

# Neuroprediction: New technology, old problems

Stephen J. Morse<sup>a</sup>

<sup>a</sup> Ferdinand Wakeman Hubbell Professor of Law, Professor of Psychology & Law in Psychiatry, Associate Director of the Center for Neuroscience & Society, University of Pennsylvania

Neuroprediction is the use of structural or functional brain or nervous system variables to make any type of prediction, including medical prognoses and behavioral forecasts, such as an indicator of future dangerous behavior for the purpose of involuntary civil commitment. This note will focus on behavioral predictions, but the analysis applies to any context. The general thesis is that using neurovariables for prediction is a new technology, but that it raises no new ethical issues, at least for now.

Institutions, including the legal system, routinely use behavioral predictions, which can have momentous implications for the life of the individual, for institutions and for society at large. In the non-legal realm consider predictive tests for admission to educational institutions or employment opportunities. In the law, predictions can have particularly grave consequences, including the loss of liberty. For examples, the following all centrally involve a prediction of future dangerous conduct: involuntary civil commitment of the mentally ill, including so-called “mentally abnormal sexually violent predators”, granting bail to a criminal defendant, sentencing a criminal defendant, potentially even to death in the United States in many states, and granting parole. In principle, neurodata alone or in conjunction with other types of variables might be used for these and myriad other predictions.

Such serious and potentially intrusive prediction decisions raise important ethical issues. Nonetheless, private, semi-public and public institutions have all made the normative judgment that employing predictions is justifiable and, indeed, it is unlikely that any of these institutions could function adequately without them. The question is what criteria are and should be used to decide if a particular prediction is justifiable. I believe that there are two major considerations: the accuracy of the prediction, which is an empirical question; and the rate and type of inevitable prediction errors that are acceptable given the interests of the predictor and subject of the prediction, including whether other important values are potentially violated by gathering the basic data necessary for the prediction. We shall discuss these generally in order and then will turn to how neuroprediction bears on both.

No predictive method is error free, much like all diagnostic tests. All produce both false positives and false negatives, and low base rate behavior, such as suicide or homicide, is particularly prone to false positive predictions. At present, three methods are primarily used

for behavioral prediction: clinical judgment guided by the clinician’s own training and experience; structured professional judgment in which the clinician typically uses a validated actuarial or semi-actuarial tool but then can modify the result using professional judgment; and actuarial, in which the data gathered are largely objective and then a statistically validated algorithm is applied. The comparative efficacy of these three methods is much-studied and the conclusion is inescapable that clinical judgment is the least accurate. There is dispute about whether actuarial is superior to structured professional judgment, but both are more accurate than clinical judgment. Even the best validated predictive methodologies still have substantial error rates, however, which motivates the search for better tools. Interestingly, the search does not require theoretical understanding of the causes of behavior, although such understanding might well improve predictions. It is sufficient if a large data base provides reasonably accurate markers even if the reason that a variable predicts accurately is not understood.

Deciding what rate and types of error are justifiable is a normative issue that can be resolved only by balancing the various interests implicated by the prediction, including the consequences to the subject and society and the cost of producing the prediction. Consider the example of involuntary civil commitment based on a prediction of future dangerousness. Forced hospitalization and mental health treatment are an enormous intrusion on liberty, but avoiding harm to self or others is an entirely worthy goal. The more serious the harm predicted, the greater weight that must be given to harm-reduction, but how many people may appropriately be hospitalized who would not cause serious harm in order to prevent one person who will? Do we prefer to hospitalize unnecessarily to prevent a few harms or to hospitalize infrequently knowing that some preventable harms will result? These and similar questions can only be resolved morally, politically and legally taking into account actual error rates and costs for various types of prediction, assuming that such data are available.

Gathering the information necessary to make a prediction may involve the intrusion on other important values even if the variables involved may increase the accuracy of a prediction. Consider the example of privacy. A subject’s genetic background, alcohol consumption and sexual activity may all increase the accuracy of some types of prediction, but obtaining these data

will surely intrude on the subject's privacy. Moreover, many types of data can be put to nefarious uses. These are familiar issues in bioethics that are often raised when considering health insurance and gathering public health information generally. Whether the intrusion is justifiable is again a normative question that can be answered only by balancing the interests involved.

Now let us turn to the implications of neuroprediction for the above considerations. At present, there are very few neuropredictive tools for behavioral prediction. There are some "proof of concept" studies indicating that adding a neural variable to behavioral measures can increase slightly the accuracy of predictions of anti-social conduct, but these techniques are simply not sufficiently established to be used for public policy purposes. There are interesting studies suggesting that neural variables can help predict which mental health treatments will be effective with some types of patients, but such predictions do not raise the types of ethical concern that predictions of anti-social conduct or other types of socially consequential predictions do. In short, neuroprediction for public policy purposes is at present more of a hope than a reality, but future studies will certainly provide better data. Moreover, we can predict quite confidently, that as the neuroprediction tools become more refined and produce a larger data base, we will be able to have a sense of how accurate they are alone or in tandem with other predictive methods. This will require expensive, methodologically difficult studies to perform, however, so I suspect that well-validated neuroprediction tools will not be produced in the near future. Note that there is nothing unique about neuroprediction in this respect. We can ask the question of comparative accuracy about any technique.

Assuming that we have a sufficient data base to know the error rates of various types of neuroprediction, will it be justified to use this methodology? We have already decided as a society that predictions are normatively justified. If neural variables increase the accuracy of such predictions, are not unduly costly to obtain and do not intrude on other values we endorse, what possible argument could there be for not using neural variables? How could we possibly justify engaging in a practice less successfully when a technique is available to do it better? Those who worry that neuroprediction may someday be "too" accurate have a substantial burden to explain why too much accuracy will undermine the well-established normative justifiability of predictive practices. Note again that the same questions can be raised about any technique that increases accuracy, whether it is genetic, neural or behavioral.

Is there something unique about neuroprediction that raises new ethical issues? Obtaining structural and functional brain scans is now quite expensive, but so are some behavioral measures, such as obtaining a psychopathy score using the *Hare Psychopathy Checklist-Revised*, and the costs of scans and other neural measurement techniques will surely decrease over

time. By analogy, consider how much less it costs to sequence an individual's genome than when the technique was first devised. And yet again, considering the cost of neuroprediction raises no new issues. Cost must always be balanced against the potential increase in accuracy for any prediction tool.

Does collecting neural data for prediction intrude more fundamentally on privacy and dignity than other techniques? Virtually all conceivable neural measures will involve "brain reading", not mind reading. Although there are now relatively accurate neural measures that can identify beyond chance whether, for example, a subject is adding or subtracting, or looking at a place or a face, these techniques do not identify the particular content of the cognition. They do not indicate what numbers are being added or subtracted or what precise face or place is being observed. They simply identify those regions of brain activity that are associated with the general activity in question. The same will be true of neuroprediction. Neural variables associated with the predicted behavior will not provide access to the content of the subject's mental states. If neural techniques could genuinely read minds, a hitherto unimaginable ethical challenge would be raised, but such an ability is science fiction at present.

There are a host of other issues neuroprediction raises that are thoroughly familiar to bioethicists because they are also raised by other techniques, such as the right response to incidental findings, whether the technique can be used without the subject's cooperation, and whether the data collected can be put to illegitimate uses. Even in the unlikely event that neuroprediction were to raise such issues more acutely than other techniques, only the magnitude of the issue and not its novelty is raised. At a certain point, of course, changes in quantity produce changes in quality, but there is little reason to believe that neuroprediction raises such a possibility.

In conclusion, neuroprediction may or may not become a useful tool, but if it does, we already fully possess the ethical theoretical resources to address any resulting ethical and legal challenges.

---

#### Correspondence

Stephen J. Morse, J.D., Ph.D.  
University of Pennsylvania Law School  
3501 Sansom Street  
Philadelphia, PA 19104-6204  
Philadelphia, PA  
U.S.A.

E-mail: smorse[at]law.upenn.edu