Non-Invasive Brain Stimulation in the Detection of Deception: Scientific Challenges and Ethical Consequences

Bruce Luber, Ph.D.*, Carl Fisher, B.S.†,
Paul S. Appelbaum, M.D.‡,§,
Marcus Ploesser, M.D.§
and Sarah H. Lisanby, M.D.†,§

Tools for noninvasive stimulation of the brain, such as transcranial magnetic stimulation (TMS) and transcranial direct current stimulation (tDCS), have provided new insights in the study of brain–behavior relationships due to their ability to directly alter cortical activity. In particular, TMS and tDCS have proven to be useful tools for establishing causal relationships between behavioral and brain imaging measures. As such, there has been interest in whether these tools may represent novel technologies for deception detection by altering a person’s ability to engage brain networks involved in conscious deceit. Investigation of deceptive behavior using noninvasive brain stimulation is at an early stage. Here we review the existing literature on the application of noninvasive brain stimulation in the study of deception. Whether such approaches could be usefully applied to the detection of deception by altering a person’s ability to engage brain networks involved in conscious deceit remains to be validated. Ethical and legal consequences of the development of such a technology are discussed. Copyright © 2009 John Wiley & Sons, Ltd.

*Correspondence to: Bruce Luber, Ph.D., Division of Brain Stimulation and Therapeutic Modulation, Department of Psychiatry, Columbia University College of Physicians and Surgeons/New York State Psychiatric Institute, 1051 Riverside Drive, Unit 21, New York, NY 10032, U.S.A.
E-mail: luberbr@pi.cpmc.columbia.edu

†Division of Brain Stimulation and Modulation, Department of Psychiatry, Columbia University College of Physicians and Surgeons and New York State Psychiatric Institute.
‡Psychiatry, Law and Ethics, New York State Psychiatric Institute.
§Department of Psychiatry, Columbia University College of Physicians and Surgeons.

Contract/grant sponsor: Defense Advanced Research Projects Agency (DARPA).
The fields of law and security have had a longstanding interest in methods of detecting deception. Ascertaining whether witnesses are lying is in some sense what the adversarial process is all about (Appelbaum, 2007). The inaccuracy of juries and judges in assessing veracity has led many parties to court proceedings to seek out technological approaches to distinguishing truth from falsehood, often very soon after a new technology has presented itself as a possible means of lie detection. The first approaches to deception detection relied on peripheral, anxiety-induced autonomic indicators, using polygraphs measuring pulse, blood pressure, respiration, and galvanic skin response. Unfortunately, concerns about the accuracy of these methods, especially the ability of individuals to deceive investigators by controlling their autonomic reactivity, have limited their use and acceptance in courts of law (see, e.g., Keckler, 2005; Stern, 2002). Basically, the problem is that peripheral physiological responses are secondary to the primary central nervous system processes generating acts of deception.

In recent years, researchers have turned to brain imaging techniques to study deceptive processes in the brain itself. Electrophysiological methods have shown great promise, with some components of the event-related electrical response having been shown to provide a useful index of the presence of concealed memories (see, e.g., Farwell & Donchin, 1991). A great deal of attention is now being paid to the possibility that brain imaging will prove to be the sort of reliable lie detection technology that will be able to provide admissible evidence in courts of law (Keckler, 2005; Meegan, 2007; Tovino, 2007). Indeed, some researchers have formed companies that claim to provide just such fMRI evidence (e.g. www.NoLieMRI.com). A number of studies have used functional MRI (fMRI) and positron emission tomography (PET) to identify the neural substrate supporting deceptive behavior. These have used a number of deceptive tasks and scenarios, including guilty knowledge tasks (GKT: Langleben et al., 2002, 2005; Phan et al., 2005), mock crime scenarios (Kozel, Padgett, & George, 2004; Kozel et al., 2005; Mohamed et al., 2006), feigned memory impairment (Lee et al., 2002, 2005), and autobiographical or experienced events (Abe et al., 2006; Abe, Suzuki, Mori, Itoh, & Fuji, 2007; Ganis, Kosslyn, Stose, Thompson, & Yurgelun-Todd, 2003; Nuñez, Casey, Egner, Harre, & Hirsch, 2005; Spence et al., 2001).

**LIMITATIONS OF BRAIN IMAGING IN DETECTING DECEPTION**

While measuring brain activity directly through brain imaging is considered to have more scientific promise than the peripheral measures used in existing tests, this approach is also fraught with limitations (for reviews, see Appelbaum, 2007; Sip, Roepstorff, McGregor, & Frith, 2007; Wolpe, Foster, & Langleben, 2005). Across the variety of deception tasks studied, there has been a certain amount of agreement in cortical regions activated, notably anterior cingulate cortex, dorsolateral and ventromedial prefrontal cortex, and parietal cortex. However, there are also considerable differences in activated regions across studies. For example, in two studies subjects participated in simulated “crimes” before being scanned. In one, subjects were instructed to steal a certain object, and then lie about it during the MRI session (Kozel et al., 2005). Deceptive responses activated primarily prefrontal...
cortex. In another study, a group of subjects fired a gun and were told to lie about doing it (Mohamed et al., 2006). In this case, the only frontal activation was in the anterior cingulate cortex, while much of the activation was posterior, in ventral and left lateral occipital visual areas. A potential source of the apparent discrepancies across these and other studies may lie in the differential demands of the deception tasks used, such as processing auditory or visual stimuli, responding verbally or with button presses, and working with various sorts of content, such as playing cards, autobiographical knowledge, or rehearsed scenarios. Each different task requires varying levels of executive processing and working memory load, which may alter the degree to which prefrontal and anterior cingulate regions are activated. Brain activations that are so heavily task dependent may not reveal generalizable deception-related signals that are task independent and applicable to real-world deception, in which the task-related demands may be unknown.

The number of approaches to deception tasks taken in various imaging studies underscores the complexity of the processing involved with a deceptive act. These involve both the cognitive and emotional processes necessary to generate the rationale, intent, and strategies for deception within a given context (Sip et al., 2007), as well as those needed to execute a response that is incompatible with the truth (Johnson, Barnhardt, & Zhu, 2003). The former set of processes shows a great degree of individual variability (Keckler, 2005), while the latter set involved in execution most likely involves varying levels of general executive processes, including those that identify and inhibit the truthful response, select and carry out a response that conflicts with the truth, and monitor internal consistency and external results of the deceptive act. This processing complexity, and the concomitant complexity of its neural underpinnings, has been acknowledged (see, e.g., Nuñez et al., 2005), and some studies have attempted to differentiate component executive processes used in the deceptive act with manipulations of working memory load (Ganis et al., 2003) or memory content (Nuñez et al., 2005). Nevertheless, the number of processes and their potential variations in usage between individuals, task types, and contexts makes the establishment of brain–behavior relationships in imaging of deceptive acts problematic. The complexity of deceptive acts and its variability across individuals and contexts makes a fallback position of carrying out discriminant analyses on the imaging data recorded during deceptive and non-deceptive acts (see, e.g., Kozel et al., 2005), without regard to underlying processes, problematic as well (Sip, Roepstorff, McGregor, & Frith, 2008).

The interpretation of imaging data becomes even more fundamentally problematic because it is correlational, and reverse inferences from physiological activations to mental states cannot be made with certainty (Sip et al., 2007). A cortical area can be involved with more than one cognitive function, and its activation during a deceptive act does not prove that it is directly involved in deception processing. This is certainly true of the prefrontal, parietal, and anterior cingulate regions commonly activated in deception studies, as these areas are generally activated in studies of executive processing when no deception is involved. Thus the greatest limitation of brain imaging approaches is that they may not actually be informative about deception selectively, calling their validity into question. Put another way, merely showing that a region of the brain activates when a person is trying to lie does not prove that region of the brain is involved in deception. It might be a sign of anxiety, for example, or some other nonspecific function that
accompanies deception but is not in and of itself conclusive of deception. This limitation in interpretation of brain imaging findings means that these techniques can only reveal correlations between brain states and behaviors—they cannot prove causation. In other words, even if an individual demonstrates a brain imaging finding correlated with deception, this does not constitute proof that the person is lying. This limitation of brain imaging, that it can only show correlations and not prove causation, is recognized throughout neuroscience as a universal limitation of this technique.

A NOVEL APPROACH TO EXAMINE CORTICAL PROCESSING UNDERLYING DECEPTION: NON-INVASIVE BRAIN STIMULATION

However, new developments in neuroscience now permit us to directly test the implications of brain imaging findings, providing the necessary data to demonstrate that the brain imaging result is truly causally linked to a specific behavior such as deception. Two such techniques, called transcranial magnetic stimulation (TMS) and transcranial direct current stimulation (tDCS), are noninvasive means of temporarily changing the function of precise regions of the brain. While brain imaging provides a measure of brain activity, TMS and tDCS change brain activity. This is why these two forms of brain stimulation can move beyond correlation to test causal relationships. Brain stimulation can thus be used to test the validity of brain imaging findings associated with deception.

TMS is a noninvasive method of focally altering cortical brain activity through the use of targeted magnetic fields that temporarily disrupt ongoing neural processing. A TMS device emits brief pulses of current through a stimulating coil held on the head. The current flow lasts less than a millisecond, and it produces a rapidly changing magnetic field around the coil. This magnetic field in turn induces current flow in cortical tissue near the surface of the head, stimulating neurons in a focal region, generating a brief modulation of neural activity. The actual neurobiological effects depend on the intensity of the magnetic field, the coil shape and its orientation, the geometry of the underlying cortex in relation to the field, the latency of stimulation relative to ongoing cortical processes, and the frequency of the pulses in the stimulation train. On the other hand, tDCS noninvasively modulates cortical excitability through the use of small, generally imperceptible DC electrical currents flowing between anodal and cathodal scalp electrodes. Anodal tDCS is thought to increase cortical excitability in the targeted brain region, while cathodal stimulation decreases cortical excitability. The excitability shifts are thought to be due to subthreshold neuronal membrane depolarization (Priori, 2003).

Because it can act directly on cortical circuitry in a noninvasive, spatially focal way with a very fine temporal resolution, TMS has proven to be a useful tool for establishing causal relationships between behavioral and brain imaging measures. A review of the neuroscience applications of TMS can be found in Luber, Peterchev, Nguyen, Sporn, & Lisanby, 2007b). For example, a single pulse from a magnetic coil placed over the occipital cortex can successfully mask a visual target stimulus when the magnetic pulse follows the visual presentation by 80–100 ms, the time when visual information reaches the cortex from the eye (Amassian et al., 1989). TMS
used to study cortical processing by disrupting ongoing processing in this way has been called a “virtual lesion” method (Pascual-Leone, Walsh, & Rothwell, 2000). Using TMS in this way can reveal detailed information about how the brain processes information. In one example of this in the sensory domain, by applying single pulses of TMS to motion-sensitive occipital regions while moving stimuli were shown to subjects, it was determined that the brain can analyze direction of motion separately from speed, which informed theoretical models of motion processing (Matthews, Luber, Qian, & Lisanby, 2001). Another form of modulation using TMS can alter neural network behavior to facilitate performance. For example, TMS applied to parietal cortex at five pulses per second during the seven second retention period of a working memory task sped up the reaction time of responses to a memory probe (Luber et al., 2007a). TMS has already been used in these ways to modulate cortical activity and thus study a wide range of motor, sensory, and cognitive processes. Studies of motor systems include investigations of motor planning (Kennerlet, Sakai, & Rushworth, 2004; Serrien, Strens, Oliviero, & Brown, 2002), motor imagery (Pulvermüller, Hauk, Nikulin, and Ilmoniemi, 2005), consolidation of motor learning (Muellbacher et al., 2002), cognitive and sensory interactions with motor cortex (Bonnard, Camus, de Graaf, and Pailhous, 2003) and eye movement (O’Shea, Muggleton, Cowey, & Walsh, 2004). The large amount of cerebral cortex devoted to it has made the study of visual perception quite amenable to TMS techniques. Many aspects of visual processes have been studied, from binocular (Saint-Amour, Walsh, Guillemot, Lassonde, & Lepore, 2005) and motion discrimination (Matthews et al., 2001) to word recognition (Lavidor & Walsh, 2003), illusory contours (Brighina, Ricci, & Piazza, 2003) and visuospatial processing (Aleman et al., 2002). Beyond motor and sensory systems, TMS effects have also been uncovered in higher order cognitive processing, in areas such as attention (Rushworth, Paus, & Sipila, 2001), task switching (Rushworth, Hadland, Paus, & Sipila, 2002), response selection (Hadland, Rushworth, Passingham, Jahanshahi, & Rothwell, 2001), and memory (both short term, Rami, Gironell, Kulisevsky, Garcia-Sanchez, & Estevez-Gonzalez, 2003; Sandrini, Cappa, Rossi, Rossini, & Miniussi, 2003, and long term, Rossi et al., 2001), as well as mental imagery (Kosslyn et al., 1999) and language processing (Shapiro, Pascual-Leone, Mottaghy, Gangitano, & Caramazza, 2001). Although tDCS has only more recently been applied to studying cognitive processing, it has already been used to affect capacities such as verbal fluency (Iyer et al., 2005), categorization learning (Kincses, Antal, Nitsche, Bartfai, & Paulus, 2004), motion perception (Antal et al., 2004), and working memory (Boggio et al., 2006).

USING NON-INVASIVE BRAIN STIMULATION TO STUDY COMPONENT COGNITIVE PROCESSES USED IN DECEPTION

As detailed in the earlier section on the limitations of brain imaging in detecting deception, it is implausible that there is a single and simple “deception network” to be discovered. Instead, it is likely that deceptive behavior results from a systematic combination of cognitive processes, including those necessary to generate the
rationale, intent, and strategies for deception within a given context, as well as those needed to execute a response that is incompatible with the truth. A deeper understanding of these component processes, as well as an ability to manipulate them, will be necessary to transform the findings of neurophysiological research in deception into an effective deception detection technology, and brain stimulation has made some inroads in this regard.

In one useful taxonomy of cognitive processes involved in interpersonal deception, Sip and colleagues (2007) listed four categories: information management, risk management, impression management, and reputation management. While little has yet been done with brain stimulation in the latter two categories, which have to do with building trust within interpersonal communication and the moral dilemma of justifying the deception to self and others, much has been done in the first two. Information management has to do with the executive processes involved with identification and inhibition of truthful responses, selection, execution, and memory maintenance of responses that conflict with the truth, and monitoring the internal consistency and external results of the deceptive act. A large number of brain stimulation studies have successfully modified the processes of working memory, planning, and response inhibition and selection necessary to information management (Luber et al., 2007b). For example, TMS has been shown to improve the maintenance of working memories (Luber et al., 2007a) and to disrupt processing of self/other distinctions and relationships (Lou et al., 2004). Both of these processes are necessary to information management of deceptive acts.

A second category of deceptive processes studied using brain stimulation is risk management, which is a necessary component of deception, as the costs of being caught in the act or of not being believed may have dire consequences, and must be carefully weighed. The decision processes involved with risk taking have been studied using brain stimulation, with both TMS (Knoch et al., 2006; van’t Wout, Kahn, Sanfey, & Aleman, 2005) and tDCS (Fecteau et al., 2007). Knoch and colleagues used a well known gambling paradigm called the Risk Task (Rogers et al., 1999), in which subjects are shown six boxes colored pink or blue (in ratios of 5:1, 4:2, or 3:3), and are asked which color box a target token will be found in. The ratio of pink to blue determines the probability of finding the token in a box of a particular color and thus defines the “level of risk” of the subject’s choice. In addition, the number of reward points for a correct choice or penalty points for an incorrect one also appears during each presentation, in ratios of 90:10, 80:20, 70:30, or 60:40, with the higher reward or penalty always being for the lower probability choice. Thus for example in the case of a presentation with five pink boxes and one blue box, with a payoff ratio of 90:10, a subject would be given (or lose) ten points if he or she chose the safer option (the higher probability “pink”), or would be given (or lose) 90 points if “blue” was the response. The goal was to accumulate as large a point total as possible. Before the task, 1 Hz rTMS was applied for 15 minutes. rTMS delivered at this low frequency acts to lower cortical excitability and thus disrupt neural processing in the targeted region for a short period after the train is completed. After stimulation of the right dorsolateral prefrontal cortex (DLPFC), subjects engaged in significantly riskier decision making, resulting in lower overall point totals, than subjects who received stimulation to left DLPFC or sham stimulation to either site. Those receiving right side rTMS chose the high risk targets more often than the left side and sham groups, especially when the payoff ratios were 90:10 and 80:20. The
authors interpreted these results to indicate that right (but not left) DLPFC plays a crucial role in suppressing superficially seductive options.

This conclusion was supported in a second brain stimulation study, this time using tDCS (Fecteau et al., 2007). Once again the Risk Task was used, but here groups of subjects received anodal tDCS to the right DLPFC and cathodal tDCS to the left DLPFC, cathodal tDCS to the right and anodal to the left, or sham tDCS. As anodal tDCS is thought to increase cortical excitability (and thus facilitate neural processing) of the targeted area, and given the hypothesized role of right DLPFC in suppressing risky behavior, anodal stimulation to the right was expected to result in decreased risk taking compared to the other conditions. This was exactly what happened, as subjects receiving anodal right tDCS chose the less risky option at a significantly higher rate than those receiving anodal left or sham, which resulted in their amassing much higher overall point totals than the other two groups. Moreover, the safer, higher probability choices by right anodal subjects were made with much less regard to the payoff ratios. These results again suggest that the right DLPFC has a role in suppressing seductive but unproductive options, and thus of modulating risk taking behavior. Together these TMS and tDCS studies demonstrate the usefulness of brain stimulation for analyzing some of the neural underpinnings of component processes necessary to risk management in deception.

USING NON-INVASIVE BRAIN STIMULATION TO DIRECTLY AFFECT DECEPTIVE BEHAVIOR

To date, only a few studies have used brain stimulation to directly investigate such processes through manipulations of deceptive behavior. In one study, TMS was delivered to the left and right motor cortices of eight subjects before and immediately after a question was answered with a false response (Lo, Fook-Chong, & Tan, 2003). A TMS pulse of sufficient intensity applied to the motor cortex will elicit activity in muscles controlled by the motor neurons stimulated, as measured at the muscle as motor evoked potentials. The amplitude of the motor evoked potential elicited by TMS was larger after a ‘lie’ response than before, an effect not seen with ‘truth’ responses. The effect was somewhat general, as amplitudes were the same regardless of complexity of the deceptive behavior (simple yes/no responses versus complex free form answers) or hemisphere of stimulation, and “lie” amplitudes were larger than “truth” amplitudes regardless of whether they were recorded pre- or post-stimulation. These results suggest that generating a false response was associated with increased excitability in the corticospinal tract, perhaps reflecting increased motor readiness or a general effect of arousal. While this demonstrated that TMS could be used to detect differences in cortical excitability between the generation of deceptive and truthful responses, the study did not examine the neural substrate of deceptive processing by testing whether TMS could alter deceptive responding.

Another study used tDCS in an attempt to directly alter deceptive responses (Priori et al., 2007). In this study, a deception task known as the Guilty Knowledge Test (GKT: Lykken, 1960; MacLaren, 2001) was used. GKTs require subjects to answer a series of yes/no questions, with instructions to answer some questions truthfully and other questions untruthfully. They are considered to be one of the most promising tasks for detecting deception (Ben-Shakhar and Elaad, 2003) and
are well suited to psychophysiological paradigms (Sip et al., 2008). Here, subjects selected five cards that had pictures on them from a set of ten. They were then shown the ten pictures one by one in random order, and asked “Do you have this picture?”. This was done over 80 trials. Before each trial, subjects were cued to either lie or tell the truth. This resulted in four conditions, two deceptive—denying one had the picture when one did (a denial lie), and pretending one had the picture when one did not (a false fact generation)—and two truthful. Anodal tDCS over the bilateral DLPFC significantly increased reaction time (RT) of 15 subjects when they denied for ‘denial lies’ but not for ‘false fact generation’ or ‘truth’ responses. This result suggested that noninvasive brain stimulation could indeed modulate deceptive processing, here by noticeably slowing the responses of participants when they attempted to lie. However, tDCS may not be an ideal technique for examining the complex processing involved in deception compared with TMS, due to its relatively lower spatial resolution and lack of temporal specificity. It should be noted that the spatial resolution in this study was decreased by using bilateral stimulation. This may have been unnecessary, as EEGs recorded during deception tasks have shown a decided right hemisphere advantage (Malcolm and Keenan, 2005).

A TMS study of deception was recently performed in our own laboratory (Luber et al., in revision). We tested the hypothesis that single pulses of TMS could influence cortical networks involved in a simplified deception task. By stimulating cortical areas activated by deception tasks in imaging studies, we sought to disrupt the neural circuits necessary for the formation of deceptive responses, and thus interfere with deceptive behavior. We adapted a paradigm from fMRI studies of deception using a playing card GKT (Langleben et al., 2002, 2005). Subjects were given a hand of six playing cards. Digital images of playing cards were displayed one at a time on a computer screen. Subjects were instructed to confirm or deny their possession of a given card using response buttons as fast as they could. Three of the six cards were designated as “lie” cards, which subjects were to deny having in their hand by pressing the no button. They were to continue pressing “yes” when any of the other three cards appeared, and “no” to not-in-hand cards. A single pulse of TMS was delivered at 0, 80, 160, 240, or 480 milliseconds in random order after the onset of presentation of each card (the time between the card onset and the TMS pulse is called a stimulus onset asynchrony, or SOA). Selection of the cortical areas targeted for stimulation was based on the work of Langleben et al., which found activations in left dorsolateral prefrontal cortex (DLPFC) and precuneus. Those two sites were chosen for stimulation as the active sites, while a third site, right DLPFC, which was not activated by the task, was chosen as a control site.

In 15 subjects stimulation to midline parietal cortex substantially slowed (by 20%) responses in trials when deception was required compared with trials when deception was not required, suggesting a critical role for the parietal cortex in intentional false responding. This effect was attributable to longer mean response times for lie cards than for truth cards or other cards at 240 ms SOA. Prefrontal stimulation, on the other hand, did not show this slowing with deceptive responses. The lack of specific effects in the DLPFC was unexpected, given that the region is generally the most strongly activated in imaging studies involving deception. This particular result may have been due to the nature of the GKT used here, which was relatively easy to perform. The most difficult aspect of the GKT (and a critical process in deception) is keeping track of which of the arbitrary stimuli are to be
responded to truthfully or not. The parietal cortex appears to be critical in such mental searches (e.g. in TMS, Luber et al., 2007a; in imaging of deception, Lee et al., 2002). On the other hand, no deception effects caused by TMS were found in the two frontal sites. This can either mean that the single TMS pulses at the latencies used were not sufficient to disrupt deceptive processing there, or that any processing associated with deception was not critical to response production. Given that the GKT we used had probably the simplest form a deception task can take, with a framework of simple categorization using well known stimuli, this may well have been the case.

FUTURE RESEARCH

While the studies described above reveal the promise of brain stimulation in uncovering and influencing the neural substrate of deceptive processing, the technology already exists for markedly improving such research. Spatial targeting would be significantly improved by first obtaining functional brain images specific to the deception task used with fMRI, and then to use the individual brain images to guide the choice of TMS targets on an individual subject basis (Herwig et al., 2003). We have found fMRI-guided TMS, informed by individual differences in network expression, to be useful in the modulation of higher cognitive processes such as working memory (Luber et al., 2007a, 2008). Likewise, temporal targeting of TMS pulses can be guided by electrophysiological recordings, which have similar millisecond time resolution as TMS. For example, the timing of pulses could be chosen to match the peak activity of specific ERP components generated by the particular task used. This approach could take advantage of the well developed body of electrophysiological research into deception detection, especially with late positive event-related components (see, e.g., Farwell & Donchin, 1991; Johnson et al., 2003). Another approach is to use trains of repetitive TMS (rTMS) pulses to affect processing by modulating local cortical excitability. The right set of frequency, duration, and timing of rTMS trains has been shown to have profound effects on cortical activity, especially when guided by individual electrophysiological dynamics (Klimesch, Sauseng, & Gerloff, 2003; Huang, Edwards, Rounis, Bhatia, & Rothwell, 2005).

Use of this sophisticated technology will allow the exploration of the complex processing of deceptive behavior. The use of playing card GKT in our TMS study informed us about the dynamics and neural substrate necessary for the execution of a simple deceptive response, corresponding most directly to bluffing or deceiving in a card game. The GKT used minimized the other aspects involved with deception, related to cognitive and emotional processes in generating a lie in a larger context and in relation to the truth, as well as monitoring its effects. A more ecologically valid approach might examine deception using real-world knowledge, both autobiographical and more general. In law enforcement and courtroom settings, detained persons and witnesses often will be responding to questions about themselves and their past behavior, and about their environment. Using more complex knowledge representations could illuminate more pre-frontal processes of interference and conflict resolution, response inhibition, and higher level cognitive control (such as awareness that a false act has been done and awareness it was not a mistake) that may
be more central to understanding and manipulating real-world deception. In addition, contrasting autobiographical and more general knowledge may also highlight such processes due to their relevance to an individual’s personal circumstance.

**IMPLICATIONS OF NON-INVASIVE BRAIN STIMULATION APPLIED TO DECEPTION**

The initial indication from TMS research is that, under certain circumstances, magnetic stimulation can slow deceptive responses, distinguishing them from truthful ones. In the TMS study by Luber and colleagues, as well as the tDCS study by Priori and colleagues, the slowing effect was found in the group mean data. Although this sort of effect is useful to research exploring brain mechanisms of deceptive behavior, it is a far cry from an effect observable in an individual, on a trial-by-trial basis, that is necessary in a true lie detector. Nevertheless, these studies represent fledgling steps. Given the existing technology to improve spatial and temporal targeting with EEG- and fMRI-guided TMS, and with sufficient manipulation of the timing parameters of TMS delivery, it is certainly possible that in the future TMS or other non-invasive brain stimulation technologies may be applied as lie detectors, by interfering with neural processes necessary for generating deceptive responses, making the behavioral difference clear between an individual’s lies and truth telling. If this scientific challenge is ever successfully met in the future, legal, ethical, social, and professional challenges must be considered as well, as discussed below.

**PRACTICAL CONSIDERATIONS: MEETING THE DAUBERT STANDARDS**

For TMS to be used in deception detection, practical issues of reliability, validity, and implementation must be addressed. In the case of law enforcement, for example, most states apply the Daubert standards for the admissibility of scientific information. The U.S. Supreme Court decision in *Daubert v. Merrell Dow Pharmaceuticals* required parties attempting to introduce evidence in federal court based on a scientific technique to provide indicia of the technique’s reliability and validity, such as tests of the falsifiability of the theory on which it is based; peer review of data generated by its use; documentation of a known error rate; and general acceptance by experts in the field (*Daubert v. Merrell Dow*, 1993; Garland & American Association for the Advancement of Science, 2004). Repeated concerns about traditional polygraph tests have been raised (Stern, 2002), and the design of TMS studies of deception will face the same challenges. If the Daubert standards are to be met, and to be confident in the results of TMS for deception detection, further questions must be answered. As outlined in the section on “Future Research” above, a good deal of work in developing reliable techniques remains to be performed. In particular, generalizability must be shown more conclusively both within individual behavioral context and over less tightly controlled populations. Investigations to date have relied on a controlled laboratory environment, in which there has been strict regulation of stimulus
conditions and researchers have controlled when and whether the subjects were to lie. As detailed by Sip and colleagues (2007), general deceptive behavior requires that the subject has the freedom to choose both the content of his/her deceptions and when it is expressed. Furthermore, there is similarly close control of subject selection and screening. The performance of healthy volunteers in a research study may not be immediately generalizable to that of subjects in the criminal justice or military arenas, who may have health (especially neurologic) problems, psychiatric conditions, substance abuse issues, or other potential confounds.

Moreover, subjects attempting to use countermeasures may successfully reduce the efficacy of TMS for deception detection. It has long been recognized that countermeasures against polygraphy are effective across multiple test conditions (Honts, Raskin, & Kircher, 1994). Similar measures could be used by subjects who were familiar with the test procedure for TMS-based deception detection; for example, if subjects knew that reaction time is an important outcome measure in these studies, they could intentionally “jitter” their responses to create statistical uncertainty and obscure difference between lie and truth responses. Furthermore, TMS has high spatial resolution, thus subtle movements of the head could interfere with experimental results. Simply moving one’s head away from the coil by just 0.5 to 1 cm would likely cause significant changes but would be difficult to observe in practice. Techniques to block the countermeasures, perhaps based on field testing, would be necessary to address these concerns.

**PRACTICAL CONSIDERATIONS:**

**INFORMED CONSENT ISSUES**

If the *Daubert* standards are met and these questions are sufficiently addressed, additional practical considerations such as safety and informed consent must be considered before this technology is adapted for use in the general public. Single-pulse TMS is considered to be minimally invasive in properly screened adults without risk factors for seizure, but repetitive TMS (rTMS), which might also be employed for the study of deception, does carry both a small risk of seizure and the possibility of unknown effects on neural functioning. Because of these risks, and other potential unknown longitudinal effects, the informed consent process must be particularly rigorous with regard to full disclosure of potential risks and side-effects and capacity assessment (Illes, Gallo, & Kirschen, 2006). TMS has followed safety guidelines since 1998, especially as regards screening for subjects with an increased risk of seizure or other neurological vulnerabilities (Wassermann, 1998). Even freely consenting subjects seeking TMS for legal purposes must still be rigorously screened and appropriately excluded if they have such vulnerabilities. tDCS, on the other hand, has an even more benign risk profile than TMS. tDCS has not been demonstrated to induce seizures, and has no known serious adverse effects. Of course, the possibility of using either technology against a subject’s will raises even more complex ethical and legal challenges.

**LEGAL CONSIDERATIONS**

Assuming that the *Daubert* or equivalent standard is met and the other practical concerns are addressed, the application of brain stimulation techniques would raise
other legal questions in the U.S. system. Use of TMS for lie detection might take place under two general scenarios: subjects *volunteering* to receive brain stimulation, for example to prove their veracity, or subjects being *compelled* to undergo testing with the technology (Appelbaum, 2007). Even in the case of voluntary use, the evidence would only be admissible if the judge found that its probative value outweighed any tendency it might have to confuse or mislead the jury. Whether a defendant’s decision not to undergo brain stimulation deception-detection would be admissible as evidence intended to impeach his or her credibility would need to be decided. The Fifth Amendment’s right against self-incrimination would be the obvious source of protection in this case, and would likely prevent this information from being admissible in legal proceedings (Greely, 2004). However, in the case of widespread application of this technology—for example, if opposing witnesses all submitted to brain stimulation deception tests—the potential exists for people to be subtly pressured into agreeing or judged untrustworthy for not doing so. The use of “DNA dragnets,” in which all individuals in a community receive genetic testing and a refusal to submit to testing can be seen as an implicit admission of guilt, is a particularly relevant comparison (Cho & Sankar, 2004).

Thornier constitutional issues arise in the case of subjects being compelled to receive brain stimulation. Under our current system, witnesses can be compelled to testify, or even to undergo testing such as medical examinations, so there may be precedent to compel subjects to submit to brain stimulation deception-detection. A key consideration will be whether TMS-based testing is considered analogous to a blood test, which can be taken from an unwilling defendant, or to an utterance, which cannot be compelled. Several constitutional protections might be invoked to oppose compulsory lie detection technology (Greely, 2004). In addition to the Fifth Amendment protections against self-incrimination just mentioned, a witness might also claim Fourth Amendment protection against unreasonable search and seizure. Doing so to protect the contents of one’s mind, so to speak, is unprecedented but may be increasingly relevant as the ability to detect truthful or false responses develops. First Amendment claims might even be used, as interference with free responses to questions could be proposed as a violation of the right to free mentation and speech. This would be the first time that the limits of “cognitive liberty” would be defined and tested, so there is no clear legal precedent (Wolpe et al., 2005).

The use of brain stimulation-based deception-detection outside of the U.S. criminal justice system—for example, in cases of national security—would raise yet another set of concerns. In this situation, we suspect that the role of medical professionals in TMS studies would become central to legal considerations. In practice, physicians must evaluate and monitor subjects receiving brain stimulation. Recently the role of medical professionals in Guantanamo Bay interrogations has attracted much attention both within the profession and in the popular press (Bloche & Marks, 2005a). The ensuing debate has clarified the role of doctors in interrogation as defined by the Geneva Conventions and other instruments of international law. There is strong consensus that physicians should not use drugs or “other biologic means” to subdue enemy combatants or extract information from detainees, they should not be party to interrogation practices contrary to human rights law or the laws of war, and their role in legitimate interrogation should not extend beyond limit setting as guardians of detainee health (UN, 1982). Whether or not these guidelines speak against the involvement of physicians in the use of brain
stimulation technology is open to question and likely depends on context. The U.S. military has long insisted that the medical profession has a social role that encompasses the needs of the criminal and civil justice system in addition to the individual provision of care (Bloche & Marks, 2005b). This stance is said to be consistent with the involvement of forensic psychiatrists in employee evaluations of fitness, and overall the consideration of limited resources and of protecting the public health occasionally trumping individual rights or preferences (Janofsky, 2006). Critics counter by noting that police, intelligence, or military interrogators routinely use deceptive techniques and other coercive measures to obtain information, but such tactics are contrary to a physician’s ethical and professional duties (Rubenstein, Pross, Davidoff, & Iacopino, 2005; Janofsky, 2006). Indeed, the American Medical Association (2006) and the American Psychiatric Association (2006) have rejected a direct role of any sort for physicians in interrogations in general. If these technologies were to be developed for use in the military against enemy combatants or other non-U.S. citizens, the role of physicians would likely come under great scrutiny.

ETHICAL CONCERNS

Outside of the legal sphere, brain stimulation deception-detection raises broader ethical issues. As a general principle, a particularly relevant consideration to the use of brain stimulation technologies is that of autonomy, a critical component of bioethics. Brain stimulation works by directly modulating brain function, in essence directly manipulating a subject’s behavior. Of course, brain function has been manipulated for years through more conventional means, as people’s minds are quite well equipped to exert personal and social influence through the sending and decoding of information (Walsh, Desmond, & Pascual-Leone, 2006). Nevertheless, the use of a technology that is capable of directly modulating brain function and behavior raises important ethical and even philosophical questions about the status and position of autonomy in our understanding of human behavior (Gazzaniga & Steven, 2004; Morris, 2007). A thorough treatment of neuroscience and free will is far beyond the scope of this discussion; suffice to say that the use of brain stimulation technologies even on freely consenting volunteers must be tempered by the potential for infringing on individual autonomy.

There are also social and ethical considerations regarding the use and spread of neuromodulatory approaches to detecting deception. Neuroscience can powerfully influence people’s understanding of psychological phenomena, even when the science itself is completely irrelevant to the underlying explanation (Weisberg, Keil, Goodstein, Rawson, & Gray, 2008). Previous authors have documented two logical traps in scientific communication that affect the public’s understanding of neuroscience: “neuro-essentialism,” reducing complex psychological phenomena to simplistic explanations; and “neuro-realism,” the uncritical acceptance of results simply because they contain neuroscience data (Racine, Bar-Ilan, & Illes, 2005). Already, many “neuromarketing” companies are selling fMRI and other neuroscience techniques to find the “buy button” in the mind of consumers (Brammer, 2004). Scientists studying deception should be careful of the way they communicate their results so as to avoid hype and premature responses to the ultimately unavoidable pressure to develop and promulgate new technology.
Finally, the use of brain stimulation to study deception raises the specter of other, perhaps more troubling applications that might be used to alter subjects’ behaviors and not simply probe for the presence or absence of information. Other authors have suggested that brain stimulation might be used to “improve the social relationship between interrogators and suspected terrorists” (Canli et al., 2007). Putting aside the low potential for efficacy of such measures, this application certainly would be a more troubling infringement on personal autonomy and free choice. Similarly, the neuropeptide oxytocin has been shown to increase trust in laboratory studies of human and other animal behavior (Kosfeld, Heinrichs, Zak, Fischbacher, & Fehr, 2005). Given that this substance is already available in the U.S. in injectable and nasal spray forms, it is less far fetched to see an immediate potential application for oxytocin as a trust modulator in the clinical setting and the legal sphere (Penney & McGee, 2005). The use of neuromodulatory approaches to deliberately change behavior should be approached with great caution.

The studies of brain stimulation deception-detection described earlier in this article are subject to a number of important ethical considerations. Concerned observers might raise the specter of “thought control”; however, the techniques so far developed simply represent the use of brain stimulation to probe for information and not a concerted effort to change subject behavior. The present level of brain stimulation research in deception is scientifically close to, but as yet distinct from, the manipulation of trust or social behavior to extract information from subjects. For now, the relevant issues revolve around applicability, legality, professionalism, and public perceptions of neuroscience. It would behoove investigators, policymakers, and the public, however, to begin to consider the possibility of more potent neurotechnologies for legal and military uses, and the proper scope and limits to such work (Moreno, 2006).

ACKNOWLEDGEMENTS

This research was supported by a grant from the Defense Advanced Research Projects Agency (DARPA). Approved for public release, distribution unlimited. Dr. Lisanby has received research support, for topics not presented here, from Magstim Company, Neuronetics, Cyberonics, and ANS. Columbia University has applied for a patent for novel TMS technology developed in Dr. Lisanby’s laboratory, for work unrelated to the topic presented here.

REFERENCES


UN. (1982). *Principles of medical ethics relevant to the role of health personnel, particularly physicians, in the protection of prisoners and detainees against torture and other cruel, inhuman, or degrading treatment or punishment*. Adopted by UN General Assembly resolution 37/194 of 18 December 1982.


