Response to the National Academy of Science *Decadal Survey of Social and Behavioral Sciences for Applications to National Security*

Nine Areas of Neuroethical Importance

Submission by the Emerging Issues Task Force
Mark Frankel, Chair
Laura Cabrera, Adrian Carter, Philipp Kellmeyer, Karola Kreitmair, Anthony Weiss

of the

International Neuroethics Society
Judy Illes, President
June 2017

In response to the National Academy of Science *Decadal Survey of Social and Behavioral Sciences for Applications to National Security* and call for commentary, members of the Emerging Issues Task Force of the International Neuroethics Society offer their insights here on nine areas of critical importance.

Lie Detection/Truth Confirmation

Neuroimaging technology has long been explored as a means for lie detection (Wolpe, Foster & Langleben 2005; Stoller & Wolpe 2007). There is preliminary evidence to suggest that deception is associated with activity in the ventrolateral and dorsolateral prefrontal cortex, inferior parietal lobe, anterior insula, and medial superior frontal cortex (Farah, Hutchinson, Phelps & Wagner 2014). There are considerable concerns, however, with the current state of technology, such as confounding factors and problems of inference with respect to measurement (see e.g. Ganis et al. 2011), as well as transferability from experimental settings to forensic or military settings. Multiple efforts have been made to admit fMRI evidence in court, although to date no such request has been granted (Farahany 2016; Gaudet & Marchant 2016; Wagner et al. 2016). However, if in national security settings, the purpose of this technology is not to convict or exonerate individuals, but merely to gather information, we might think that the bar for validity may be lower. But how low is too low?
If the accuracy of such lie detection technologies became reliable, what sorts of purposes might be appropriate for its use? Might such technology be ethically used in the interrogation of suspects of a crime, persons applying to work for the intelligence agencies, or immigrants and refugees? What role, if any, would physicians play in this process, given ethical prohibition of their involvement in coercive interrogation or torture (Miles 2009)? Furthermore, simple countermeasures (such as moving vigorously in the scanner) can easily invalidate the fMRI measurement, such that the real applicability of this technology remains doubtful.

Beyond neuroimaging, pharmacological agents such as oxytocin have shown some promise of inducing an attitude of "trusting behavior" (Baumgartner et al 2008), which may be employed in interrogation settings, and for which similar questions of validity and appropriateness apply. However, these findings have only been demonstrated in laboratory settings that are unlikely to be significantly relevant in the real world.

If lie detection/truth confirmation technology or pharmacology were indeed to rise to a reliable level of validity, it certainly seems conceivable that it could be exploited to force individuals to reveal information in ways that violate ethical guidelines of information-gathering. Individuals who may have a right to non-self-incrimination under international guidelines may be forced to divulge information against their will. This technology may allow the meaningful decoding of brain activity that goes beyond what has previously been possible, raising questions concerning cognitive liberty. Some have argued that the mind ought to be a protected domain of an individual, and it should not be be penetrated against that individual's wishes (Wolpe, Foster & Langleben 2005). However, one might argue that the consequentialist rewards of interference trump any such deontological qualms.

Neurointerventions in Warfare

Neuroenhancement: What are the ethical implications of using neurotechnologies to help soldiers remain awake and alert during extended combat or to improve their critical thinking and reaction times? The Defense Advance Research Project Agency (DARPA) has several active research programs on this. For example, Electrical Prescriptions (ElectRx) program is using ultraminiaturized devices, which could be delivered through minimally invasive procedures, to help the human body heal itself through neuromodulation of organ functions (DARPA website). Another program relevant for neuroenhancement is the Targeted Neuroplasticity Training (TNT) program which seeks to enhance learning of a wide range of skills to reduce cost and duration of training as well as to improve outcomes in the battlefield.

Neuroimpairment: If during military conflict the brain could be manipulated to produce temporary unconsciousness, dizziness, or psychotic effects in order to confuse the enemy or make them less effective warriors, when should such action be employed? How do neurobiological weapons differ from other forms of chemical or biological warfare? (Many chemical agents, such as nerve gas, already interfere with the nervous system). In the Moscow theatre hostage siege in 2002, the potent opioid fentanyl was used to sedate the hostage-takers, while the soldiers were inoculated against the drug with an opioid antagonist. If done via mass applications, what ethical considerations should apply to civilians located nearby? What physical or mental side effects should be assessed? How can an ethical research design be constructed to determine the impacts of various approaches (e.g., a spray, electrical signals)? There is also the possibility of using neurotechnologies to detect deficiencies in a warfighter's neurological level of functioning and use this information to deploy mechanisms to restore or normalize the warfighters neurological function.

Memory Manipulation

In principle, different methods, such as pharmacological or electrophysiological interventions, are available or in development to interfere with memory at different stages (formation, consolidation or recall). Beta blockers like propranolol, which function by inhibit the activity of the sympathetic nervous system, have been shown to dampen the affective component of traumatic memories (Henry et al 2007). This may be beneficial in preventing, or perhaps even treating, PTSD, which can help stem the suffering experienced by warfighters. However, as some have pointed out (Hurley 2007.), there may be a risk of eroding a person's moral sensibilities, if the emotional content of a memory is divorced from the factual. Perhaps this intervention can block feelings of guilt or regret associated with combat or military experiences, but is such blocking always desirable, or do warfighters depend on such affective

states to make moral decisions in complex situations?

Deception

How might neuroscience be used to fuel or impede deception aimed at mass populations via, for example, propaganda, "fake news," or other communication strategies aimed at distorting one's perception of reality during heightened national security concerns? What ethical considerations should guide the use of such strategies or efforts to combat them?

In terms of impeding mass deception a readily available tool is education. Using training programs that are designed to improve specific neurocognitive functions, such as working memory, may have effects on overall brain function and may improve a person's ability to critically assess information she received (Ansari, Smedt & Grabner 2011). However, living in highly technological and information loaded societies, the capacity of the human brain might be limited to assess all the information needed. This is where the use of neuroenhancers may play a role, from pharmaceuticals to neurointerfaces, that enable people to better judge different facts, context and make connections that might facilitate knowing when they are being deceived. There could also be external technologies, like augmented reality apps, where people can input chains of events and through advanced algorithms that double check facts get a result about the accuracy of their perception of reality.

There is however a potential dual use of this increased knowledge about neurocircuitry to devise advanced methods for mass deception. This might even include the insertion of false memories, which has already been achieved with the use of optogenetics in animals (Ramirez et al, 2013).

Brain-to-brain Communication

If two individuals are being connected via Brain-Computer Interfaces to enable communication or to manipulate each other (e.g. via extra- or intracranial electrical stimulation), such direct "brain-to-brain communication" could interfere with an individual's ability to distinguish his or her "own" brain activity (which manifest as thoughts or feelings) from the "inception" or "intrusion" by another individual's transmitted brain activity. Does this violate some fundamental freedom of humans to be allowed to think whatever they want (e.g., cognitive liberty)? How should such concerns be weighed against the need to hold someone accountable for their thoughts? Imagine, for example, if during the interrogation of a suspect who's brain activity was measured with EEG and directly connected to an interrogator's EEG (who perhaps has the capability of sending electric stimulation to the suspect's head), the interrogator uses brain-activity-dependent (non-painful) stimulation to coerce the suspect to reveal his or her thoughts. Would the end of receiving vital information justify the means of this neurointervention, or would we consider this to be a form of non-justified interrogation method that unduly violated their human rights, or possibly amount to torture?

AI/Machine Learning

There are efforts to complement or even replace human decision-making with machine learning/AI in intelligent and autonomous weapon systems. When decisions get offloaded on AI, what kinds of consequences does that have for responsibility with respect to the decisions made and potential catastrophic side effects (such as civilian casualties)? Does AI ever have to make decisions that require moral reasoning? If so, how can such reasoning be built into robot design and behavior?

Robotic Combat

If AI leads to robots capable of substituting for human warfighters, what ethical constraints should be considered for their use? Depending on the level of intelligence and autonomy of such advanced robotic systems, should they have similar rights, e.g. with regard to their structural integrity and continuity of existence, as human warfighters? If not, at what point along the human/robot spectrum of assistive wearable to "cyborg" to robot do these rights cease to exist? Moreover, what should such distinctions be based on?

Predictive Analytics

The use of advanced machine learning methods, such as "deep learning", to predict people's inclination toward terrorist behavior or some other threat to national security raises concerns about how such techniques would be used to restrict individual liberties (e.g., increased monitoring by the government, travel constraints). A recent study using a new machine learning method looked for multiregional brain activity patterns that could collectively predict "culpable" mental states (Vilares et al 2017). Such algorithms have been shown to incorporate inherent biases that adversely affect individuals from certain minority groups that result from historical social and cultural inequity. Do these predicitve analytics need to correct for these inherent biases? What rights do individuals have to know what data are being used against them and to correct inaccuracies? Does the government have a responsibility to restore a person's reputation or compensate victims for the incorrect or intentional misuse of the data? Now, just before the impending advent of individualized consumer wearable technology for recording brain activity (with EEG or NIRS), there is a historic opportunity to safeguard neural data against widespread dissemination and misuse. How should the use and security of neural data be governed and legislated? Should there be a law that prohibits the selling of neural data (analogous to the prohibition against selling one's organs)?

Neurofeedback for Weapons or Drone Control

Neurological control of weapons or drones became a possibility when in 2007, researchers taught a monkey to neurologically control a walking robot on the other side of the world by means of electrochemical measurements of motor cortical activity. As less invasive technologies, like neurofeedback, are improved these can be more readily used in the continuous trend to create distance between parties involved in conflict.

References

Ansari, D., Smedt, B., & Grabner, R. H. (2011). Neuroeducation – A Critical Overview of An Emerging Field. Neuroethics, 5(2), 105–117. http://doi.org/10.1007/s12152-011-9119-3

Baumgartner, T., Heinrichs, M., Vonlanthen, A., Fischbacher, U., & Fehr, E. (2008). Oxytocin shapes the neural circuitry of trust and trust adaptation in humans. *Neuron*, *58*(4), 639-650.

Farah, M. J., Hutchinson, J. B., Phelps, E. A., & Wagner, A. D. (2014). Functional MRI-based lie detection: scientific and societal challenges. *Nature Reviews Neuroscience*, *15*(2), 123-131.

Farahany NA. Neuroscience and behavioral genetics in US criminal law: An empirical analysis. Journal of Law and the Biosciences 2016;2:485–509.

Ganis, G., Rosenfeld, J. P., Meixner, J., Kievit, R. A., & Schendan, H. E. (2011). Lying in the scanner: covert countermeasures disrupt deception detection by functional magnetic resonance imaging. *Neuroimage*, *55*(1), 312-319.

Gaudet LM, Marchant GE. Under the radar: Neuroimaging evidence in the criminal courtroom. Drake Law Review 2016;64:57.

Henry, M., Fishman, J. R., & Youngner, S. J. (2007). Propranolol and the prevention of post-traumatic stress disorder: Is it wrong to erase the "sting" of bad memories?. *The American Journal of Bioethics*, *7*(9), 12-20.

Hurley, E. A. (2007). The moral costs of prophylactic propranolol. *The American Journal of Bioethics*, *7*(9), 35-36.

Ramirez, Steve, et al. "Creating a false memory in the hippocampus." *Science* 341.6144 (2013): 387-391.

Stoller, S. E., & Wolpe, P. R. (2007). Emerging neurotechnologies for lie detection and the fifth amendment. Am. JL & Med., 33, 359.

Vilares, I. Wesley, M.J., Ahn, W. et al. 2017. Predicting the knowledge–recklessness distinction in the human brain. PNAS, 114 (12): 3222-3227.

Wagner, A. D., Bonnie, R. J., Casey, B. J., Davis, A., Faigman, D. L., Hoffman, M. B., ... & Richeson, J. A. (2016). fMRI and Lie Detection. MacArthur Foundation Research Network on Law and Neuroscience, Vanderbilt Law Research Paper No. 17-10.

Wolpe, P. R., Foster, K. R., & Langleben, D. D. (2005). Emerging Neurotechnologies for Lie-Detection: Promises and Perils. The American Journal of Bioethics, 5(2), 39–49. http://doi.org/10.1080/15265160590923367