may lead to a violation of ethical and social values (e.g., equality and fairness) and if—how to do that? Is our attitude to epistemic diversity the same as to neuronal diversity? Moreover, the more we know about the impact of culture on mind and brain (Farahany and Ramos 2020), the more we need to ask about if and how we should try to use this knowledge to shape people's cognitive capabilities (including those related to social cognition). For example, should we try to learn how to use different epistemic styles to achieve specific goals? Or, knowing that intergroup biases can be modulated by specific actions, e.g., motivational ones, should we try to use them to control phenomena such as xenophobia and racism (Malinowska 2016), although it can lead to abuse of power? It is also interesting to analyze the problem of epistemic values in the context of research on neuro- and cognitive enhancement (Farahany and Ramos 2020). There are many questions we should rise before we ask "if it works" (Farahany and Ramos 2020, 149).

For example, if we are going to try to expand or even change our cognitive abilities, what new epistemic norms will we accept as a society? And what epistemic values should guide or exclude such research areas? What will their relationship be to ethical or social values (again, they may be contradictory, e.g., when a specific enhancement realizes the value of epistemic utility but entails an increase in social inequalities)? Finally, due to the research possibilities offered by neuroscience, we face the need to reexamine discussions on epistemic conditions for moral responsibility as well as epistemic deontology (i.e., the question if a person has a moral obligation to develop her cognitive powers).

Although epistemological values (such as testability, reproducibility, explanatory power, empirical support, reliability) are not usually considered in the context of ethics, in the case of neuroethics (as well as bioethics) they seem to play a huge role, which actually has been highlighted by its goals formulated by Farahany and Ramos (2020). It is so, mainly because epistemological values may conflict with social or ethical values. To visualize the interplay between all of them, it is worth noting the problem with "ethnic" biases that the Authors have shortly mentioned. First, they notice that data-science approaches, e.g., AI "can encode—and thus amplify—gender, ethnic, and cultural biases, which can have adverse effects on research, commerce, and health care" (151). Second, they claim that brain data may "reiterate existing biases and generate potentially unfounded conclusions that could harm vulnerable individuals from racial and ethnic minority groups" (149).

In this first case, we assume they mean not only that these biases encoded in AI have some "adverse effects," but also that it is genuinely unjust to spread them with the help of science. The problem can be seen in an analogy to some discussions about the design of diagnostic strategies in medicine (Rzepiński 2018) or the program COMPAS, i.e., a well-known statistical method for quantitative risk assessment in criminal justice (although in contrast to medical diagnoses, legal systems in principio do not allow, with the exception of DNA evidence, to convict a suspect on the basis of the statistical evidence). In both of these cases, there are vivid discussions about including or using racial/ethnic categories to make diagnostic decisions or related to crime prevention (for the sake of argument, we assume that categories like Afro-American or white are well-defined and disjoint, although we believe it is generally false about race or ethnicity, see: (Malinowska and Żuradzki 2020)). In the case of medicine, some authors argued that using racial identification as a demographic characteristic with assumed biological implications may unintentionally contribute to racial discrimination (Malinowska and Żuradzki 2017; Perez-Rodriguez and de la Fuente 2017). Analogically, some argued that COMPAS might be biased against African-American defendants in comparison with white defendants (Angwin et al. 2016), although other researchers disagreed (Flores et al. 2016).

However, the problem is that the statement that biases may be encoded in AI used in medicine, healthcare, or crime prevention can be interpreted in at least three different ways related to three criteria of fairness. First, there is the calibration problem: the algorithm itself may not be well-calibrated, i.e., if we use probabilistic reasoning about the group of people (e.g., 30 individuals from the group consisting of 100 has a two-thirds chance of being rearrested—or having cancer recurrence—within two years), the algorithm is well-calibrated if, indeed, a 20 individuals from this 100 group have this feature (will be rearrested or will have cancer recurrence), and the condition is held simultaneously for all the groups we are interested in (e.g., African-American and white, men and women, etc.). Second, the false-positive balance: the algorithm may lead to the misidentification of low-risk individuals from one group as high-risk more often than from the other (i.e., African-Americans who are in reality in a low-risk group may be more likely to be incorrectly labeled as high-risk than they actually are, in comparison with low-risk whites). Third, the false-negative balance: the algorithm may lead to the misidentification of high-risk individuals from one group as low-risk more often than from the other (i.e., whites who are in reality in a high-risk group may be more likely to be incorrectly labeled as low-risk than they actually are, in comparison with high-risk African-Americans). However, the difficulty is that no algorithm can simultaneously solve each off the above problems and thus satisfy all three criteria of fairness, which means that risk assessments cannot be fair in general (Kleinberg et al. 2017).

In the second case, we interpret the Authors' claim about reiterating biases and harming vulnerable individuals from racial and ethnic minority groups (p. 4), as a call for either higher evidentiary standards or even to non-conducting, withdrawing or redirecting research that may be potentially harmful in that way. An excellent example of such research is investigating race/ethnicity-correlated differences in cognitive abilities trying to establish biologically-based differences by different methods.

We mention two types of generally understood harm caused by this kind of research: the first is indirect and leads to rectifying, objectifying and scientifying existing folk categories that are intertwined with discriminative practices, and in that way—indirectly—may help to maintain these practices. For example, some other researchers whose meta-analysis on the heritability of intelligence across "racial" or ethnic groups has been published in 2020 by the journal *Intelligence* were surprised that although they did not limit their search to just the United States, they received all samples only from there. They tried to explain this by many different ways (homogeneity, the lack of biometric research programs elsewhere), but they forgot about the simplest explanation: they implicitly used the US racial categorization based on OMB Policy Directive 15 that defines one ethnic category (Hispanic or Latino) and five racial categories (American Indian or Alaska Native; Asian; Black or African

American; Native Hawaiian or Other Pacific Islander; White) mistakenly and probably unconsciously treating this contingent classification as general, objective or scientific. Racial/ethnic classification is a mixture of folk racial categories based on phenotypic features like skin color, historical contingencies, and current political borders and interests. It is clearly visible if one compares this type of classifications in different legal contexts, e.g., the UK regulations distinguish between "British black Africans" and "British black Caribbeans" or between "Irish" and "British Irish"; in the most European legal regulations do not use "races", but use historic ethnic categories like "Tatars" and "Karaites" (these last two ethnicities are taken from Polish regulations).

The second type of harm is direct. How may research seeking to determine which "race" or ethnicity is better at solving IQ tests be directly harmful? Some authors argue that cognitive abilities are highly socially contingent, so just hearing repeatedly about "scientific" bases of some group's under-performance in cognitive abilities due to genetic factors may be enough to negatively affect this group performance, and therefore "scientifically" reproduce the stereotype (Dar-Nimrod and Heine 2006; Kourany 2016). Through this commentary, we presented evidence that neuroethical reflection has gone beyond traditional moral considerations. This situation requires not only deepening the studies of scientists working in its area (e.g., taking into account epistemic and epistemological values) but also their closer cooperation with representatives of other philosophical disciplines, e.g., methodology, philosophy of science, logic, and epistemology. Only establishing common goals and values by all these fields will enable us to consciously shape and develop neurosciences.

ORCID

Joanna K. Malinowska: http://orcid.org/0000-0002-8330-5234

Tomasz Żuradzki: http://orcid.org/0000-0001-6749-2124

REFERENCES

Angwin, J., J. Larson, S. Mattu, and L. Kirchner. 2016. Machine bias: There's software used across the country to predict future criminals. And it's biased against blacks. ProPublica https://www.propublica.org/article/machinebias-risk-assessments-in-criminal-sentencing (accessed May 27, 2020).

Dar-Nimrod, I., and S. J. Heine. 2006. Exposure to scientific theories affects women's math performance. Science (New York) 314(5798): 435–435. doi: 10.1126/science.1131100.

Farahany, N. A., and K. Ramos. 2020. Neuroethics: Fostering collaborations to enable neuroscientific discovery. AJOB Neuroscience 11(3): 148–154.

Flores, A. W., K. Bechtel, and C. T. Lowenkamp. 2016. False positives, false negatives, and false analyses: A rejoinder to machine bias: There's software used across the country to predict future criminals. And it's biased against blacks. Federal Probation 80(2): 38–46.

Kleinberg, J., S. Mullainathan, and M. Raghavan. 2017. Inherent trade-offs in the fair determination of risk scores. 8th Innovations in Theoretical Computer Science Conference (ITCS 2017) 67: 43:1–43:23. doi: 10.4230/LIPIcs.ITCS.2017.43.

Kourany, J. A. 2016. Should some knowledge be forbidden? The case of cognitive differences research. Philosophy of Science 83(5): 779–790. doi: 10.1086/687863.

Malinowska, J. K. 2016. Cultural neuroscience and the category of race: the case of the other-race effect. Synthese 193(12): 3865–3887. doi: 10.1007/s11229-016-1108-y.

Malinowska, J. K., and T. Z_ uradzki. 2017. The practical implications of the new metaphysics of race for a postracial medicine: Biomedical research methodology, institutional requirements, patient-physician relations. The American Journal of Bioethics 17(9): 61–63. doi: 10.1080/15265161.2017.1353181.

Malinowska, J. K., and T. Żuradzki. 2020. The ethics of constructing reference classes: Biomedical research methodology and race. Manuscript under review. NIH. 2019. The brain initiative and neuroethics: Enabling and enhancing neuroscience advances for society. https://braininitiative.nih.gov/strategic-planning/acd-workinggroups/brain-initiative-and-neuroethics-enabling-andenhancing (accessed June 5, 2020).

Perez-Rodriguez, J., and A. De La Fuente. 2017. Now is the time for a postracial medicine: Biomedical research, the national institutes of health, and the perpetuation of scientific racism. The American Journal of Bioethics 17(9): 36–47. doi: 10.1080/15265161.2017.1353165.

Resnik, D. B., and K. C. Elliott. 2016. The ethical challenges of socially responsible science. Accountability in Research 23(1): 31–46. doi: 10.1080/08989621.2014.1002608.

Rzepińnski, T. 2018. Twierdzenie Bayesa w projektowaniu strategii diagnostycznych w medycynie. Diametros 57: 39–60. doi: 10.13153/diam.1235.