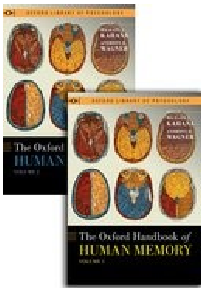

The Oxford Handbook of Human Memory, Two Volume Pack: Foundations and Applications

Michael J. Kahana (ed.), Anthony D. Wagner (ed.)

https://doi.org/10.1093/oxfordhb/9780190917982.001.0001

Published: 2024 **Online ISBN:** 9780190918019 **Print ISBN:** 9780190917982

Search in this book

CHAPTER

81 Brain-Based Memory Detection and the New Science of Mind Reading

Jesse Rissman, Emily R. Murphy

https://doi.org/10.1093/oxfordhb/9780190917982.013.81 Pages 2329–2350

Published: 18 July 2024

Abstract

Neuroimaging studies reveal distinct brain activations when people encounter something they recognize relative to something novel. Such a "neural signature" of memory could theoretically be used as a forensic tool to detect whether or not someone remembers a given entity. This chapter examines the ways that researchers have used electroencephalography and functional magnetic resonance imaging to capture temporal and spatial brain activation profiles that index different recognition memory states. Studies have addressed forensically relevant factors such as the examination of memories acquired in real-world contexts, classification of individual subjects rather than analysis of group differences, and the effect of subjects' deployment of evasive countermeasures. Recent development of multivariate analysis techniques, capable of decoding brain activity patterns on individual trials, show promise for yielding inferences about a subject's memory for specific stimuli or event details. Critical methodological shortcomings that may ultimately limit the forensic value of brain-based memory detection are discussed.

Keywords: memory detection, recognition, retrieval, autobiographical, concealed information test, guilty knowledge test, P300, multivoxel pattern analysis

Subject: Cognitive Psychology, Cognitive Neuroscience, Psychology

Series: Oxford Library of Psychology

Collection: Oxford Handbooks Online

Introduction

There has been growing interest among psychologists, neuroscientists, legal scholars, and the general public regarding the use of functional neuroimaging techniques to “read out” information about a person’s mental state.[1] Improvements in neuroimaging technology, coupled with advances in data analysis methods—including the use of machine learning algorithms to classify complex patterns of brain activity—have made it possible to decode some rudimentary information about what a person is currently perceiving (Nishimoto et al., 2011), feeling (Saarimäki et al., 2015), thinking about (Naci et al., 2013), imagining (Boccia et al., 2015), remembering (Rissman et al., 2010), attending to (Datta & DeYoe, 2009), intending to do (Gilbert & Fung, 2018), or even dreaming about (Horikawa et al., 2013). One such application that has garnered particular attention is the prospect of being able to detect the presence or absence of a memory for a past event. In a forensic context, a neuroscientific memory detection technique could conceivably be used to interrogate the brains of suspected criminals or witnesses for neural evidence that they recognize certain people, places, or objects, such as those from a crime scene. If such a procedure were sufficiently reliable, it could in principle be used to supplement, or even supplant, eyewitness testimony, which is known to be deeply vulnerable to memory distortions and erroneous recall (Loftus, 2017; Schacter & Loftus, 2013). ↳
p. 2330
Indeed, a recent U.S. governmental report highlighted the potential societal impact that something like this could have (Presidential Commission for the Study of Bioethical Issues, 2015):

> If we could accurately interrogate the brain, with a high degree of reliability, then just as DNA evidence has helped to exonerate many wrongfully accused and convicted individuals, so too might neuroscience offer greater accuracy and insights to improve our laws and policies. We should be open to the possibilities that neuroscience can bring, while ensuring the progress and responsible application of neuroscience to the legal system and policymaking. (p. 102)

Of course, for a brain-based memory detection method to be of practical value, it would need to have both high sensitivity (i.e., likely to detect a memory when one is present) and specificity (i.e., unlikely to indicate the presence of a memory for probes that an individual has not previously experienced). There are also a host of theoretical, methodological, and ethical factors that will be critically important to address before mainstream adoption in the justice system would be warranted. This chapter reviews the scientific literature on memory detection and examines the ways that functional neuroimaging techniques such as electroencephalography (EEG) and functional magnetic resonance imaging (fMRI) have been used to identify “neural signatures” of human memory, with an emphasis on understanding the strengths and limitations of each method and the challenges inherent in any effort to use brain activity as a proxy for probing memory for events of the past.

What Is “Memory Detection”?

Before proceeding, it is important to clarify the meaning of “memory detection” because this term is sometimes confused with lie detection. Lie detection involves the search for physiological (e.g., increased heart rate or skin conductance) or neural correlates of deceptive responding, typically associated with the heightened arousal and cognitive conflict that one experiences when suppressing a truthful response and generating an untruthful one (for a critical review of neuroimaging-based lie detection, see Farah et al., 2014). In contrast, memory detection refers to the probing of an individual’s memory with the purpose of determining whether a given stimulus is something that they have encountered in their past or whether it is being encountered for the first time. The core idea is that the experience of recognizing a familiar (and often personally significant) stimulus should evoke a brain response that is sufficiently different from that evoked by a novel and/or nonmeaningful stimulus. Ideally, the neural activity indicative of stimulus recognition would be evoked regardless of whether the subject is motivated to conceal this memory. However, as we discuss below, this is often not the case, as many studies have shown that subjects can be taught to adopt simple cognitive strategies known as countermeasures to diminish the likelihood that their memories will be detected. Regardless of potential susceptibility to countermeasures, memory detection approaches fundamentally differ from lie detection approaches in that the goal is not to detect a neural correlate of deceptive thinking (e.g., activity associated with the cognitive conflict and heightened arousal that occurs when one produces a response that one knows to be at odds with the truth) but, rather, to differentiate recognized stimuli from novel ones.

EEG-Based Approaches

There is now a substantial body of research using EEG-based techniques to detect memories, reporting impressive results on accuracy of detection. This section first reviews the basic technique and theory behind EEG-based memory detection, and then it reports on recent findings and remaining limitations on forensic use.

p. 2331 EEG-based memory detection protocols rely on measuring brain activity as a subject is presented with a series of stimuli, typically words or pictures shown on a computer screen, but auditory stimuli can also be used. As the subject processes each stimulus, electrical brain activity is continuously recorded through dozens of small electrodes placed directly on a subject’s scalp. Although protocols differ markedly in the timing of stimulus presentation and the subject’s task demands, the fundamental concept for EEG-based memory detection relies on comparing the brain’s responses to personally “meaningful” versus “nonmeaningful” stimuli. In this context, a stimulus that is “meaningful” is one that would be salient or significant only to someone with prior knowledge or experience with that stimulus. A stimulus that is recognized by the subject—whether by through its relation to a past crime or merely by virtue of the fact that it was studied prior to the EEG session—should evoke a characteristic profile of electrical brain activity that can be detected by the experimenter, although often only after extensive trial averaging and analysis. Because the method relies on comparing brain activity across different classes of stimuli, the composition of the stimulus set (i.e., how the set of “nonmeaningful” comparison stimuli are chosen) is a critically important methodological factor in EEG-based memory detection and one that we return to below.

Most EEG-based memory detection studies rely on indexing the magnitude of a particular event-related potential (ERP) known as the P300. An ERP is a characteristic spike or dip in voltage activity that is evoked in response to a discrete event, such as the presentation of a stimulus, and can be detected at the scalp by amplifying the collective electrical potentials of large populations of activated neurons in the underlying cortex. Relative to ongoing brain activity, an ERP is small in magnitude. Thus, for a single event, the brain signal response to that event is difficult to discern from the noise of the ongoing brain activity. As such, to detect the ERP, researchers typically average the EEG samples of repeated presentations of the same stimulus or of several stimuli in the same category. The P300 is an ERP associated with a positive deflection of voltage over the parietal cortex occurring between 300 and 1,000 ms after the onset of events that are both infrequent in their occurrence and meaningful to the participant (Polich & Kok, 1995).

Researchers have been studying the use of the P300 ERP to detect "concealed information" since the late 1980s (Rosenfeld et al., 1988). Much of this work has utilized an experimental protocol known as the Concealed Information Test (CIT), formerly referred to as the "guilty knowledge test" (Lykken, 1959). In a typical CIT study, subjects are either asked to commit a mock crime (e.g., stealing and hiding away an object) or are provided with some privileged information (e.g., the identity of a secretly chosen playing card or details about a hypothetical crime or espionage scenario) and then are instructed to conceal their knowledge of these crime-relevant or privileged details. Many CIT studies have focused exclusively on measuring psychophysiological signatures of autonomic arousal—such as changes in skin conductance, respiration, and heart rate—in response to guilty knowledge probes. These efforts have typically yielded accuracies of 80–90% in their ability to discriminate "guilty" from "innocent" subjects (Gamer, 2011). However, CIT studies using EEG-based measures such as the P300 have reported even higher discrimination accuracies, often upwards of 85% and occasionally even achieving 100% correct classification (Ben-Shakhar, 2012; Farwell, 2012; Rosenfeld, 2019).

Generally, EEG-based memory detection in the CIT relies on comparison of brain responses between three categories of stimuli: (a) infrequent "probe" stimuli with information relevant to the event of interest that would only be recognized by someone with prior knowledge (i.e., meaningful to some people but meaningless to others); (b) infrequent "target" stimuli that subjects are explicitly instructed to monitor for
p. 2332 and respond to with a unique ↳ button-press—these are stimuli that the subject is expected to know, recognize, or be familiar with (meaningful to everyone); and (c) frequent "irrelevant" stimuli, of no relevance to the event of interest nor of any particular importance to the subject (meaningless to everyone). For all subjects, the target stimuli should elicit a greater P300 response amplitude than the irrelevant stimuli, providing an internal check that the procedure is working and the data are of sufficient quality. When used in a forensic context, the brains of "guilty" examinees should show a P300 to probe stimuli that is similar in magnitude to that observed for target stimuli, indicating that their brain "knows" the probe stimuli are not irrelevant. Conversely, the brains of "innocent" examinees should show little or no P300 response to the probes, such that this waveform will look more similar to that observed for irrelevant stimuli than for target stimuli (Meegan, 2008; Rosenfeld, 2019).

Although a simple visual comparison of the average P300 waveforms evoked in response to all probe, target, and irrelevant trials can sometimes be enough to support an educated guess as to whether the subject had concealed knowledge of the probe stimuli, it is important to have a quantitative measure of the statistical confidence of this determination that explicitly takes the across-trial variability into account. A bootstrapping analysis procedure, first introduced by Farwell and Donchin (1991), is typically used to formally evaluate the likelihood that the probe trials come from the same distribution as the irrelevant trials ("information absent") and the likelihood that probes come from the same distribution as targets ("information present"). This involves repeatedly sampling random subsets of the available trials and computing the amplitude differences between trial types, and then evaluating what percentage of the samples showed each effect of interest (e.g., probe > irrelevant). An experimenter-specified cutoff can then be applied to make a determination of the subject's status. For instance, if a subject's data do not show 90% or more probe > irrelevant P300 differences, the subject is classified as innocent. If desired, a separate threshold can be specified for determining guilt, and if neither criterion is met, the protocol yields an indeterminate response (Farwell, 2012; Rosenfeld, 2019).

Much of the recent research in EEG-based memory detection has focused on subtle changes in protocol design. Those that are particularly significant for purposes of assessing forensic utility are methodological developments focused on countermeasure vulnerability; issues of ecological validity; drawing statistical conclusions about single subjects and single trials; and, most recently, application of pattern recognition techniques to analyze EEG data. The highlights of those areas of research, and consideration of the outstanding limitations for forensic use, are discussed below.

Countermeasure Vulnerability

The original CIT test was determined to be vulnerable to countermeasures (Mertens & Allen, 2008; Rosenfeld et al., 2004). For instance, subjects could learn to adopt simple strategies that would diminish the difference in the P300 response to targets and irrelevants, making it more difficult for the analysis to yield conclusive evidence about whether the guilty knowledge probes were different from the irrelevants. One such strategy is to increase the salience of the irrelevants by performing a subtle imperceptible movement, such as wiggling one's finger or toe each time an irrelevant stimulus is presented. To make their procedure more resistant to these countermeasures, Rosenfeld and colleagues (2008) developed a new variant called the Complex Trial Protocol (CTP). Rather than requiring a yes/no response to all stimuli, each trial of the CTP first presents an irrelevant stimulus or probe stimulus without requiring any response and then presents a second stimulus (typically just a string of numbers such as 11111 or 22222) that requires a target/nontarget response (e.g., if 11111 is designated as the target, participants would press the "yes" button if this was presented and the "no" button if p. 2333 ↳ a different number string was presented). By temporally separating these two trial stages by just 1 or 2 s, the attentional demands of the task are increased in such a way as to enhance the magnitude of the automatically evoked P300 responses to probe stimuli and theoretically render countermeasures largely ineffective. Several subsequent studies from Rosenfeld's group confirmed that the sensitivity and specificity of the CTP-based memory detection approach remained above 90% even when participants were attempting to use countermeasures (for a review, see Rosenfeld, 2019).

Alternative Mental Actions as Countermeasures

Many experiments evaluating the efficacy of countermeasures have used physical actions as countermeasures, such as toe or finger wiggling during stimulus presentation (Ambach & Gamer, 2018; Rosenfeld et al., 2004). But countermeasures can be even more covert. If simple mental imagery were a successful countermeasure, it is even less likely to be detected by an observer than subtle physical responses. Since 2011, Rosenfeld's group has been studying the use of mental countermeasures such as silently "speaking" one's own name or one's family member's name to oneself in response to irrelevant stimuli (Sokolovsky et al., 2011). These types of countermeasures would be nearly impossible to detect through even the closest visual or mechanical observation of a test subject.

But does using mental countermeasures slow down task performance, such that countermeasure use could be detected by experimenters? In one mock crime application of the CTP, "guilty" subjects using countermeasures were found to display slower reaction times on both crime-relevant probes and irrelevant targets (Winograd & Rosenfeld, 2011). This telltale slowing seems to work only when subjects execute the covert countermeasure separately from the required "I saw it" motor response. If subjects are trained to execute both at the same time, this "lumping" strategy eliminates the ability to use reaction time as an indication of countermeasure use (Sokolovsky et al., 2011). Thus, there may be simple countermeasures (e.g., lumping) that participants can employ to evade countermeasure detection strategies. However, even when participants engaged in lumping, the P300 ERP signal could still detect more than 80% of the "guilty" test subjects (Rosenfeld et al., 2013), suggesting that even sophisticated countermeasures by highly trained subjects may have only limited efficacy.

Voluntary Suppression of Memory

Can someone beat a memory detection test by voluntarily suppressing their memories? There is some evidence that the P300 signal for episodic information can indeed be voluntarily suppressed (Bergström et al., 2013; Hu et al., 2015). However, a more recent paper reported that with a single discrete change to the CTP (changing the ratio of target to nontarget items to make the task slightly easier), participants' attempts to suppress their recognition of probe stimuli did not yield significant differences compared to a "simply guilty" control group (Ward & Rosenfeld, 2017). This suggests that the previously demonstrated reductions in P300 may not have been due to effective memory suppression but, rather, to task demand. Another recent paper investigated the extent to which people could suppress P300 signals of well-rehearsed *semantic* memories (e.g., the subject's own name) and reported that attempts to suppress responses to probes either enhanced or had no effect on the P300 signal (Rosenfeld et al., 2017).

Currently, the effect of countermeasures on well-constructed CTP applications appears to be modest and perhaps available only to highly trained subjects. Nevertheless, a complete understanding of potential
p. 2334 countermeasure vulnerability is critical to accurate forensic ↳ application of any P300-based memory detection protocol, especially research that can differentiate the effects of task demands from mechanisms of countermeasure deployment.

Ecological Validity

Ecological validity is the extent to which laboratory conditions and results translate to those in the real world. An example of a memory detection study with low ecological validity would be having a subject study a list of random words and then asking them to respond whether or not certain words were on that list, while attempting to conceal their actual memories for the study list items by making false responses. Studying a list of words is not a task that resembles many real-world events, let alone forensically relevant ones. Alternatively, a number of studies aiming for greater ecological validity have engaged their subjects in mock crime scenarios, such as "stealing" a particular item and then attempting to conceal their "crime" from the examiners. These types of studies began with the earliest efforts at EEG-based memory detection (Farwell & Donchin, 1991; Rosenfeld et al., 1988).

The Importance of Selecting Appropriate Stimuli

The subjective meaningfulness of any probe stimulus will inevitably depend on the context in which it is encountered, so care must be taken in selecting the other "irrelevant" items in the set to ensure that the probe will only be meaningful to a subject with specific memories of that item. For example, a gun encountered amidst a series of innocuous object stimuli might automatically evoke a P300 due to its salience and distinctiveness from the other items in the set, but a gun encountered amidst an array of different weapons (e.g., dagger, switchblade, crow bar, etc.) would not automatically evoke a P300 but, rather, would only evoke a P300 if it was meaningful to the subject. Of course, this could be because the subject is a gun enthusiast or because they think that this is the "guilty knowledge" stimulus that the investigators are interested in, even if they did not commit the crime. A more optimal use of the P300 for the forensic assessment of specific weapon knowledge might be to present a series of different guns, with only one of these being the gun that was actually found at the crime scene or known to be the murder weapon. It is important to have a sufficient number of unique irrelevant stimuli (typically 10 or more) because this increases probe rareness, which is linearly related to P300 amplitude (Rosenfeld & Greely, 2009).

Likewise, it is important to consider the possibility that many of the putatively "irrelevant" stimuli could be personally meaningful to the subject by virtue of evoking episodic memories from an event that is different from the one that the investigators are aiming to probe or evoking personal autobiographical semantic memories (e.g., seeing a necklace reminiscent of one that your mother wore often in childhood). Consider the scenario in which the probe is a valuable stolen item, such as a ring, and the irrelevants are similarly valuable jewelry items, such as a watch, a necklace, or a bracelet. Perhaps the suspect—a jewelry thief or house burglar—has committed more than one crime, as yet unknown to investigators. This could cause some of the probes designated by the experimenter as "irrelevants" to be processed as relevant (personally meaningful) to the suspect, albeit relevant to a different episode(s) of the suspect's past. The result would be a decreased amplitude of the P300 for the supposed "probe" item. Conversely, category-based selection of the irrelevant items could lead to false-positive results for innocent suspects. For instance, if investigators know that a victim was bludgeoned with a baseball bat, and they put the probe "bat" in a lineup of other "weapons" that can be wielded with blunt force (e.g., a tire iron, metal candlestick, and brick), the bat could function as a salient probe for an innocent subject by virtue of being the only wooden p. 2335 object, the only ↳ piece of sports equipment, or because the innocent subject played baseball as a child. This methodological limitation should caution against forensic use of a P300 procedure in any early stage investigation, before sufficient certainty as to key events, objects, and people has been established via other means.

Use of Real-World Scenarios

A fundamental question for forensic application of memory detection technology is the extent to which lab-created memories—even those with more realistic mnemonic content than a memorized word list, such as a mock crime—are similar to or meaningfully different from real-world memories. There are many reasons such memories could be different, including the amount of attention a subject is allocating to lab-based, instructed tasks such as stealing a document versus going about normal routines of the day when something unexpected happens. A recent study sought to address the detectability of incidentally acquired, real-world memories. Meixner and Rosenfeld (2014) had subjects wear a small video camera on Day 1 and go about their routine as the camera recorded their lives. On Day 2, subjects returned to the lab and, while hooked up to the EEG, were shown words or phrases associated with activities they experienced the previous day that should have been well attended. They were also presented with irrelevant words of the same general semantic category but not relating to their personal activities from the prior day (e.g., if the autobiographical memory probe was "grocery store," the irrelevants could include "movie theater" or "mall"). The authors reported perfect discrimination between 12 "knowledgeable" subjects who viewed relevant words related to their personal activities and 12 "non-knowledgeable" subjects who simply viewed irrelevant items. Notably, this was the first P300 study to examine whether real-world autobiographical memories could be detected at an individual subject level; prior studies of autobiographical memory have used group averaged data, which is not helpful for forensic purposes. But the study design did not get at the question of whether lab-created memories are different, at some important neural level, from real-world memories—it simply suggests that this particular tool can be used to discriminate between subjects who had or did not have very discrete real-world experiences.

Innocent-But-Informed Participants

Other work focused on assessing the ecological validity of a P300-based CIT demonstrated that prior knowledge of crime details has an effect on detection rates, illustrating potential risks of using probes that may have become known to an innocent subject, such as someone who received instructions to steal a "ring" but did not go through with the crime, compared to a "naive" subject told only to steal an "item" (Winograd & Rosenfeld, 2014). These "innocent informed" participants were essentially indistinguishable from those who committed the mock crime. This suggests that when testing a person who did not participate in a criminal act but knows key information about it, there is a high chance of a false positive. Rosenfeld and colleagues counsel that crime details must remain secret, known only to police, perpetrators, eyewitnesses, and victims. This, however, may be easier said than done because it is not always possible to know the degree to which critical details about the event in question have been inadvertently disclosed. Generally, the studies investigating variables reviewed in this section have not distinguished between witnesses (who may be innocent but informed) and participants ("guilty" subjects), such that variables such as time delays and quality of encoding are poorly understood in applications of detection of witness (rather than suspect) memory.

p. 2336
Time Delay Between Event and Test

In lab studies using mock crimes, participants are often tested immediately or a few days after the incident. In the real world, interrogation about an event may occur weeks, months, or even years later. One study attempting to quantify the impact of a delay in testing asked students to "steal" an exam from a professor's mailbox (Hu et al., 2012). Some students were tested using a P300-based CTP procedure immediately after the theft, whereas others returned to the lab a month later. Researchers found no difference in detection efficiency. This is an encouraging result, but more evidence is needed. The test only used a single probe item (the stolen exam), and—as with other mock crime scenarios—subjects were instructed to engage in the theft, likely heightening the salience of the central detail. What is currently unknown is how P300-based detection fares over time for peripheral crime details, which may be less robustly encoded but more likely to be uniquely known only to a perpetrator, and thus useful to minimize a false-positive result.

Quality of Encoding

Not all information is equally well-remembered, and how well information was initially encoded has substantial implications for how well it may later be detected. The ability to detect event details that were incidentally encoded (i.e., stored into memory without any explicit intention to memorize these details) is critical for potential forensic applications of memory detection. Yet behavioral tests of people's memory for incidental details of real-world experiences show that sometimes surprisingly little is retained (Misra et al., 2018). Rosenfeld's research group acknowledges that detection sensitivities in their P300-based tests are lower with incidentally acquired information than with well-rehearsed autobiographical information, such as the subject's own name (Rosenfeld et al., 2006). Strategies to improve sensitivity of detection of incidentally acquired information under investigation include providing feedback to more intensely focus the subject's attention on the probe (e.g., by telling them that "based on your brainwaves, it seems that there is a certain stimulus that is important to you"), using an additional ERP component called the N200 that emerges approximately 100 ms earlier than the P300, and combining separately administered behavioral testing with the CTP (Hu et al., 2013).

Existing research has focused on how use of countermeasures by a guilty subject may lead to missed detection (i.e., a false negative) because of countermeasures. What has not been firmly established is how often a guilty subject—or a witness—will not show a P300 response to a probe stimulus for another reason, such as the fact that they may not have encoded the particular details of a weapon used, because of intoxication, darkness, impulsivity, or stress. For instance, consider a home break-in scenario in which the perpetrator hastily snatched and bagged any items of perceived value without having closely examined them. If the suspect was later apprehended and asked to perform a P300-based assessment to establish guilty knowledge of the stolen items, it is entirely possible some of the items would elicit no neural signature of concealed knowledge because a strong memory for their appearance was never formed. A similar issue could occur if a truly innocent suspect was subjected to P300 assessment for confirmation of details pertaining to their alibi. If the investigators were to visit the alibi location and take photos to use as probe stimuli, it is entirely possible that these images would fail to elicit a P300 response if the details of the alibi location had never been robustly encoded by the suspect. And this in turn could be wrongly interpreted by investigators as evidence that the suspect's alibi was not supported by neural interrogation of their memory. In addition to concerns about poor initial encoding of event details, natural forgetting processes—p. 2337 whether attributable to interference (Wixted, 2005; see also Chapter 40) or decay ↳ (Hardt et al., 2013)—may also lead to diminished detectability of a memory. This illustrates the challenges inherent in using the absence of a P300 response as evidence of innocence or guilt.

Pattern Recognition Analysis Techniques

Recent work has attempted to use multivariate pattern recognition techniques to analyze EEG data in tests of concealed information. This technique takes a fundamentally different approach to diagnostic classification. Instead of simply evaluating whether the magnitude of an individual subject's P300 response to probe stimuli is more comparable to that of target or irrelevant stimuli, pattern recognition methods involve training a computerized classifier algorithm on spectrotemporal features of the single-trial EEG data. In one demonstration of this approach, Mehrnam and colleagues (2017) collected EEG recordings as participants performed a CIT, in which half the participants were instructed to conceal their "guilty knowledge" of a specific probe face. Using a leave-one-subject-out cross-validation scheme, a classifier model was trained on the aggregated single-trial EEG data from all but one subject, and then it was applied to the single-trial EEG data from the held-out subject to predict the recognition status of each item. Using an adaptive thresholding technique that took into account the relative likelihood that target versus irrelevant stimuli would elicit a P300 response in each subject, the algorithm sought to determine if at least half of the probe trials for each subject showed a target-like activity pattern, in which case the subject would be labeled as guilty. The results showed that the algorithm could discriminate guilty from innocent participants nearly 92% of the time. Unfortunately, the researchers did not report decoding accuracy levels for individual trials but, rather, only used the single-trial classifier outputs to arrive at a determination of each subject's guilt. Moreover, because the experiment only used six face stimuli per participant (one probe, one target, and four irrelevant), the data are not able to speak to the ability of the method to determine which stimuli, out of a large set of possibilities, are recognized by the subject. To address this, another recent EEG study using a pattern classification approach found that the memory retrieval outcomes of single trials could be classified with above-chance accuracy, yet this accuracy level was quite modest (only 56–61% correct, where 50% would be chance-level classification) (Noh et al., 2018). Because the application of pattern recognition classifiers to EEG-based memory detection is relatively new, more studies will be needed to evaluate its potential advantages over prior P300 analysis methods, as well as its susceptibility to countermeasures.

The Need for Broader Research Efforts and Independent Replication Attempts

One potential limitation of EEG-based memory detection is that work on this topic has been largely dominated by two research groups: one academic (the laboratory of Peter Rosenfeld at Northwestern University) and one industry-based (Lawrence Farwell, who founded and runs an independent company called Brain Fingerprinting Laboratories, Inc.). Much of our discussion thus far has focused on the work of Rosenfeld and collaborators. To our knowledge, only one independent research group has sought to test the use of Rosenfeld's CTP for memory detection, and it largely replicated the core P300 findings but found some vulnerability to countermeasures (Lukács et al., 2016).

Farwell and colleagues have published extensively on the use of EEG-based methods for memory detection, with much of their work relying on a slight variant of the standard P300 measure that they refer to as the P300-MERMER (memory and encoding related multifaceted electroencephalographic response; Farwell & Smith, 2001) and which they have found to be somewhat more accurate than analyses based on the P300 alone (Farwell et al., 2013). Farwell p. 2338 ↳ has claimed that his proprietary "brain fingerprinting" methods have consistently yielded highly accurate classifications (Farwell, 2012). Indeed, in all of his published work, there has never been a report of a single guilty knowledge participant classified as innocent nor a naive participant incorrectly classified as having guilty knowledge. On rare occasions, his procedure will yield an indeterminate response indicating that the evidence for concealed knowledge of probe stimuli was ambiguous, but Farwell claims that indeterminate outcomes have only ever been produced when the P300 is analyzed alone and not with the P300-MERMER method. If such results seem too good to be true, it is worth noting that Farwell's work has come under serious critique. Specifically, the proprietary P300-MERMER data analysis and scoring methodology is criticized as being too inadequately described in published work to allow for independent replication (see Rosenfeld, 2005). Farwell's detractors also note a paucity of peer-reviewed data and selection bias in results reporting (Meijer et al., 2013). Without more transparent sharing of experimental protocols and open source analysis software, it will be challenging to resolve some of the disputes in the EEG memory detection literature and get a clearer sense of the true accuracy of these methods and their resistance to countermeasures.

Efforts to Use EEG-Based Memory Detection in Legal Contexts

Although the use of neuroscientific evidence in the US criminal justice system is sharply on the rise (Farahany, 2015; Greely & Farahany, 2019), there has been only one instance to date in which a judge formally considered evidence derived from functional neuroimaging pertaining to a defendant's memory for details of the crime scene. In that case (*Harrington v. Iowa*, 2001), the defendant's counsel, as part of a post-conviction relief action, sought to introduce evidence from Farwell's brain fingerprinting procedure to suggest that the defendant lacked any discernible memory for a number of important features of the murder that he had been found guilty of committing 23 years ago. A second test was proffered as evidence that the defendant did recognize salient details of his alibi. An Iowa district court judge held an evidentiary hearing and ruled that the brain fingerprinting evidence could be admitted. However, given that a key eyewitness to the murder also recanted his testimony, the appellate court handling the case did not end up incorporating the brain fingerprinting evidence in its decision to free the defendant (for discussion, see Rosenfeld, 2005).

Outside of the United States, an EEG-based procedure developed in India called brain electrical oscillation signature (BEOS) profiling has been used as a forensic investigatory tool in hundreds of cases. Its existence caught the attention of the international community in 2008 when a judge admitted BEOS evidence into a murder trial and seemed to rely heavily on this evidence when issuing his guilty verdict (Giridharadas, 2008), although this ruling was overturned by a higher court soon thereafter. Although the methods that underlie the BEOS procedure have not been subjected to rigorous peer review, from what is known (Mukundan et al., 2017), the technique involves presenting the subject with a series of short sentences, each

referring to crime-relevant episodic details, personally familiar autobiographical episodic details, or neutral information. The subject merely listens to each sentence without making any response, and the analysis of the EEG data searches for a series of spatiotemporal signatures of experiential knowledge that need to emerge in a specific temporal sequence. The automated analysis is purported to be calibrated to be conservative, and it is argued that experiential knowledge with a probe stimulus will rarely ever be indicated for neutral probes or guilty knowledge probes in subjects known to be innocent. Of course, without knowing the precise methods and the conditions under which the BEOS procedure has been tested, it is not possible for the international research community to critically evaluate these claims. At least one other country—
p. 2339 Singapore—has licensed the BEOS technology for use in its own ↳ forensic investigations, and its use in India appears to be ongoing, albeit as a tool for forensic assessment rather than as legal evidence of guilt.

fMRI-Based Approaches

Several attributes distinguish fMRI-based memory detection from EEG-based memory detection. fMRI can provide data from the entire brain and can do so with a spatial resolution that is far higher than EEG. This allows researchers to more precisely localize the source of the signals they are measuring. Also, whereas scalp EEG measurements are limited largely to cortical activity, fMRI is capable of measuring activity levels in deeper subcortical regions, such as the hippocampus. A major disadvantage of fMRI is that its temporal resolution (i.e., its ability to resolve neural events in time) is far worse than that of EEG. This is because fMRI relies on measuring the blood oxygenation level–dependent signal, an indirect correlate of neural activity that peaks 4–6 s after an area becomes activated. But unlike EEG measures, which tend to be quite noisy at the level of individual trials, single-trial fMRI measures can be surprisingly stable. Given that fMRI can measure the spatial profile of brain activity with such precision, the technique has proven to be highly amenable to the application of multivariate pattern classification algorithms that exploit the spatially distributed nature of memory processes and representations (Rissman & Wagner, 2012).

Early fMRI Work on True and False Memories and Concealed Information Tests

A number of fMRI studies have sought to compare the brain activity evoked during memory retrieval trials in which participants claim to remember a stimulus that they actually did study (true memories) versus remembering a stimulus that they did not study (false memories) (see Chapter 46). A review of this literature identified as a ubiquitous finding the considerable overlap in the neural networks mediating both true and false memories for recognition responses to items that share the same gist (i.e., general conceptual meaning) as items that were actually studied during encoding (Dennis et al., 2015). Many studies also find differences between true and false memories—notably, increased activity in regions related to sensory processing for true compared to false retrieval, supporting the sensory reactivation hypothesis that true memories are associated with the bringing back to mind of more sensory and perceptual details than false memories (Slotnick & Schacter, 2004). Overall, empirical support for the intuitively appealing sensory reactivation hypothesis is mixed, and other dissociations, such as different patterns of activity in the medial temporal lobe and prefrontal cortices, have been reported (for a review, see Dennis et al., 2015).

Other work testing episodic memory retrieval in an fMRI scanner 1 week after viewing a narrative documentary movie reported that coactivations of the left medial temporal lobe regions along with temporal and parietal cortices (indicative of coordinated functional communication between these regions) were greater when subjects responded correctly versus incorrectly to factually accurate statements about the movie (Mendelsohn et al., 2010). However, these coactivations did not differ between trials when participants incorrectly endorsed fictitious statements about the movie versus when they correctly rejected these fictitious statements, calling into question whether these coactivations can be taken as a proxy for memory veracity. Collectively, this work provided some suggestion that activation patterns could differentiate between true and false recognitions, based on distinct memory processes, although no clear potential for diagnostic assessment of true or false memories emerged. In other words, there is no particular "spot" in the brain that serves as a litmus test for whether a memory is true or false.

Newer fMRI Methods: Advanced Techniques Reveal the Biological Limitations of Memory Detection

A major limitation of earlier fMRI analysis paradigms is that they involve comparing activity levels across task conditions (e.g., true vs. false memories, guilty knowledge probes vs. irrelevants) in an effort to identify focal brain regions that show reliable differences. However, memories are encoded and stored in a highly distributed manner across networks of brain regions (Rissman & Wagner, 2012), and thus searching for localized areas whose activity can be interpreted as a signature of recognition may be a misguided effort. But newer fMRI analysis methods can use machine learning algorithms to make better use of the massive amounts of data acquired in each brain image, allowing the characterization of subtle patterns of activity within and across regions on individual trials rather than activation in local areas averaged across many trials. Such classifier-based methods were discussed briefly above as the pattern recognition techniques applied to EEG data.

Multivariate pattern analysis is a substantially more powerful way to analyze complex data, and the remainder of this review focuses on fMRI studies that employ such methods. These methods offer the best hope for reliable forensic memory detection, the greatest insights into the biological substrates for and limitations of memory detection, and the subtlest challenges for a fact finder to assess the credibility of the techniques used to make claims about the presence or absence of a memory of interest.

Multivoxel Pattern Analysis and Machine Learning Classifiers

Classic fMRI analysis examines effects within discrete "chunks" of the brain: clusters of voxels and regions of interest. By analyzing each voxel separately (and then pooling the outcomes across the voxels in a cluster), this univariate analysis approach ignores the rich information that is encoded in the spatial topography of distributed activation patterns—that is, it can highlight particular areas showing significantly increased or decreased levels of activity, but it misses patterns in the broader "landscape" of brain activity. A newer multivariate technique, multivoxel pattern analysis (MVPA), tries to exploit the information that is represented in the distributed patterns throughout a brain region or even across the entire brain. It is a more sensitive method of analysis because it is more adept at detecting distributed representations of information that may be missed by classic univariate (i.e., voxel-by-voxel) fMRI analysis approaches. For example, there is not any one location of the brain that shows significantly greater activity when a subject perceives the letter X versus the letter O, and yet distributed fMRI activity patterns within visual brain areas are capable of not only being used to decode which letter a participant is perceiving on a given trial but also which of the two letters a participant is imagining (Stokes et al., 2009). As with pattern recognition analysis in the EEG studies discussed above, MVPA techniques typically use machine learning algorithms to train classifiers on data patterns from one or more subjects, and these classifiers learn the distributed neural signatures that differentiate unique mental states or behavioral conditions. Then the classifier is tested on new data (that it has not been trained on) to determine whether it can accurately classify the condition of a subject's brain on a given trial based solely on brain data information (Haynes, 2015; Lewis-Peacock & Norman, 2013). Simply stated, there is more informational content in fMRI activity *patterns* than is typically detected with conventional univariate fMRI analyses. The accuracy with which the classifier can discriminate trials from Condition A and Condition B gives a quantitative assessment of how reliably two putatively distinct mental states are differentiated by their brain activity patterns.

MVPA has enabled significant advances in memory detection research. In 2010, the first paper to apply an
p. 2341 MVPA approach to memory detection in fMRI evaluated whether ↳ individuals' subjective memory experiences, as well as their veridical experiential history, can be decoded from distributed fMRI activity patterns evoked in response to individual stimuli (Rissman et al., 2010). Prior to entering the scanner, participants studied 200 faces for 4 s each. In the scanner approximately 1 hr later, they were presented with the 200 studied faces interspersed with 200 unstudied faces, and they pressed a button to indicate whether or not they recognized a given face. Participants' judgments were accurate approximately 70% of the time, giving researchers the ability to examine their brain activity both when they were correct and when their memory failed them. Participants also had to indicate how confident they were about their response and whether they merely found the face to be familiar or actually *recollected* the face (i.e., whether their memory for the face was accompanied by bringing back to mind contextual details associated with the initial encounter, such as recalling a particular thought they had had about that particular face). The classifier was first trained to differentiate brain patterns responding to old faces that subjects correctly recognized from brain patterns responding to new faces that they correctly judged to be novel—both situations in which the subjective experience and objective reality were identical. The classifier performed well above chance at decoding the old/new status of individual trials, with a mean classification accuracy of 83%, rising to 95% if only the classifier's "most confident" guesses were considered. But because subjects did not perform perfectly—sometimes misidentifying new faces as having been previously seen or rejecting old faces as novel—the classifier could also be tested on the subjective mnemonic experience. In those scenarios, the classifier performed relatively poorly when applied to detect the true experiential history of a given stimulus on trials for which participants made memory errors. When tested on its ability to distinguish true recognition from false recognition (controlling for participants' confidence in their judgments), the classifier could succeed only 59% of the time. The classifier fared even worse—indeed no better than chance-level guessing—when tested on its ability to determine which of the faces that participants rated as unfamiliar had actually been studied. In other words, although the classifier proved to be very good at decoding a participant's *subjective* memory state, it was not nearly as good at detecting the true, veridical, *objective* experiential history of a given stimulus. Classification performance also dropped to chance levels in a second experiment in which participants encountered old and new faces but were asked merely to judge the male/female gender of each face rather than make explicit memory judgments. This indicates that when the experiential history of the stimuli was not relevant to one's behavioral goals, the neural distinction between recognition and novelty was not pronounced in the observer's brain. From a forensic standpoint, such data suggest that fMRI memory detection might only be useful insofar as there is value in confirming a person's explicit and subjective recognition experience. But of course, most foreseeable forensic applications would be predominantly concerned not with whether a person claims to recognize a given stimulus but, rather, whether that recognition is objectively true or false.

Three other aspects of this study deserve separate mention because they are particularly relevant to assessing the forensic capabilities and limitations of MVPA-based fMRI memory detection. First, although many stimuli were shown to subjects in order to train the classifier, once trained, it could be applied to single trials—that is, a face shown just one time provided sufficient neural information for the classifier to make a decision about whether or not it was recognized. This is a significant advantage over EEG-based memory detection paradigms that require multiple presentations to detect an ERP, as well as univariate fMRI analyses that assess averages of trials across conditions. A subsequent study confirmed that an MVPA classifier could reliably decode the memory status of individual retrieval trials, and it further attempted to
p. 2342 assess the vulnerability of this single-trial assessment to countermeasures by instructing ↳ participants who had studied a set of faces to attempt to conceal their true memory state while in the scanner (Uncapher et al., 2015). During some of the scanning runs, participants were instructed to feign the subjective experience of novelty for any faces they actually recognized (e.g., by focusing on superficial attributes of the photographic stimulus that they may not have noticed the first time they saw it) and to feign the experience of recognition for any faces they did not recognize (e.g., by recalling someone whom the novel face

reminded them of). Using this easy-to-implement countermeasure strategy, most participants were able to prevent the experimenters from accurately differentiating brain responses to studied versus novel faces, with the mean classification accuracy declining to chance level. That is, although MVPA classification of fMRI data can enable single-trial memory assessment, it may still be vulnerable to simple mental countermeasures.

Second, in the study by Rissman et al. (2010), the researchers specifically considered whether the same classifier would work across subjects by training it on data from some individuals but then testing it on data from others. This would be an important feature of any technology to be forensically applied, because it would enable development and use of a normative classifier that is then applied to individuals of interest. However, this feature depends on an assumption that there is some unknown degree of consistency in memory-related neural patterns across different people. Consistent with this possibility, the classifier worked well across individuals, such that cross-participant classification performance levels were similar to those of the corresponding within-participant analyses.

Third, recall that the “classifier” is not a person making a judgment call; it is a machine learning algorithm—a regularized logistic regression in Rissman et al. (2010)—examining patterns of neural data. This study suggests that memories that feel true but are objectively false may often elicit indistinguishable neural patterns from memories that feel true and are veridical. If that finding holds, it has substantial implications for the ability to use fMRI to detect true memories and to discriminate them from false memories.

Advances in Experimental Paradigms with MVPA Analysis: Real-World Life Tracking, Single-Trial Memory Detection, and Boundary Conditions Revealed

Lab-created memories such as studying a series of faces may be relatively impoverished compared to real-life memories that have potential for context and meaning. It is possible that where MVPA-based fMRI detection met limits for lab-created memories (in particular, the inability to distinguish objectively false but subjectively recognized stimuli and also vulnerability to simple countermeasures), additional information available to the classifier from a richer memory set may make real-life memories more distinguishable. This enrichment of the memory experience, and delays between the experience and time of test, also addresses concerns about ecological validity of research for potential forensic application. Wearable cameras, as mentioned previously in regard to the Meixner and Rosenfeld (2014) study, offer one promising means to automatically capture a first-person record of people's experiences in naturalistic settings and later use the resulting images or videos to probe autobiographical memory (for a review, see Chow & Rissman, 2017). In an effort to evaluate fMRI-based memory detection for real-world experiences, Rissman et al. (2016) had participants wear a necklace-mounted digital camera for a 3-week period while going about their daily lives. Participants returned approximately 1 week later to be scanned while making memory judgments about sequences of photos from their own life or from others' lives. Each four-photograph sequence depicted the temporal unfolding of a single potentially memorable episode (captured within a 5-min interval). After viewing each sequence, participants made a self/other judgment and p. 2343 indicated how strong their experience of recollection or familiarity was for photos judged to be from ↳ their own life or their degree of certainty about photos judged to be from someone else's life. Behaviorally, participants were extremely good at this task, with approximately 80% of the self/other judgments being correct, and the remaining 20% split between incorrect or unsure. Indeed, participants performed so well that there were too few false memories to permit assessment of whether the classifier could distinguish true memories (i.e., photos from one's own camera that were recognized) from objectively false memories (i.e., photos from another's camera that were mistakenly recognized as one's own). This is perhaps a side effect of the fact that real-life memories are richer than lab-created memories, such as pictures of faces or words, and perhaps less susceptible to spontaneous false-memory effects. Importantly, however, the classifier performed extremely well at distinguishing own life events that were recognized (i.e., hits) from others' life events that were correctly perceived as novel (i.e., correct rejections), being accurate 91% of the time on average (with no subject's classification performance below 80%). When the classifier was required to distinguish between trials in which participants reported recalling details and those in which they reported only familiarity for the event, it still performed well above chance, with mean accuracy of 72%. One feature of regularized logistic regression machine learning classifiers is the ability to create "importance maps," which permit researchers to assess which voxels (and thus networks of brain areas) are important to the classifier's decision. Based on these importance maps, self/other classifier distinctions relied on brain areas associated with mnemonic evidence accumulation and decision processes, whereas recollection/familiarity distinctions showed a very different pattern, involving brain regions associated with retrieval of contextual details about an event.

The Rissman et al. (2016) study design also addressed the issue of retention interval—that is, the amount of time that elapses between the formation of a memory and the probing of that memory in the MRI scanner. Here, the classifier performed just as well when tested on memories for events that occurred 3 weeks before the scan, at the beginning of the time participants started wearing the cameras, as on memories for events 1 week before the scan. This suggests that over a 1-month time period, recent memories are not necessarily easier to decode than remote memories.

As in Rissman et al. (2010), Rissman et al. (2016) also confirmed that a classifier trained on the neural data from some individuals performs well when tested on neural data from other participants, suggesting that the underlying brain activity patterns are fairly consistent across subjects. Moreover, the researchers directly addressed the question of whether lab-based memories are different from autobiographical memories by using the classifier trained on the data from the 2010 face memory study and tested on brain data from the self/other life photograph study, and vice versa. In both situations, the classifier still succeeded at predicting a subject's mnemonic judgment over two-thirds of the time, well above chance (for related work examining autobiographical episodic vs. autobiographical semantic memory decisions about photos from one's camera, see Brown et al., 2018).

Because participants in the Rissman et al. (2010, 2016) studies were only asked to make honest memory judgments and had no incentives to deceive the experimenters by concealing their recognition, those studies were not designed to address whether MVPA-based analyses could be used to ascertain the presence or absence of guilty knowledge. As mentioned above, the Uncapher et al. (2015) study attempted to address this question by instructing participants to conceal their recognition using a simple countermeasure strategy, and under these conditions mean old/new classification performance declined to chance levels. But that study involved concealing one's recognition of faces that had only been encountered twice before (each time for only 2 s). Thus, the probed memories were quite weak and lacked rich contextual associations, which likely made them easier to conceal.

p. 2344 To investigate the ability of MVPA to decode fMRI activity patterns associated with guilty knowledge of richer experiential memories, a study by Peth and colleagues (2015) used a mock crime scenario similar to the EEG-based CIT. In this experiment, one group of subjects (guilty intention) planned a realistic mock crime (a theft of money and a compact disc with important study information) but did not actually commit it. Another group (guilty action) planned and executed the "crime." And a third group (informed innocent) was informed of half of the relevant details in a neutral context but did not engage in any planning intention. Subjects were scanned while performing a CIT behavioral task, which included probes that would be unknown to participants in all three group. The MVPA analyses showed that although it was possible to reliably determine whether or not individual subjects possessed knowledge of crime-relevant details, the classifier was far less accurate (and indeed not significantly better than chance) at discriminating between the subjects in the three groups. In other words, the classifier could not determine—at least not consistently—whether the presence of crime-related memories had been obtained by way of crime execution, crime planning, or merely reading about the crime-relevant details. Thus, much like the comparable EEG study discussed above (Winograd & Rosenfeld, 2014), even the informational richness of whole-brain fMRI brain activity patterns may be insufficient to prevent the risk of false-positive identifications of innocent-but-informed individuals.

Of course, researching real-life memories in a laboratory setting is methodologically complex. What about the fact that looking at pictures is itself an autobiographical experience? How do "laboratory memories" really diverge from "real-world memories," and can a classifier tell the difference? In an effort to address these questions, an fMRI study by Chow et al. (2018) examined whether there is a detectable neural difference between the experience of viewing a photograph and the experience of actually living a particular experience. Like their earlier study (Rissman et al., 2016), the experiment used necklace-mounted cameras to capture participants' life events during a 3-week period, but this time the researchers permitting participants to preview a portion of the photographs—some from their own life and some from another participant's life—1 day prior to scanning. The critical question was whether distributed fMRI activity patterns within two putatively distinct brain networks—identified via meta-analysis (McDermott et al., 2009) as preferentially associated with the retrieval of autobiographical memories and of laboratory-based memories—would show differential sensitivity to the source of the event photographs (i.e., whether or not they were from one's life) and their pre-exposure status (i.e., whether or not the photographs had been previewed the day before the scan). This is indeed what was found; the classification analyses revealed a dissociation between the diagnosticity of each of two different brain networks. Specifically, activity patterns within the "autobiographical memory network" (involving bilateral regions of the medial temporal lobe, medial prefrontal cortex, posterior cingulate/retrosplenial cortex, and left angular gyrus) were significantly more diagnostic than those within the “laboratory-based network” (involving largely left-lateralized regions of the middle frontal and inferior frontal gyri, along with the posterior lateral parietal cortex and precuneus) as to whether photographs depicted one’s own personal experience versus another’s experience, regardless of whether they had been viewed prior to scanning. In contrast, activity patterns within the laboratory-based memory network were significantly more diagnostic than those within the autobiographical memory network as to whether photographs had been previewed, regardless of whether they were from the participant’s own life. This dissociability provides some evidence for separate neural processes for retrieval of firsthand experience versus secondhand knowledge—a finding that has significant implications for how, in a forensic context, stimuli are selected and whether or not they can or should be previewed to subjects.

p. 2345

So Where Is Brain-Based Memory Detection Now?

The most advanced brain-based memory detection work leverages algorithmic classification of distributed patterns of EEG or fMRI activity to identify neural signatures of recognition on single trials. As additional progress is made in computational approaches to the analysis of EEG (Hubbard et al., 2019; King & Dehaene, 2014) and fMRI (Cohen et al., 2017; Wang et al., 2020) activity patterns, it is possible that brain-based memory detection accuracy will further improve. Moreover, combining sophisticated data analytic techniques with experimental paradigms aimed at capturing real-world expressions of memory—such as wearable camera studies (Chow & Rissman, 2017)—will continue to advance our understanding of the conditions under which memory detection is most likely to succeed or fail.

However, the research reviewed above has revealed a number of salient limitations. Even with cutting-edge technology able to detect distinctions in different types of autobiographical or episodic retrieval processes, there may be no way for neuroimaging data to reliably (a) distinguish between a false but subjectively believed memory and an objectively true memory; (b) detect, on a single-trial level, the deployment of simple mental countermeasures to conceal or feign recognition; and (c) distinguish between someone who participated in a particular event and one who has knowledge of, or even intention to, but did not participate in a particular event. These limitations may ultimately have more to do with the constructive nature of human memory rather than with inadequacies in the fidelity of neuroimaging data or the power of the analysis tools. If cognitive neuroscientists are correct that the act of retrieval involves accessing fragments of a stored memory and filling in any missing elements with schematically plausible details (e.g., Barry & Maguire, 2019; Schacter & Addis, 2007), this is likely to be a persistent problem for potential forensic applications of brain-based memory detection.

Some recent progress has been made in using fMRI activity patterns to decode, or even to reconstruct a rudimentary representation of, the actual contents of what is being retrieved (e.g., Lee & Kuhl, 2016; Thakral et al., 2017; Wimmer et al., 2020; see also Chapters 37 and 38). In most cases, this content decoding is limited to determining whether a given stimulus is triggering the recall of associated information of a specific category, such as a face or scene or object, without being able to specify what specific item is being recalled. But representational similarity analysis approaches have yielded some success at indexing the recall of item-level representations (e.g., Wing et al., 2015). However, these methods do not yet have sufficient reliability to be able to decode which specific item a participant is bringing back to mind on an individual trial. Even if some day the technology were to advance to the point at which it was possible to get a readout of the specific details that are coming back to a person’s mind in response to a given memory probe, it would still be questionable whether this would have much utility in forensic contexts because the issues of false remembering and deliberately manipulated retrieval would still remain.

Conclusion

Experimental efforts to detect neural signatures of memory have demonstrated that some aspects of recognition can be inferred from functional neuroimaging data with surprisingly high reliability. The results of EEG-based studies have perhaps achieved the greatest success at detecting the presence of “guilty knowledge” for probe stimuli, even when participants are ostensibly trying their best to conceal this knowledge. There is also reason to believe that some EEG-based approaches, such as those using Rosenfeld’s CTP, might be quite robust to participants’ attempted use of countermeasures. More work, ideally from independent laboratories, will be needed to confirm these effects and examine how well they
p. 2346 extend to real-world event ↳ memories acquired more than 1 day prior to EEG testing. Furthermore, most of the EEG memory detection work has been limited to binning data across categories of stimuli (e.g., targets, irrelevant, and probes) to yield an overall assessment of whether a given participant harbors knowledge of the probes. Only a few recent studies have attempted to decode the recognition status of individual stimuli from EEG data, and the results have been mixed, in some cases yielding decoding performance barely above chance-level guessing. Unless these issues can be addressed in future work, EEG-based memory detection may ultimately be restricted to the labeling of broad classes of stimuli or the differentiation of subject groups, which would seriously constrain its forensic utility.

In contrast, fMRI-based measures have proven to be more capable of facilitating the detection of memories on individual retrieval trials. When analyzed with multivariate pattern classification techniques, fMRI activity patterns appear to be highly diagnostic of whether a participant is experiencing a subjective sense of recognition or novelty in response to a given probe stimulus. Furthermore, classifiers can be trained to differentiate recollection-based retrieval from familiarity-based recall, as well as the subjective strength of one’s sense of recollection, familiarity, or novelty. This has been shown for both laboratory-encoded stimuli, such as faces, and real-world stimuli, such as photographs of first-person experiences captured by wearable cameras. The effects appear to be stable even for memories that are up to 1 month old. However, few fMRI studies have attempted to classify true versus false memories on individual trials, and those that have typically find this to be much more challenging than classifying subject memory states. The work conducted to date appears to suggest that the large-scale fMRI activity patterns that can be decoded on individual retrieval trials are heavily dominated by brain processes associated with the subjective experience of remembering (and potentially also the decision processes involved in performing an explicit memory task). Moreover, these brain activity patterns can be willfully manipulated by participants who deploy simple countermeasures, so as to minimize the difference between recognized and novel stimuli. Brain activity patterns associated with objective recognition, when this is experimentally decoupled from subjective experience, appear to be considerably weaker, if not sometimes impossible to decode. This severely limits the practical value that fMRI-based memory detection could have in the legal system, as the objective truth of a witness or suspect’s experiential history is typically what the jury is after. Of course, eyewitness testimony is often riddled with inaccuracies of its own, but it is unlikely that fMRI will ever swoop in to save the day and provide a veridical readout of what a person experienced in the past.

Both EEG- and fMRI-based memory detection approaches are plagued by the additional challenge of needing the experimenters to construct appropriate stimulus sets that ensure that the probes are only meaningful and/or recognized by those individuals whose brains indeed have experienced the specific person, place, object, or event detail being probed. There are countless scenarios in which an innocent person might process a given probe stimulus in a manner that suggests personal meaningfulness or evokes vivid recall of a memory that has nothing to do with the event in question. In all research conducted thus far, researchers had perfect access to the veridical truth about participants’ experiential history—for instance, by controlling the mock crime scenario or selecting the photographs from the wearable cameras. What remains unknown is how such technologies would work when investigators have varying degrees of uncertainty about which stimuli should match a person's recollection or trigger recognition. For example, would photos of a terrorist training camp trigger recognition or recollection if they were from a vantage point that a suspect had never seen? Would a years-old photograph of an associate's face, with a different hairstyle, eyewear, and countenance, elicit recognition? Were details known only to a crime participant and investigators, such as the p. 2347 ↳ paisley-patterned couch the victim was found on, really encoded by the perpetrator? Or might memories be more reliably detected for acts such as trafficking and financial crimes, where a perpetrator has repeated exposures to a particular place, face, or documents?

Finally, it is known that there are substantial individual differences in people's autobiographical memory abilities (Palombo et al., 2018) and the associated brain processes (Petrican et al. 2020), which may further complicate efforts to develop general-purpose memory detection algorithms. However, as discussed above, there is some evidence that large-scale fMRI activation patterns appear to be sufficiently consistent across people to allow a classifier trained on one group of subjects to succeed at labeling whether a new subject's brain is experiencing recognition. Such findings will need to be critically examined in studies with much larger samples to assess the degree to which age differences and other sociodemographic factors may impact the across-participant consistency of memory-related brain activity patterns.

By highlighting the progress of memory detection efforts to date, as well as the shortcomings and unresolved issues, we hope to encourage future research aimed at fostering a more complete understanding of the boundary conditions that will constrain potential forensic applications.

Note

1. Portions of this chapter have been adapted from Murphy, E. R., & Rissman, J. (2020). Evidence of memory from brain data. Journal of Law and the Biosciences, *7*(1), lsaa078. https://doi.org/10.1093/jlb/lsaa078.

References

Ambach, W., & Gamer, M. (2018). Physiological measures in the detection of deception and concealed information. In J. P. Rosenfeld (Ed.), *Detecting concealed information and deception* (pp. 3–33). Elsevier.
Google Scholar Google Preview WorldCat COPAC

Barry, D. N., & Maguire, E. A. (2019). Remote memory and the hippocampus: A constructive critique. *Trends in Cognitive Sciences*, 23(2), 128–142.
Google Scholar WorldCat

Ben-Shakhar, G. (2012). Current research and potential applications of the concealed information test: An overview. *Frontiers in Psychology,* 3, Article 342.
Google Scholar WorldCat

Bergström, Z. M., Anderson, M. C., Buda, M., Simons, J. S., & Richardson-Klavehn, A. (2013). Intentional retrieval suppression can conceal guilty knowledge in ERP memory detection tests. *Biological Psychology*, 94(1), 1–11.
Google Scholar WorldCat

Boccia, M., Piccadi, L., Palermo, L., Nemmi, F., Sulpizio, V., Galati, G., & Guariglia, C. (2015). A penny for your thoughts! Patterns of fMRI activity reveal the content and the spatial topography of visual mental images. *Human Brain Mapping*, 36(3), 945–958.
Google Scholar WorldCat

Brown, T. I., Rissman, J., Chow, T. E., Uncapher, M. R., & Wagner, A. D. (2018). Differential medial temporal lobe and parietal cortical contributions to real-world autobiographical episodic and autobiographical semantic memory. *Scientific Reports,* 8, Article 6190.
Google Scholar WorldCat

Chow, T. E., & Rissman, J. (2017). Neurocognitive mechanisms of real-world autobiographical memory retrieval: Insights from studies using wearable camera technology. *Annals of the New York Academy of Sciences*, 1396(1), 202–221.
Google Scholar WorldCat

Chow, T. E., Westphal, A. J., & Rissman, J. (2018). Multi-voxel pattern classification differentiates personally experienced event memories from secondhand event knowledge. *NeuroImage,* 176, 110–123.
Google Scholar WorldCat

Cohen, J. D., Daw, N., Engelhardt, B., Hasson, U., Li, K., Niv, Y., Norman, K. A., Pillow, J., Ramadge, P. J., Turk-Browne, N. B., & Willke, T. L. (2017). Computational approaches to fMRI analysis. *Nature Neuroscience*, 20(3), 304–313.
Google Scholar WorldCat

Datta, R., & DeYoe, E. A. (2009). I know where you are secretly attending! The topography of human visual attention revealed with fMRI. *Vision Research*, 49(10), 1037–1044.
Google Scholar WorldCat

Dennis, N. A., Bowman, C. R., & Turney, I. C. (2015). Functional neuroimaging of false memories. In D. R. Addis, M. Barense, & A. Duarte (Eds.), *The Wiley Handbook on the Cognitive Neuroscience of Memory* (pp. 150–171). Wiley.
Google Scholar Google Preview WorldCat COPAC

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, 123–131.
Google Scholar WorldCat

Farahany, N. A. (2015). Neuroscience and behavioral genetics in US criminal law: An empirical analysis. *Journal of Law and the Biosciences*, 2(3), 485–509.
Google Scholar WorldCat

p. 2348 Farwell, L. A. (2012). Brain fingerprinting: A comprehensive tutorial review of detection of concealed information with event-related brain potentials. *Cognitive Neurodynamics*, 6(2), 115–154.
Google Scholar WorldCat

Farwell, L. A., & Donchin, E. (1991). The truth will out: Interrogative polygraphy (“lie detection”) with event-related brain potentials. *Psychophysiology*, 28(5), 531–547.
Google Scholar WorldCat

Farwell, L. A., Richardson, D. C., & Richardson, G. M. (2013). Brain fingerprinting field studies comparing P300-MERMER and P300 brainwave responses in the detection of concealed information. *Cognitive Neurodynamics*, 7(4), 263–299.
Google Scholar WorldCat

Farwell, L. A., & Smith, S. S. (2001). Using brain MERMER testing to detect knowledge despite efforts to conceal. *Journal of Forensic Sciences*, 46(1), 135–143.
Google Scholar WorldCat

Gamer, M. (2011). Detecting concealed information using autonomic measures. In B. Verschuere, G. Ben-Shakhar, & E. Meijer (Eds.), *Memory detection: Theory and application of the Concealed Information Test* (pp. 27–45). Cambridge University Press.
Google Scholar Google Preview WorldCat COPAC

Gilbert, S. J., & Fung, H. (2018). Decoding intentions of self and others from fMRI activity patterns. *NeuroImage,* 172, 278–290.
Google Scholar WorldCat

Giridharadas, A. (2008, September 14). India’s novel use of brain scans in courts is debated. The New York Times.

Greely, H. T., & Farahany, N. A. (2019). Neuroscience and the criminal justice system. *Annual Review of Criminology,* 2, 451–471.
Google Scholar WorldCat

Hardt, O., Nader, K., & Nadel, L. (2013). Decay happens: The role of active forgetting in memory. *Trends in Cognitive Sciences*, 17(3), 111–120.
Google Scholar WorldCat

Harrington v. Iowa, No. PCCV 073247 (Dist. Ct. Iowa, March 5, 2001); discussed in *Harrington v. Iowa*, 659 NW 2d 509 (Iowa 2003).

Haynes, J. D. (2015). A primer on pattern-based approaches to fMRI: Principles, pitfalls, and perspectives. *Neuron*, 87(2), 257–270.
Google Scholar WorldCat

Horikawa, T., Tamaki, M., Miyawaki, Y., & Kamitani, Y. (2013). Neural decoding of visual imagery during sleep. *Science*, 340(6132), 639–642.
Google Scholar WorldCat

Hu, X., Bergstrom, Z., Gagnepain, P., & Anderson, M. (2015). Suppressing unwanted autobiographical memories reduces their automatic influences: Evidence from electrophysiology and an implicit autobiographical memory test. *Psychological Science*, 26(7), 1098–1106.
Google Scholar WorldCat

Hu, X., Pornpattananangkul, N., & Rosenfeld, J. P. (2013). N200 and P300 as orthogonal and integrable indicators of distinct awareness and recognition processes in memory detection. *Psychophysiology*, 50(5), 454–464.
Google Scholar WorldCat

Hu, X., Rosenfeld, J. P., & Bodenhausen, G. V. (2012). Combating automatic autobiographical associations: The effect of instruction and training in strategically concealing information in the autobiographical implicit association test. *Psychological Science*, 23(10), 1079–1085.
Google Scholar WorldCat

Hubbard, J., Kikumoto, A., & Mayr, U. (2019). EEG decoding reveals the strength and temporal dynamics of goal-relevant representations. *Scientific Reports*, 9(1), 1–11.
Google Scholar WorldCat

King, J. R., & Dehaene, S. (2014). Characterizing the dynamics of mental representations: The temporal generalization method. *Trends in Cognitive Sciences*, 18(4), 203–210.
Google Scholar WorldCat

Lee, H., & Kuhl, B. A. (2016). Reconstructing perceived and retrieved faces from activity patterns in lateral parietal cortex. *Journal of Neuroscience*, 36(22), 6069–6082.
Google Scholar WorldCat

Lewis-Peacock, J. A., & Norman, K. A. (2013). Multi-voxel pattern analysis of fMRI data. In M. S. Gazzaniga & G. R. Mangun (Eds.), *The cognitive neurosciences* (5th ed., pp. 911–919). MIT Press.
Google Scholar Google Preview WorldCat COPAC

Loftus, E. F. (2017). Eavesdropping on memory. *Annual Review of Psychology,* 68, 1–18.
Google Scholar WorldCat

Lukács, G., Weiss, B., Dalos, V. D., Kilencz, T., Tudja, S., & Csifcsak, G. (2016). The first independent study on the complex trial protocol version of the P300-based concealed information test: Corroboration of previous findings and highlights on vulnerabilities. *International Journal of Psychophysiology,* 110, 56–65.
Google Scholar WorldCat

Lykken, D. T. (1959). The GSR in the detection of guilt. *Journal of Applied Psychology*, 43(6), 385–388.
Google Scholar WorldCat

McDermott, K. B., Szpunar, K. K., & Christ, S. E. (2009). Laboratory-based and autobiographical retrieval tasks differ substantially in their neural substrates. *Neuropsychologia*, 47(11), 2290–2298.
Google Scholar WorldCat

Meegan, D. V. (2008). Neuroimaging techniques for memory detection: Scientific, ethical, and legal issues. *American Journal of Bioethics*, 8(1), 9–20.
Google Scholar WorldCat

Mehrnam, A. H., Nasrabadi, A. M., Ghodousi, M., Mohammadian, A., & Torabi, S. (2017). A new approach to analyze data from EEG-based concealed face recognition system. *International Journal of Psychophysiology,* 116, 1–8.
Google Scholar WorldCat

Meijer, E. H., Ben-Shakhar, G., Verschuere, B., & Donchin, E. (2013). A comment on Farwell (2012): Brain fingerprinting: A comprehensive tutorial review of detection of concealed information with event-related brain potentials. *Cognitive Neurodynamics*, 7(2), 155–158.
Google Scholar WorldCat

p. 2349 Meixner, J. B., & Rosenfeld, J. P. (2014). Detecting knowledge of incidentally acquired, real-world memories using a P300-based concealed-information test. *Psychological Science*, 25(11), 1994–2005.
Google Scholar WorldCat

Mendelsohn, A., Furman, O., & Dudai, Y. (2010). Signatures of memory: Brain coactivations during retrieval distinguish correct from incorrect recollection. *Frontiers in Behavioral Neuroscience,* 4, Article 18.
Google Scholar WorldCat

Mertens, R., & Allen, J. J. (2008). The role of psychophysiology in forensic assessments: Deception detection, ERPs, and virtual reality mock crime scenarios. *Psychophysiology*, 45(2), 286–298.
Google Scholar WorldCat

Misra, P., Marconi, A., Peterson, M., & Kreiman, G. (2018). Minimal memory for details in real life events. *Scientific Reports*, 8(1), 16701.
Google Scholar WorldCat

Mukundan, C., Sumit, S., & Chetan, S. (2017). Brain Electrical Oscillations Signature Profiling (BEOS) for measuring the process of remembrance. *EC Neurology,* 8, 217–230.
Google Scholar WorldCat

Naci, L., Cusack, R., Jia, V. Z., & Owen, A. M. (2013). The brain's silent messenger: Using selective attention to decode human thought for brain-based communication. *Journal of Neuroscience*, 33(22), 9385–9393.
Google Scholar WorldCat

Nishimoto, S., Vu, A., Naselaris, T., Benjamini, Y., Yu, B., & Gallant, J. (2011). Reconstructing visual experiences from brain activity evoked by natural movies. *Current Biology*, 21(19), 1641–1646.
Google Scholar WorldCat

Noh, E., Liao, K., Mollison, M., Curran, T., & de Sa, V. (2018). Single-trial EEG analysis predicts memory retrieval and reveals source-dependent differences. *Frontiers in Human Neuroscience,* 12, Article 258.
Google Scholar WorldCat

Palombo, D. J., Sheldon, S., & Levine, B. (2018). Individual differences in autobiographical memory. *Trends in Cognitive Sciences*, 22(7), 583–597.
Google Scholar WorldCat

Peth, J., Sommer, T., Hebart, M., Vossel, G., Buchel, C., & Gamer, M. (2015). Memory detection using fMRI: Does the encoding context matter? *NeuroImage,* 113, 164–174.
Google Scholar WorldCat

Petrican, R., Palombo, D. J., Sheldon, S., & Levine, B. (2020). The neural dynamics of individual differences in episodic autobiographical memory. *eNeuro*, 7(2), ENEURO.0531-19.2020.
Google Scholar WorldCat

Polich, J., & Kok, A. (1995). Cognitive and biological determinants of P300: An integrative review. *Biological Psychology*, 41(2), 103–146.
Google Scholar WorldCat

Presidential Commission for the Study of Bioethical Issues. (2015). *Gray matters: Topics at the intersection of neuroscience, ethics, and society* (*Vol. 2*).
Google Scholar Google Preview WorldCat COPAC

Rissman, J., Chow, T., Reggente, N., & Wagner, A. (2016). Decoding fMRI signatures of real-world autobiographical memory retrieval. *Journal of Cognitive Neuroscience*, 28(4), 604–620.
Google Scholar WorldCat

Rissman, J., Greely, H. T., & Wagner, A. D. (2010). Detecting individual memories through the neural decoding of memory states and past experience. *Proceedings of the National Academy of Sciences of the USA*, 107(21), 9849–9854.
Google Scholar WorldCat

Rissman, J., & Wagner, A. D. (2012). Distributed representations in memory: Insights from functional brain imaging. *Annual Review of Psychology,* 63, 101–128.
Google Scholar WorldCat

Rosenfeld, J. P. (2005). Brain fingerprinting: A critical analysis. *Scientific Review of Mental Health Practice*, 4(1), 20–37.
Google Scholar WorldCat

Rosenfeld, J. P. (2019). P300 in detecting concealed information and deception: A review. *Psychophysiology*, 57(7), e13362.
Google Scholar WorldCat

Rosenfeld, J. P., Ben-Shakhar, G., & Ganis, G. (2013). Detection of concealed stored memories with psychophysiological and neuroimaging methods. In L. Nadel & W. P. Sinnott-Armstrong (Eds.), *Memory and Law* (pp. 263–305). Oxford University Press.
Google Scholar Google Preview WorldCat COPAC

Rosenfeld, J. P., Biroschak, J. R., & Furedy, J. J. (2006). P300-based detection of concealed autobiographical versus incidentally acquired information in target and non-target paradigms. *International Journal of Psychophysiology*, 60(3), 251–259.
Google Scholar WorldCat

Rosenfeld, J. P., Cantwell, B., Nasman, V., Wojdac, V., Ivanov, S., & Mazzeri, L. (1988). A modified, event-related potential-based guilty knowledge test. *International Journal of Neuroscience*, 42(1–2), 157–161.
Google Scholar WorldCat

Rosenfeld, J. P., & Greely, H. T. (2009). Deception, detection of, P300 event-related potential (ERP). In A. Jamieson & A. A. Moenssens (Eds.), *Wiley Encyclopedia of Forensic Science.* Wiley.
Google Scholar Google Preview WorldCat COPAC

Rosenfeld, J. P., Labkovsky, E., Winograd, M., Liu, M., Vandenboom, C., & Chedid, E. (2008). The Complex Trial Protocol (CTP): A new, countermeasure-resistant, accurate, P300-based method for detection of concealed information. *Psychophysiology*, 45(6), 906–919.
Google Scholar WorldCat

Rosenfeld, J. P., Soskins, M., Bosh, G., & Ryan, A. (2004). Simple, effective countermeasures to P300-based tests of detection of concealed information. *Psychophysiology*, 41(2), 205–219.
Google Scholar WorldCat

Rosenfeld, J. P., Ward, A., Drapekin, J., Labkovsky, E., & Tullman, S. (2017). Instructions to suppress semantic memory enhances or has no effect on P300 in a concealed information test (CIT). *International Journal of Psychophysiology,* 113, 29–39.
Google Scholar WorldCat

Saarimäki, H., Gostopoulos, A., Jääskeläinen, I., Lampinen, J., Vuilleumier, P., Hari, R., Sams, M., & Nummenmaa, L. (2015). Discrete neural signatures of basic emotions. *Cerebral Cortex*, 26(6), 2563–2573.
Google Scholar WorldCat

p. 2350 Schacter, D. L., & Addis, D. R. (2007). The cognitive neuroscience of constructive memory: Remembering the past and imagining the future. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences*, 362(1481), 773–786.
Google Scholar WorldCat

Schacter, D. L., & Loftus, E. F. (2013). Memory and law: What can cognitive neuroscience contribute? *Nature Neuroscience*, 16(2), 119–123.
Google Scholar WorldCat

Slotnick, S. D., & Schacter, D. L. (2004). A sensory signature that distinguishes true from false memories. *Nature Neuroscience*, 7(6), 664–672.
Google Scholar WorldCat

Sokolovsky, A., Rothenberg, J., Labkovsky, E., Meixner, J., & Rosenfeld, J. (2011). A novel countermeasure against the reaction time index of countermeasure use in the P300-based complex trial protocol for detection of concealed information. *International Journal of Psychophysiology*, 81(1), 60–63.
Google Scholar WorldCat

Stokes, M., Thompson, R., Cusack, R., & Duncan, J. (2009). Top-down activation of shape-specific population codes in visual cortex during mental imagery. *Journal of Neuroscience*, 29(5), 1565–1572.

Google Scholar WorldCat

Thakral, P. P., Wang, T. H., & Rugg, M. D. (2017). Decoding the content of recollection within the core recollection network and beyond. *Cortex,* 91, 101–113.
Google Scholar WorldCat

Uncapher, M. R., Boyd-Meredith, J., Chow, T., Rissman, J., & Wagner, A. (2015). Goal-directed modulation of neural memory patterns: Implications for fMRI-based memory detection. *Journal of Neuroscience,* 35 (22), 8531–8545.
Google Scholar WorldCat

Wang, X., Liang, X., Jiang, Z., Nguchu, B., Zhou, Y., Wang, Y., Wang, H., Li, Y., Zhu, Y., Wu, F., Gao, J., & Qui, B. (2020). Decoding and mapping task states of the human brain via deep learning. *Human Brain Mapping*, 41(6), 1505–1519.
Google Scholar WorldCat

Ward, A. C., & Rosenfeld, J. P. (2017). Attempts to suppress episodic memories fail but do produce demand: Evidence from the P300-based Complex Trial Protocol and an implicit memory test. *Applied Psychophysiology and Biofeedback*, 42(1), 13–26.
Google Scholar WorldCat

Wimmer, G. E., Liu, Y., Vehar, N., Behrens, T. E., & Dolan, R .J. (2020). Episodic memory retrieval success is associated with rapid replay of episode content. *Nature Neuroscience*, 23(8), 1025–1033.
Google Scholar WorldCat

Wing, E. A., Ritchey, M., & Cabeza, R. (2015). Reinstatement of individual past events revealed by the similarity of distributed activation patterns during encoding and retrieval. *Journal of Cognitive Neuroscience*, 27(4), 679–691.
Google Scholar WorldCat

Winograd, M. R., & Rosenfeld, J. P. (2011). Mock crime application of the Complex Trial Protocol (CTP) P300-based concealed information test. *Psychophysiology*, 48(2), 155–161.
Google Scholar WorldCat

Winograd, M. R., & Rosenfeld, J. P. (2014). The impact of prior knowledge from participant instructions in a mock crime P300 Concealed Information Test. *International Journal of Psychophysiology*, 94(3), 473–481.
Google Scholar WorldCat

Wixted, J. T. (2005). A theory about why we forget what we once knew. *Current Directions in Psychological Science*, 14(1), 6–9.
Google Scholar WorldCat