



Electrophysiological markers of working memory usage as an index for truth-based lies

Yu-Hui Lo^{1,2} · Philip Tseng^{1,3}

Published online: 18 July 2018
© Psychonomic Society, Inc. 2018

Abstract

People prefer to lie using altered truthful events from memory, perhaps because doing so can increase their credibility while reducing cognitive and working memory (WM) load. One possible way to counter such deceptive behavior is to track WM usage, since fabricating coherent lies or managing between truth and lies is likely to involve heavy WM load. In this study, participants memorized a list of words in the study session and used these old words to provide deceptive answers when cued later, in the testing session. Our behavioral results showed that people needed more time to make a deceptive response during the execution stage, and this prolonged deceptive reaction time (RT) was negatively correlated with each participant's WM capacity. Event-related potential findings showed a more negative-going frontal amplitude between the lie and truth conditions during the preparation stage, suggesting that WM preparatory processes can be detected long before a deceptive response is verbalized. Furthermore, we observed a larger positive frontal-central amplitude during the execution stage, which was negatively correlated with participants' lie–truth RT differences, suggesting that participants' efficiency in producing deceptive responses can be readily traced electrophysiologically. Together, these findings suggest that WM capacity and preparation are crucial to efficient lying and that their related electrophysiological signatures can potentially be used to uncover deceptive behaviors.

Keywords Short-term memory capacity · Working memory capacity · Guilty knowledge test · Concealed information test

The ability to deceive others is a result of complex interactions between different cognitive faculties and is arguably one of the hallmarks of cognitive development (e.g., Talwar & Lee, 2002). Decades of psychological research from an information-processing approach have made it possible for researchers to look into the inner workings of these various cognitive mechanisms, and thereby to detect lies by tracking these cognitive processes that occur behind the scenes. For example, although deception is hard to detect even with cutting-edge neuroscientific tools, decoding memory traces from neural signals is less cognitively complex and perhaps less difficult to monitor (e.g., Chadwick, Hassabis, Weiskopf, & Maguire, 2010). Therefore, by using our existing knowledge of the cognitive or

electrophysiological components behind deception, it is possible to spot lies by detecting recognition memory. Excellent examples of tools for such work are the Guilty Knowledge Test (GKT) and the Concealed Information Test (CIT).

Detecting concealed information

Most of the lie detection studies to date have focused on information concealment, and in this regard the field has made great progress. In this context, the goal is to dig out traces (electrophysiological, memory, or behavioral) of critical knowledge that the participants or suspects deny having. In the GKT, such memory traces can be verified by computing the differences in reaction times (RTs) or P300 amplitudes between three kinds of items: target, irrelevant, and probe. Participants perform a binary judgment task, in which they are instructed to press the “yes” button when the word or picture on the screen is a known item (i.e., target), or to press “no” button when the word is not known (i.e., irrelevant) or crime-related (i.e., probe). The rationale is that the investigators actually know the ground truth for both target and irrelevant items, hence they serve as appropriate baselines

✉ Philip Tseng
philip@tmu.edu.tw

¹ TMU—Research Center of Brain and Consciousness, Taipei Medical University, No. 250, Wuxin St, Taipei City 11031, Taiwan

² Shuang-Ho Hospital, Taipei Medical University, Taipei, Taiwan

³ Graduate Institute of Humanities in Medicine, Taipei Medical University, Taipei, Taiwan

for known and unknown responses that the probe response can be compared against. Using this setup, Seymour, Seifert, Shafto, and Mosmann (2000) first reported slower RT for the probes than irrelevant items, suggesting that denial of meaningful probe information requires extra stages of inhibitory processing. A discriminant function analysis also correctly classified 89% of the guilty participants and 100% of the innocent participants. Combining GKT with event-related potentials (ERPs), Farwell and Donchin (1991; see also Rosenfeld, Angell, Johnson, & Qian, 1991) used an oddball paradigm and designed crime-related information as a deviant (low-frequency) stimulus for the suspects, since innocent participants do not have the knowledge to tell them apart from a standard (high-frequency) irrelevant stimulus. They observed a larger P300 component at the parietal region in the guilty over innocent participants. Since then, the P300 amplitude has been regarded as an index of recognition memory in GKT that can be used to infer deception, and its protocols and parameters have been optimized to work against countermeasures (Rosenfeld et al., 2008).

Following the same logic, other variations of the memory task have also been used to take advantage of recognition processes (Brown dyke et al., 2008; Johnson, Barnhardt, & Zhu, 2003, 2004; Tardif, Barry, Fox, & Johnstone, 2000; van Hooff, Brunia, & Allen, 1996). In one study by Tradif et al., participants were instructed to feign memory impairment in order to mimic what often happens in medical settings for financial gain. Participants performed a recognition memory task and had to discriminate whether a word was old (i.e., seen in the study phase) or new (not seen in the study phase). The authors found that truth-telling participants had a larger positive amplitude to old words at 450 ms after word onset, and interestingly, that the malingering group had an even earlier old–new difference (0- to 450-ms window) from the left frontal region, even though such a pattern was not present in the control group. In another study, Wu, Hu, and Fu (2009) demonstrated that the P300 can be useful in detecting both forced and self-directed lies (though the magnitude of the component is smaller in forced lies). Finally, researchers have also observed a larger deception-related contingent negative variation (CNV) component—which is indicative of anticipation and response preparation—around the 1,000-ms window when participants had to deny recognition of familiar target faces (Fang, Liu, & Shen, 2003) or knowledge of a mock crime (Suchotzki, Crombez, Smulders, Meijer, & Verschuere, 2015), identify fake banknotes as genuine (Sun, Mai, Liu, Liu, & Luo, 2011), and rate an unattractive face as attractive (Dong, Wu, & Lu, 2010). These results imply that some preparatory processes that separate deceptive from truth-telling participants can start as early as before the 400-ms window, or as late as around the 1,000-ms window, which then propagates to later time intervals, such as response execution.

Truth-based lying, cognitive load, and working memory

Beyond the concealment of truth, lies that involve partial truth (i.e., lies that are somewhat fact-based instead of pure fabrication) or misleadingly formulated truth are also frequently used in interviews or interrogations (e.g., Leins, Fisher, & Ross, 2013). This scenario would be applicable to the interrogation that would take place after a culprit has been outed by a successful GKT: Imagine that an interrogatee confesses upon seeing his or her GKT results; the logical next step is to extract useful information, such as identification of culprits, aliases, and so on. However, although pictures of culprits can be identified via the GKT or variants of the implicit association task (Verschuere, Suchotzki, & Debey, 2015), the interrogatee can still lie about recognized faces, photos, or keywords, especially with partially truthful stories, in order to lead the investigation astray (e.g., “Yes I know him, but his name is Sam, not John,” “He is an acquaintance from somewhere else, not a member of the group,” etc.). Indeed, in one study by Leins et al., 67% of the participants chose to use their previous *truthful* experiences to construct their lies, and that number went up to 86% when the participants were informed that the interviewer would check the credibility of their reports. Therefore, deception can be achieved even with truthful events from memory, if such memory is, for instance, narrated in a misleading way, with intentionally omitted (but important) details, or perhaps falsely coupled with other truthful but nonapplicable information (e.g., wrong time, place, etc.). From the Leins et al. study, such truth-based lies seemed to be people’s preferred choice to increase the credibility of their lies. This is perhaps because truth-based stories can minimize people’s working memory (WM) load, as it would be too much information to maintain if one were to invent an entire event from scratch. Furthermore, lies in this category often involve spoken lies that are complex and elaborate, rather than simple denial or binary yes–no answers.

In the context of such truth-based and slightly more elaborate lies that are meant to mislead, recognition memory paradigms are less fitting, since recognition is already confirmed. However, WM usage may be a viable approach to investigating this type of lie. This is because fabricating lies takes time and requires constant WM maintenance in order to stay coherent as the number of scripts increases. Even for truth-based lies, the contents and details need to be weaved in a way that is internally consistent (Granhag & Hartwig, 2008), externally consistent (i.e., consistent with what the interrogator knows; DePaulo et al., 2003), and multimodal (i.e., depicting visual, verbal, or other nonverbal events; Vrij, Mann, Leal, & Fisher, 2010). This places greater cognitive load on liars, because truthful events from episodic memory are recalled and placed in the episodic WM buffer (Baddeley, 2000) for further processing and verbalization, and thus should involve greater

engagement from the central executive. This idea is captured well by the activation–decision–construction model (ADCM), proposed by Walczyk, Roper, Seemann, and Humphrey (2003) to account for the extra number of stages of processing behind lying. In this model, the difference between truth-telling and lying occur in the decision and construction stages, where the interrogatee has to decide or remember whether or not to lie, and if so, to construct a plausible reply, all of which are extra cognitive processes that truth-telling does not entail. Specifically, the decision component involves remembering relevant context and deciding whether or not to lie (i.e., decision making), whereas the construction component requires attention in order to construct a coherent response and prevent utterance of the truth (i.e., inhibitory control), all of which are notable elements of Baddeley's WM model. In the context of real-life scenarios, the decision-making component would be applicable to situations in which the interrogatee has to withhold the script until the question is heard. If it is then more advantageous to tell the truth, the rehearsed lie in WM has to be suppressed (i.e., the construction component). Indeed, recent studies have begun to demonstrate WM as a series of processes that requires maintenance, decision making, and inhibitory control, as opposed to maintenance alone (e.g., Unsworth, Fukuda, Awh, & Vogel, 2014; see Johnson et al., 2004, for an example in deception). In this study we aimed to investigate the impact of these processes, by separating them temporally into the preparation stage (i.e., maintenance) and the execution stage (decision and inhibition), which we describe in more detail below.

In light of the differences in cognitive load between truth-telling and lying, it makes sense that this load-based approach has been gaining support lately and has also been implemented by inducing cognitive load on participants (e.g., Vrij, Fisher, Mann, & Leal, 2006; for a review, see Walczyk, Igou, Dixon, & Tcholakian, 2013). Notably, in a series of studies by Vrij and colleagues (Vrij, Fisher, Mann, & Leal, 2008; Vrij & Granhag, 2012; Vrij, Mann, et al., 2008; Vrij et al., 2010), these authors found that load-inducing techniques such as asking unanticipated questions, maintaining eye contact with the interviewer, recalling events in reverse chronological order, and dual-tasking can all produce more behavioral clues to deception, such as increased numbers of eye blinks and speech errors and hesitations. Therefore, to detect spontaneous or truth-based deception beyond the level of recognition memory, WM usage and load may be useful measures for doing so.

The present study

In the present study, we aimed to investigate the behavioral and neurophysiological markers of lying by tracking WM usage instead of inducing cognitive load. The task here presents a simplified version of the misleading lies that

are often used during an interrogation. Instead of inducing cognitive load in participants, here we passively monitored people's WM usage as they performed the deception task. As such, we expected to see a wide range of individual differences in performance on the deception task. But, crucially, we hypothesized that such variability in performance should be related to people's WM capacity. To do this, we measured each person's verbal and visual WM capacity in two separate tasks, and we expected to find a correlation between people's WM scores and their performance on the deception task. This relationship should also be true for the electrophysiological data.

In this task, participants first went through a study phase of ten words per block, which constituted the factual words they would later use for lying. In the testing phase, participants were cued to either lie or tell the truth using the words they had previously committed to memory in the study phase. Crucially, in the deception trials, participants would see a familiar word on screen (i.e., a word from the study phase), but, instead of denying having seen it before (i.e., a pure denial lie), participants had to lie about the word identity by saying another truthful word from memory (i.e., a truth-based lie). Therefore, the participants were lying using actual materials from memory and from the same event (i.e., the study phase), but in an incorrect or deceptive manner. With this setup we aimed to mimic situations in which the suspect needs to keep several factual answers or truth-based scripts ready in WM, which would allow us to investigate the cognitive load and individual differences behind this type of truth-based deception. Furthermore, to investigate the effects of the different components of WM, each trial was structured to consist of a preparation (i.e., WM maintenance) and execution (i.e., WM decision and inhibition) stage, so that our event-related potentials (ERPs) could be contextualized in terms of the cognitive processing that was going on at the time.

Method

Participants

Twenty-seven participants were recruited in this experiment (14 females, 13 males; mean age = 24.67 years). All participants have no previous neurological history, and have normal or corrected-to-normal vision. All experimental procedures were approved by the Institutional Review Board of the Taipei Medical University, Taiwan, and all participants gave informed consent prior to their participation. No participants were excluded from analysis of the behavioral data, but three were excluded from the ERP analysis due to too many ocular artifacts in their EEG data (retained for ERP analysis: 11 females, 13 males; mean age = 24.67 years).

Task and procedures

In this experiment, participants first performed the main deception task, then performed two separate complex WM span tasks—the operation span and symmetry span tasks (short version; Foster et al., 2015), which were designed to measure each participant’s verbal and visual WM capacity, respectively.

The main task consisted of 20 blocks, and each block was composed of two phases: a study phase and a test phase (Fig. 1). During the study phase, ten words were displayed sequentially (5 s per word, with a 1-s interstimulus interval), and the participants were instructed to memorize as many of the words as possible. There were a total of 242 words (22 practice, 220 formal session), and all were two-character concrete nouns that were controlled for frequency of usage using the Academia Sinica Corpus (Chen, Lo, Liu, & Cheng, 2016; Huang, Ahrens, & Chen, 1998).

During the test phase, each test phase contained six trials, which began with either a lie or truth color cue (3 s; preparation stage), followed by a blank interval (1 s), and finally the target word (5 s; decision and execution stage). The color cues were red and green for lie and truth, respectively, for 18 participants, and these colors were reversed for the remaining nine participants. Participants had to say their response to a microphone, with their vocal response latencies being recorded using E-Prime and their accuracies calculated by the experimenter. Upon seeing the target word, participants had three types of responses to make. New words (i.e., target words that had not been part of the ten-word study list) were always catch trials, and participants simply had to say “didn’t see” upon seeing a new target word, regardless of the cue that preceded it. This was done to mimic the occasional occurrences in which, despite one’s intention and preparation to lie, suspects sometimes switch to truth-telling upon hearing a harmless question. That is, a liar can prepare multiple scripts in WM, but the final decision to lie or tell the truth is not decided upon until one hears the interrogator’s question, which would be akin to our word on the screen during the execution stage. In the truth condition, if a prestudied target word appeared on screen, participants were instructed to simply say the word’s true identity (e.g., the word “desk” in Fig. 1). Critically, in the lie condition, upon seeing an old word, participants had to lie about its identity by saying another old word from memory (e.g., substituting “desk” with “phone” in Fig. 1). This was done to mimic the situations in which an object, name, or script from the same cohort as the target is used to achieve a credible lie. In summary, there were three types of responses for the four conditions: truth cue and old target word (*truth old*; i.e., say the target word as is), lie cue and old target word (*lie old*; i.e., lie with another, unused old word), truth cue and new word (*truth new*; respond “didn’t

see”), and lie cue and new word (*lie new*; respond “didn’t see”). To prevent the participants from saying the same old word in every lie trial, they were told to use each old word only once while lying. The time limit for responding to the target word was 5 s, until the word’s offset. At the end of the study phase, there was a 10-s waiting period before the test phase began.

In each test phase there were six trials, one of which was a catch trial (i.e., a new target word). The remaining five trials were all old words, consisting of three truth old and two lie old trials in ten blocks, and three lie old and two truth old trials in another ten blocks, all presented in randomized order (at both the trial and block levels). This ratio created a medium level of difficulty, and not all ten words from the study phase were exhaustively used. Before the formal session began, participants completed two practice blocks in order to familiarize themselves with the task. The total time required to complete the main task was about 40 min.

Independent WM measures

After the main task, participants then completed two more WM tasks, the *operation span* and *symmetry span* tasks (Foster et al., 2015), which were designed to measure verbal and visual WM capacity, respectively. In both span tasks, the to-be-remembered letter (operation span) or location (symmetry span) and distractors appeared alternately in a sequence (Fig. 2). The numbers of to-be-remembered items and distractors varied between two and seven in each trial, in random order. In the operation span task, letters were to be remembered in correct order, while distracting mathematical statements (participants had to judge whether the statement is true or false) appeared between the letters. In the symmetry span task, the locations of a red square in a 4×4 grid were to be remembered and recalled in correct order, while symmetrical or asymmetrical figures appeared in between (participants had to judge whether or not the figure was symmetrical along the vertical axis). When participants finished the span tasks, the partial scores were calculated by summing the number of letters or blocks that were recalled correctly, such that the highest score in both tasks was 25 (sum of spans 3, 4, 5, 6, and 7). Therefore, a score of 25 indicated that the participant has not made a single error in recalling any of the stimuli.

Electroencephalography recording

EEG was continuously recorded from 64 Ag/AgCl electrodes (following the 10/20 system) with a reference electrode between Fz and Cz. All electrodes were mounted on a BrainCap electrode cap (Brain Products GmbH, Munich,

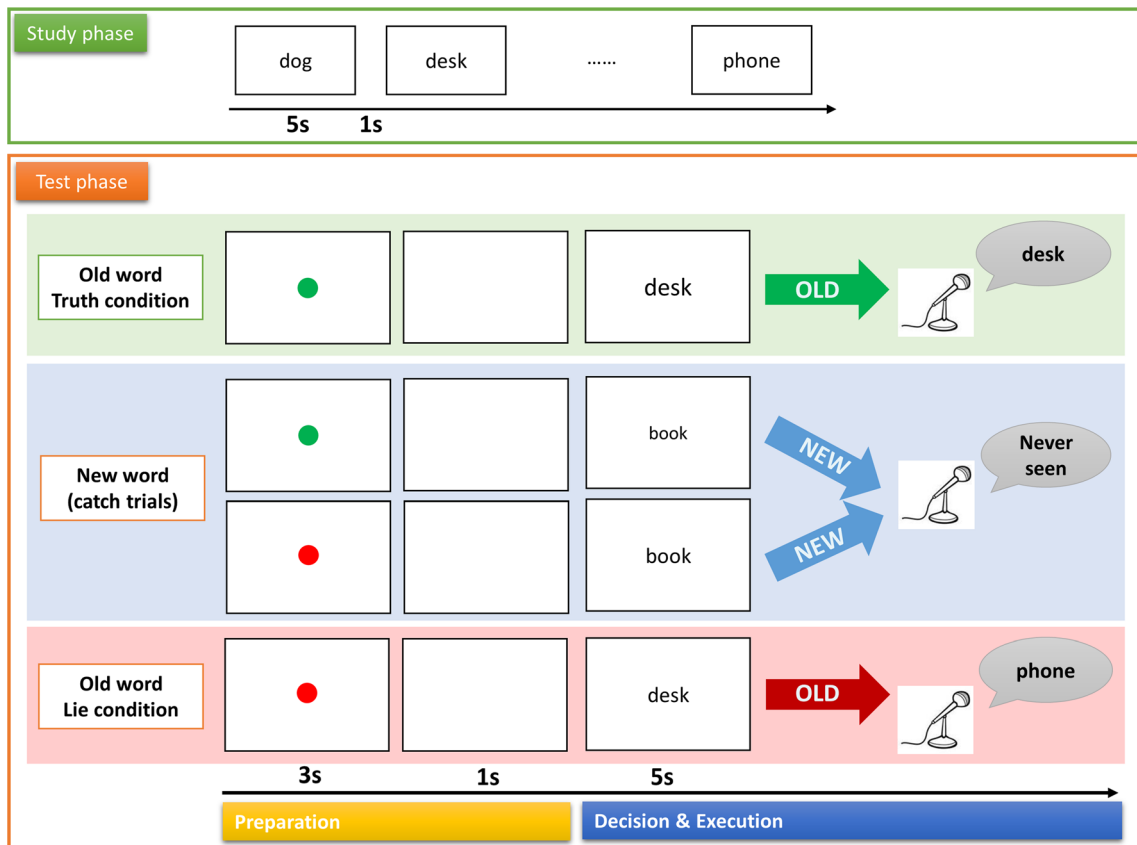


Fig. 1 Experimental procedure. Participants had to use pre-studied words to lie in the lie old condition (bottom row). Note that the actual experiment consisted of two-character Chinese nouns that were equated in length and usage frequency, and the English display here is for illustration only.

Germany). All signals were amplified using the BrainAmp amplifier (Brain Products, Munich, Germany). The signals were digitized at a sampling rate of 1000 Hz. Two sets of

electrodes were placed on the upper and lower sides of the right eye and on the canthi of both eyes, to measure participants' vertical and horizontal eye movements.

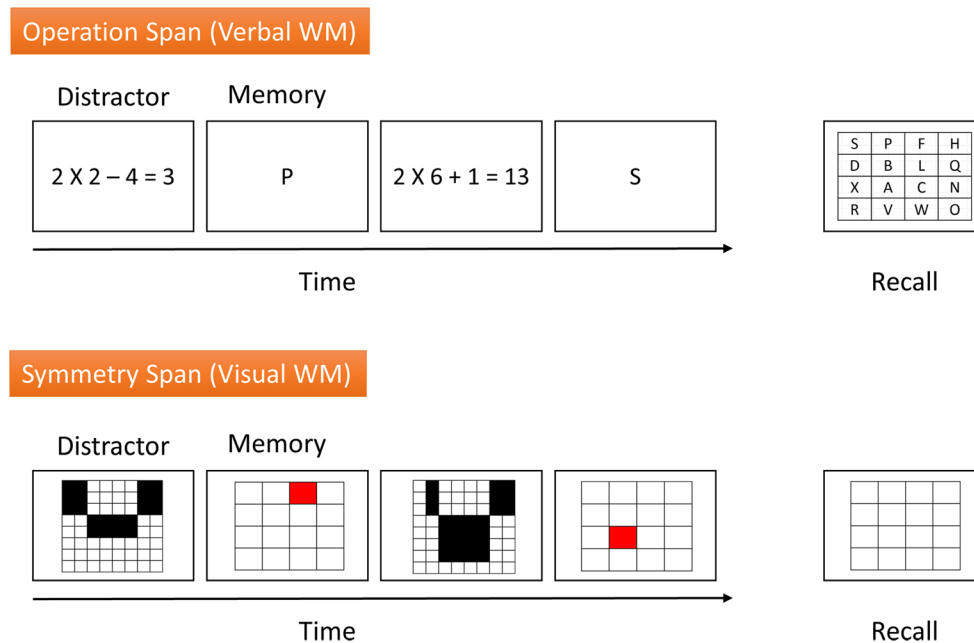


Fig. 2 Procedure of the operation span and symmetry span WM tasks. The memory materials were interleaved with distractor tasks (judging whether a mathematical statement was true vs. judging whether the block figure was symmetrical between left and right).

ERP data analysis and averaging

Eye blinks in the EEG data were corrected using the linear regression method in the Brain Vision Analyzer 2.1 software (Gratton, Coles, & Donchin, 1983). The continuous oculo-corrected EEG data were first offline re-referenced to the average of the electrodes at the left and right mastoids (M1 and M2). The EEG data were then segmented into epochs from 200 ms before stimulus onset to 1,000 ms after stimulus onset. A digital low-pass filter of 30 Hz (24 dB/octave) was applied to filter out high-frequency noise. Baseline correction was executed using a prestimulus interval (200 ms). Epochs with artifacts fluctuating over $\pm 70 \mu\text{V}$ or with incorrect verbal responses were rejected. The remaining trials [averaged numbers of trials: for the preparation stage—lie = 43.88 ($SE = 1.71$), truth = 49.88 ($SE = 1.67$); execution stage—lie old = 37.96 ($SE = 1.02$), truth old = 43.25 ($SE = 0.99$), lie new = 9.04 ($SE = 0.21$), truth new = 9.38 ($SE = 0.22$)] were averaged according to stimulus type. Each trial was divided into two segmented epochs including cue-locked (i.e., preparation stage) and target-locked (i.e., decision and execution stage) waveforms. There were two stimulus types (truth and lie) in cue-locked waveforms. In target-locked waveforms, there were four stimulus types: the truth old, lie old, truth new, and lie new conditions. Since the purpose of the experiment was to probe deceptive preparation and execution, we mainly compared the ERPs from the lie condition to the truth condition. Moreover, the memory old–new effect was investigated as well, by comparing the ERPs between the truth old, truth new, lie old, and lie new conditions. Three scalp regions were chosen to perform the statistical analysis, creating the within-subjects factor of anterior–posterior electrode: frontal (F3, FZ, F4), central (C3, CZ, C4), and parietal (P3, PZ, P4) regions. Another within-subjects factor was laterality: left (F3, C3, P3), middle (FZ, CZ, PZ), and right (F4, C4, P4).

Results

Behavioral results

Participants' RTs were analyzed with a repeated measures 2×2 ANOVA with cue type (lie vs. truth) and target type (old vs. new words) as within-subjects factors. The RT and accuracy results are visualized in Fig. 3. Note that the new words here included both the lie new and truth new words. There was a significant main effect of cue type [$F(1, 26) = 37.675, p < .001$], a marginally significant effect of target type [$F(1, 26) = 3.949, p = .058$], and, critically, a significant interaction between cue type and target type [$F(1, 26) = 27.376, p < .001$]. This interaction was driven by the longest RTs occurring in the lie old condition (1,405.12 ms), as compared to much shorter RTs in the truth old condition (1,082.13 ms).

Moreover, the RTs in the lie new condition (1,182.16 ms) and the truth new condition (1,157.77 ms) were not significantly different. [lie old vs. truth old: $t(26) = 6.387, p < .001$; lie new vs. truth new: $t(26) = 0.940, p = .356$].¹

The accuracy data were also submitted to a repeated measures 2×2 ANOVA with cue type (lie vs. truth) and target type (old vs. new words, where the new words included both the lie new and truth new words) as within-subjects factors. We found significant main effects of cue type [$F(1, 26) = 51.677, p < .001$] and target type [target type: $F(1, 26) = 100.059, p < .001$], as well as an interaction between them [$F(1, 26) = 20.218, p < .001$]. The mean accuracy in the lie old condition (80.07%) was significantly lower than the mean accuracy in the truth old condition (92.00%) [lie old vs. truth old: $t(26) = -7.359, p < .001$], and there was no significant difference between the lie new and truth new conditions [lie new: 97.04%; truth new: 98.89%; lie new vs. truth new: $t(26) = -1.412, p = .170$], which reflects the same pattern as we observed in the RT data.²

The averages of the participants' partial scores in the operation span task and symmetry span tasks were 22.26 and 11.33, respectively. We performed a Pearson correlation analysis between these partial scores and participants' behavioral performance in the main memory-based task (in RTs and accuracy). This was done to probe whether WM is an important contributor to the underlying processes of such truth-based verbal deception. If so, is visual WM, verbal WM, or both important to the task at hand? To this end, we observed a significant correlation between participants' RTs from the lie old condition and their operation span (verbal WM) performance [$r = -.552, p = .003$], but not for their symmetry performance (visual WM) [$r = .070, p = .727$], and this difference was significant using Pearson and Filon's (1898) z [$z = -2.6367, p = .0084$]. Therefore, verbal WM is particularly relevant to the mental effort in formulating truth-based lies. Importantly, when we computed the difference in RTs between the lie old and truth old conditions, this difference (i.e., the time cost for lying) was also negatively correlated with participants' verbal WM scores [$r = -.608, p = .001$], but it did not correlate with visual WM scores [$r = .189, p =$

¹ Because three participants were excluded from our ERP analyses, we also analyzed the remaining 24 participants' behavioral data, to make sure that all patterns remained the same.

In RTs, a 2×2 ANOVA revealed a significant main effect of cue type [$F(1, 23) = 33.469, p < .001$], a marginally significant effect of target type [$F(1, 23) = 3.105, p = .091$], and a significant interaction between cue type and target type [$F(1, 23) = 27.715, p < .001$]. The mean RT in the lie old condition (1,404.03 ms) was again significantly slower than the mean RT in the truth old condition (1,074.44 ms) [$t(23) = -6.325, p < .001$].

² In the smaller sample ($n = 24$), a 2×2 ANOVA showed significant main effects of cue type [$F(1, 23) = 43.442, p < .001$] and target type [target type: $F(1, 23) = 94.393, p < .001$], as well as an interaction between them [$F(1, 23) = 17.268, p < .001$]. The mean accuracy in the lie old condition (79.83%) was again significantly lower than the mean accuracy in the truth old condition (91.83%) [$t(23) = -6.720, p < .001$].

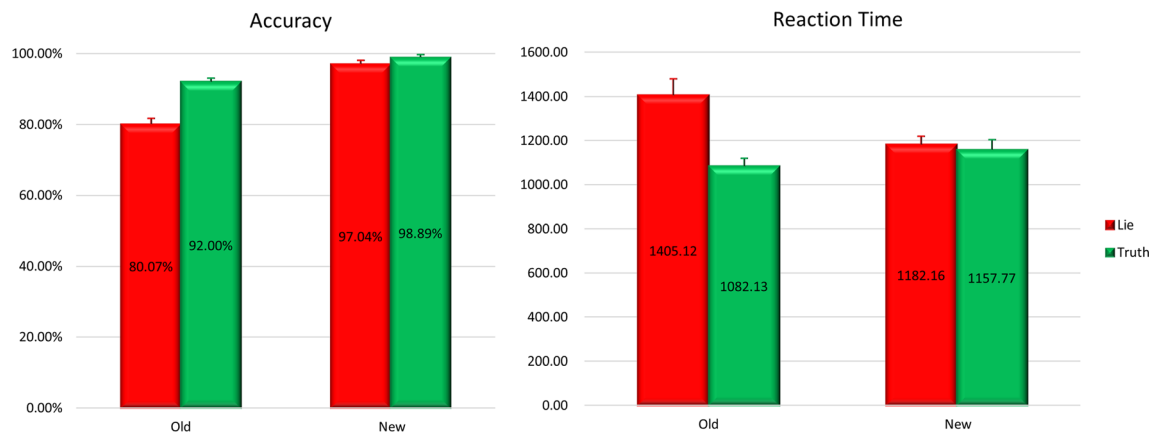


Fig. 3 Behavioral results of the memory-based task. The accuracy and RT from the lie old condition were significantly different from all other conditions.

.344; Pearson and Filon's z : $z = -3.6179$, p value = .0003]. The correlation remained significant if we eliminated the participant from the upper left corner (of Fig. 4) using a more stringent outlier criterion [$r = -.467$, $p = .016$]. These results suggest that people with higher verbal WM capacity are faster liars that have lower RT costs (Fig. 4).³ Note that the highest score in the verbal WM task was a score of 25; therefore, the correlation might have been stronger if some of the high performers could have scored higher for operation score.

In terms of accuracy, the overall accuracy of the main task was not correlated with any of the WM scores. However, if we look at the progression of accuracy from Trial 1 to Trial 6 in each block, it is quite clear that accuracy started to decrease from Trials 4 to 6 (Fig. 5, upper panel), presumably because participants had exhausted those truth-based words they could remember well. Therefore, later trials in the block, such as Trials 4–6, might be more telling in terms of whether an individual can hold multiple accounts or scripts in his or her mind. Indeed, a significant correlation between accuracy and verbal WM [$r = .383$, $p = .049$], and not for visual WM [$r = .157$, $p = .434$; Pearson and Filon's z : $z = 2.1122$, $p = .0347$], emerged when we looked only at the accuracy from Trials 4 through 6 (Fig. 5, lower panel).

ERP results

Each trial was divided into two stages: preparation and execution. These two stages were time-locked to cue onset and target onset, respectively. The assumption was that during the preparation stage, participants had to come up with a few legitimate words in the lie condition, and in the executions stage they had to look at the target first and then decide

³ In the smaller sample ($n = 24$), we observed a significant correlation between participants' RTs from the lie old condition and their operation span performance [$r = -.521$, $p = .009$], but not their symmetry performance [$r = .015$, $p = .946$]. The differences in RTs between the lie old and truth old conditions was also negatively correlated with participants' operation span scores [$r = -.593$, $p = .002$].

whether to execute their prepared deceptive response or to tell the truth.

Preparation stage To explore the neural processes behind the preparation stage, a repeated measures $2 \times 3 \times 3$ ANOVA with cue type (lie vs. truth), anterior–posterior electrode (frontal vs. central vs. parietal regions), and laterality (left vs. middle vs. right) as within-subjects factors was conducted. The Geisser–Greenhouse correction was applied to repeated measures with more than one degree of freedom. The main effect of cue type, the interaction between cue type and laterality, and the three-way interaction were not statistically significant. But there was a significant interaction between cue type and anterior–posterior electrode [$F(2, 46) = 7.782$, $p = .004$]. Separate comparisons of cue type showed that the amplitude of negative-going waveforms was larger in the lie condition than in the truth condition between 500 and 900 ms (Fig. 6, highlighted window) after lie–truth cue onset, but only in the frontal regions [frontal: $t(23) = 2.137$, $p = .043$; central: $t(23) = -0.736$, $p = .469$; parietal: $t(23) = -0.657$, $p = .518$], which consisted of F3, Fz, and F4. This possibly reflects the WM processes of maintaining multiple old words before the participants

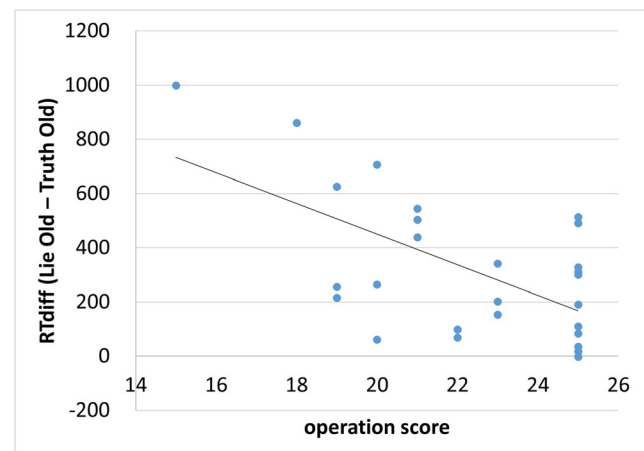


Fig. 4 Correlation between the RT cost of lying and verbal WM capacity.

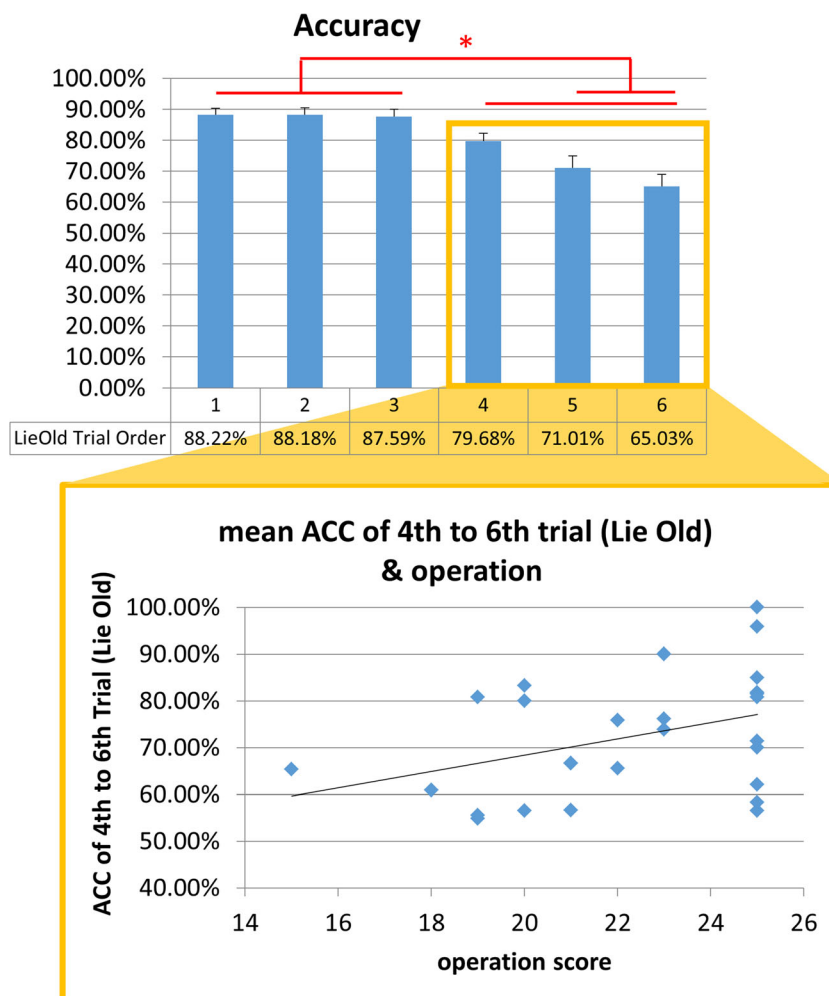


Fig. 5 Correlation between mean accuracy from Trials 4–6 and verbal WM capacity. This correlation suggests that participants with higher verbal WM capacity can lie with fewer errors.

actually see the target word, although we did not observe any significant correlation between the frontal amplitude and participants' verbal WM capacity (frontal: $r = -.248$, $p = .243$).

Execution stage For the execution stage, we mainly compared waveforms between the lie old and truth old conditions, excluding the catch trials. We found a significant main effect of cue type [$F(1, 23) = 4.504$, $p = .045$], due to a more positive amplitude in the lie old condition than in the truth old condition, from 300 to 550 ms after target onset. The interaction between cue type and anterior–posterior electrode was also significant [$F(2, 46) = 19.285$, $p < .001$], mainly because the lie–truth difference in the 300-ms to 550-ms window was mostly prominent in the frontal [frontal: $t(23) = -2.981$, $p = .007$] and central [central: $t(23) = -2.494$, $p = .020$] regions, but not in the parietal region [parietal: $t(23) = -0.555$, $p = .584$] (Fig. 7, left panel). Both the two-way interaction between cue type and laterality and the three-way interaction were not statistically significant [two-way: $F(2, 46) = 0.830$, $p = .436$; three-way: $F(4, 92) = 2.206$, $p = .109$].

In contrast with the frontal amplitude from the preparation stage, here we actually observed a significant correlation between participants' frontal and central amplitude and their behavioral performance (Fig. 7, right panel). Specifically, frontal and central amplitudes within the 300- to 550-ms window were used for the analysis, and they significantly correlated with the RT cost associated with lying (frontal: $r = -.426$, $p = .038$; central: $r = -.441$, $p = .031$), as well as with participants' verbal WM capacity measured by the operation span task (frontal: $r = .417$, $p = .043$; central: $r = .405$, $p = .049$). It is worth noting, however, that the correlation with the RT cost was not observed when we used the difference in ERP amplitudes between the lie old and truth old conditions (frontal: $r = -.086$, $p = .691$; central: $r = -.038$, $p = .861$). Together, the positive correlation between ERP amplitude and WM capacity and the negative correlation between ERP amplitude and RT cost seem to suggest that participants with larger frontal and central amplitudes in the lying condition were high performers who were faster in generating a deceptive response.

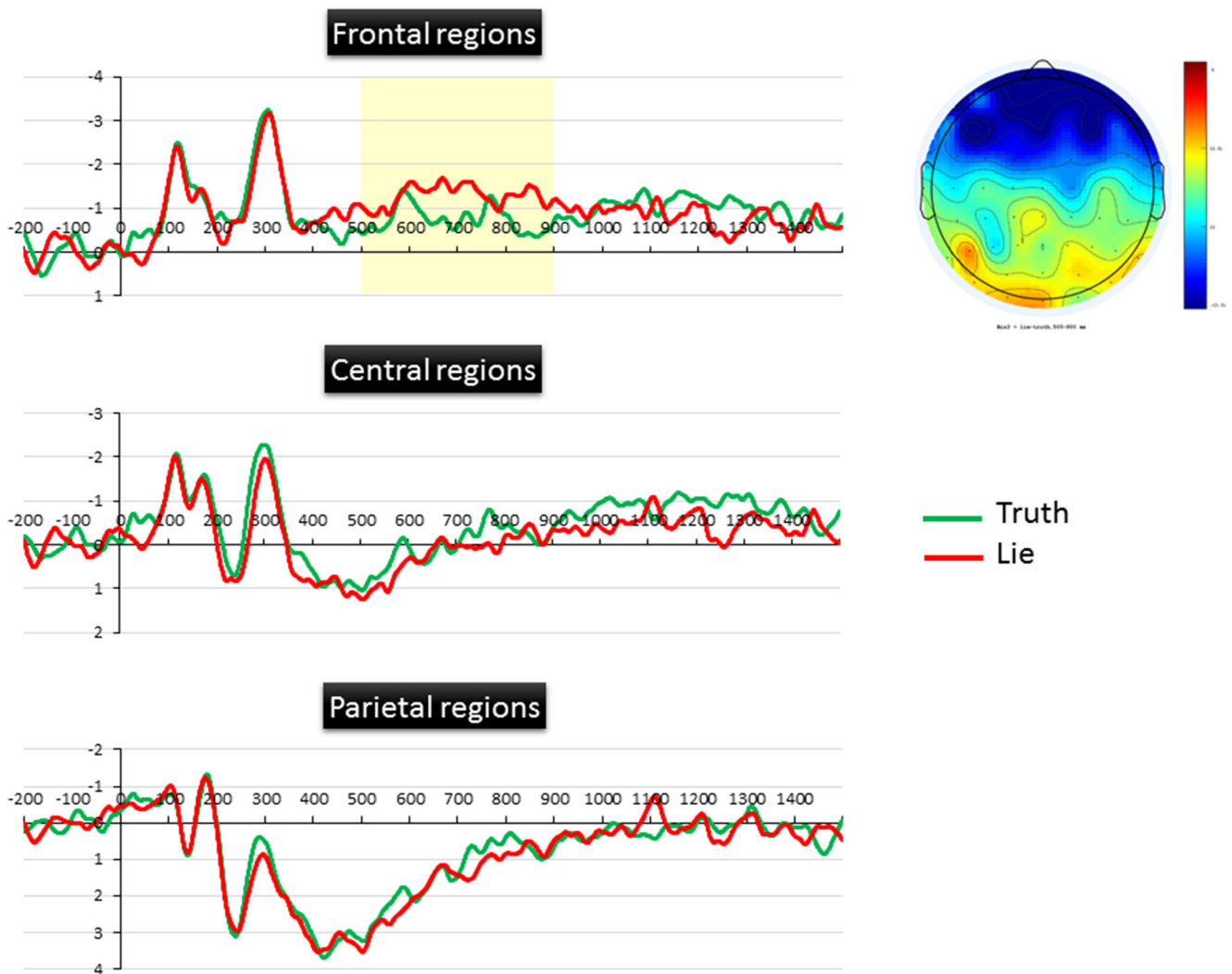


Fig. 6 Waveforms in the lie and truth conditions during the preparation stage. All graphs depict averaged amplitudes from three electrodes (e.g., F3, FZ, and F4 for the frontal region). A significant difference between the two conditions occurred at the frontal region between 500 and 900 ms.

Old–new recognition memory effect Although old–new recognition memory was not the main focus of the present study, our design did allow for old–new comparisons, to see whether memory traces existed for previously studied items. To this end, we analyzed waveforms between the old and new conditions to see whether previously seen words could still be traced electrophysiologically.

We first compared the truth old and truth new conditions via a three-way repeated measures ANOVA with old–new condition, anterior–posterior electrode (frontal vs. central vs. parietal region), and laterality (left vs. middle vs. right) as within-subjects factors. There was a significant old–new main effect [$F(1, 23) = 5.596, p = .027$], a significant interaction between anterior–posterior electrodes and old–new condition [$F(1, 23) = 26.917, p < .001$], and a significant three-way interaction [$F(4, 92) = 5.164, p = .002$]. The interaction between old–new and laterality was not statistically significant [$F(1, 23) = 0.101, p = .887$]. In contrast to the WM-related

frontal effects reported above, separate comparisons showed that the words from the old and new conditions were statistically different at parietal (Fig. 8), but not at frontal and central, regions within the 350- to 650-ms window [P3: $t(23) = 5.092, p < .001$; PZ: $t(23) = 4.247, p < .001$; P4: $t(23) = 3.690, p = .001$]. This contrast was particularly strong at P3, which was similar to the left parietal old–new ERP effect in the memory-retrieval literature (e.g., Curran, 2000).

The same three-way ANOVA was also conducted for the lie old versus either the lie new or the truth new condition. Here we found a significant main effect of condition [vs. lie new: $F(1, 23) = 14.542, p = .001$; vs. truth new: $F(1, 23) = 11.822, p = .002$], a significant interaction between condition and anterior–posterior electrode [vs. lie new: $F(1, 23) = 7.582, p = .007$; vs. truth new: $F(1, 23) = 6.163, p = .016$], and finally, a significant three-way interaction [vs. lie new: $F(4, 92) = 2.896, p = .042$; vs. truth new: $F(4, 92) = 3.349, p = .021$]. In the cases of both interactions, separate comparisons showed

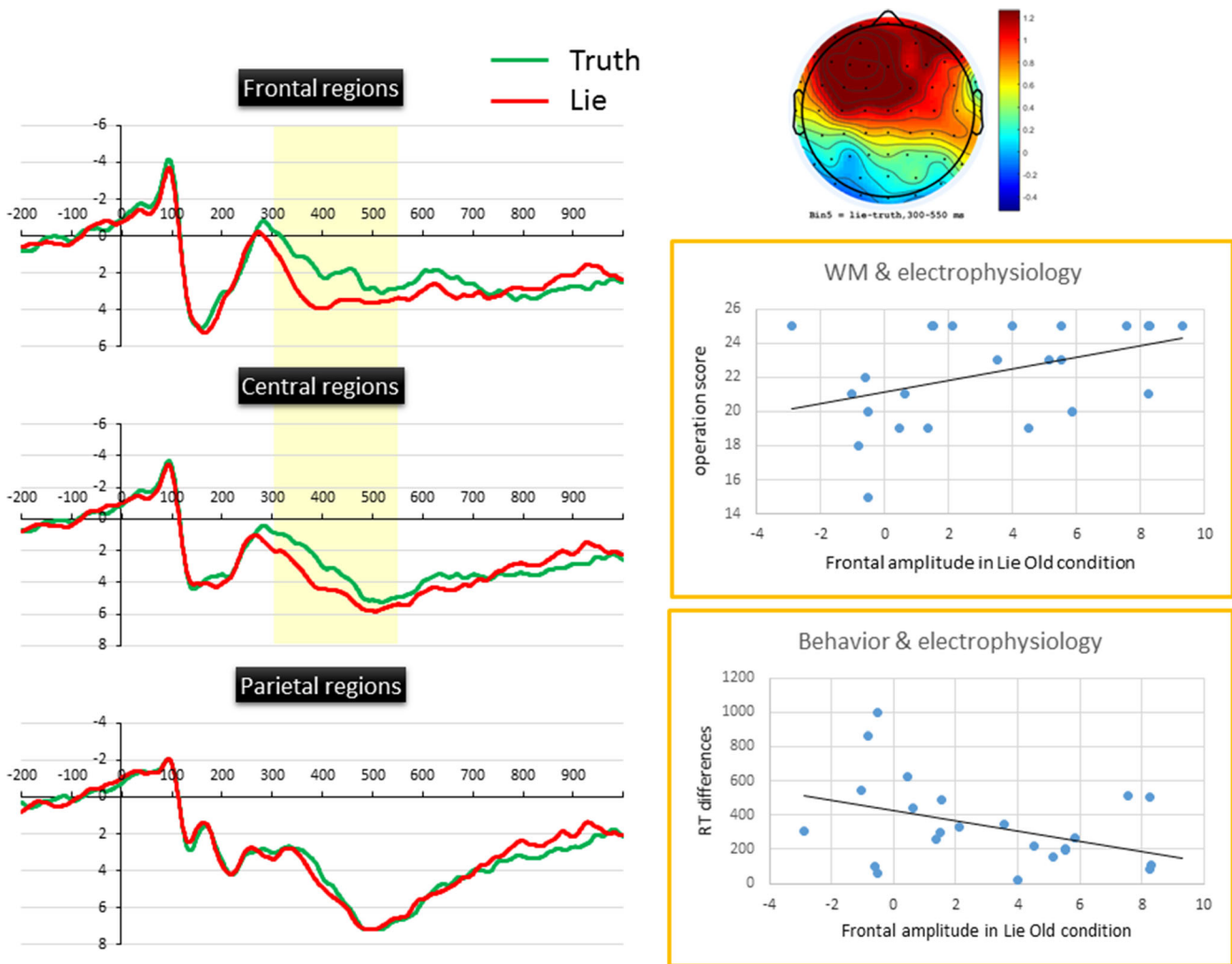


Fig. 7 Waveforms during the execution stage (left panel) and their correlations with participants' verbal WM capacity (right panel). All graphs and correlations depict averaged amplitudes from three electrodes (e.g., F3, FZ, and F4 for the frontal region). There is a notable difference in ERP amplitudes between the lie and truth

conditions between the 300- and 550-ms marks, and such amplitude differences could be predictive of participants' behavioral performance: People with larger waveform amplitudes had smaller RT costs when lying, and vice versa.

that the strongest difference between old and new within the 350- to 650-ms window occurred at P3 [vs. lie new: $t(23) = 5.357$, $p < .001$; vs. truth new: $t(23) = 4.161$, $p < .001$]. Therefore, regardless of which new condition (truth new or lie new) was used to contrast with the lie old condition, the left parietal old–new effect was quite robust, implicating a GKT-like recognition memory effect upon seeing a familiar target word.

Importantly, was this recognition memory effect sufficient to pick out deceptive responses in the present paradigm? We compared the amplitudes at parietal regions between the lie old and truth old conditions. However, no significant lie–truth difference emerged at any of the parietal sites. There was also no significant correlation between parietal amplitudes in the 350- to 650-ms window and participants' operation or symmetry span WM scores. These results are consistent with the

rationale of the GKT, and perhaps highlight its strength and weakness, as well as the need for the present paradigm. Namely, the GKT is great at separating recognized versus unrecognized targets, but it is unsuitable when deception is based on preexisting or truth-based knowledge.

rationale of the GKT, and perhaps highlight its strength and weakness, as well as the need for the present paradigm. Namely, the GKT is great at separating recognized versus unrecognized targets, but it is unsuitable when deception is based on preexisting or truth-based knowledge.

Discussion

In this study, we investigated the electrophysiological signatures of a simplified form of truth-based deception during its preparation stage and execution stages: Participants had to lie about the identity of a prestudied word by using another prestudied word. This scenario was designed to mimic situations in which recognition memory is confirmed via the GKT, but suspects can still provide false accounts based on

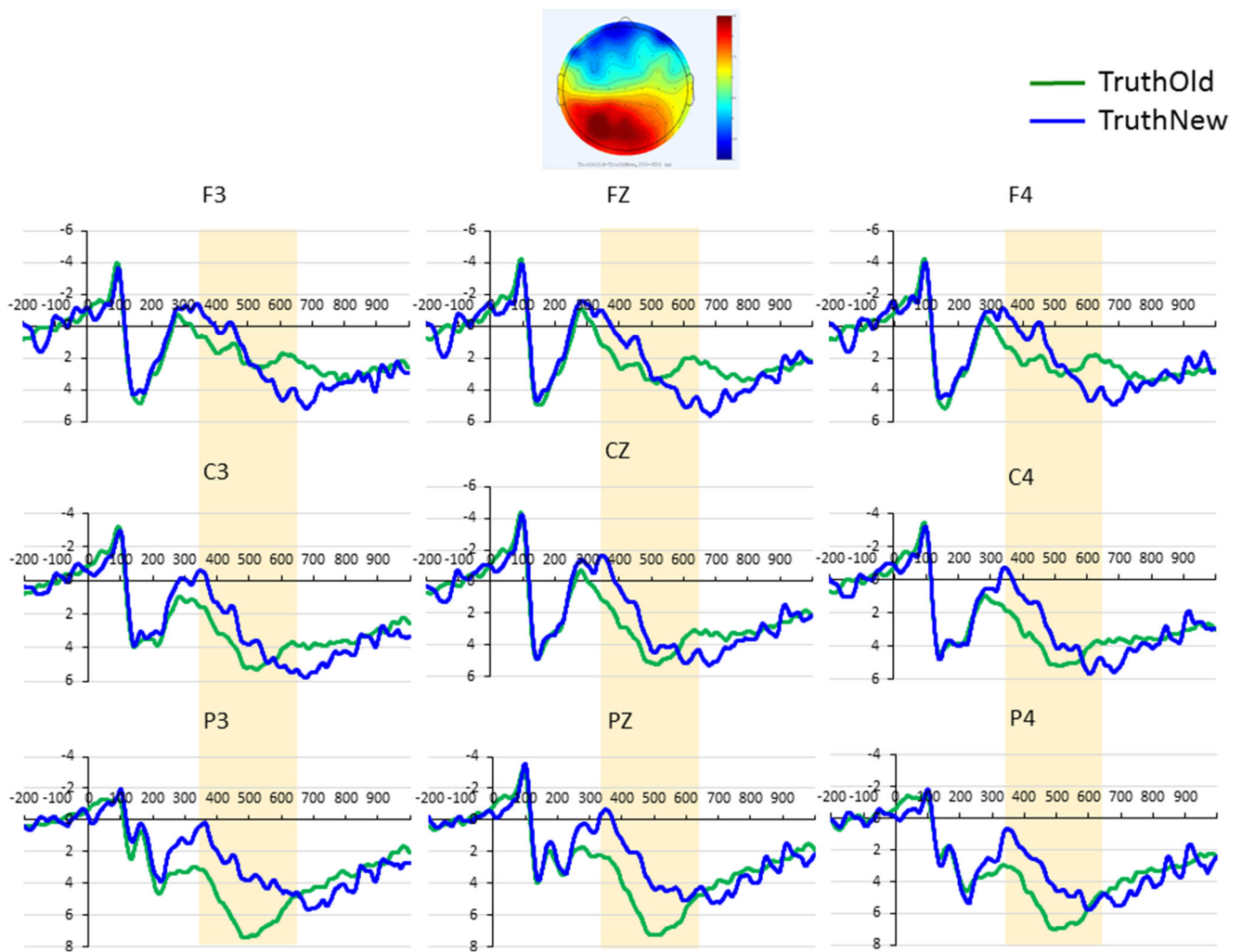


Fig. 8 Waveforms from the truth old and truth new conditions. The highlighted area depicts the time window between 350 and 650 ms after target onset. The brain topographic distribution at the top shows the contrast between the averaged amplitudes of the truth old and truth

news conditions. The hot topographic distribution shows that the amplitude from the truth old condition is larger than those from the truth new condition.

recognizable (truth-based) events, in the interest of minimizing legal responsibility or misleading an investigation. Indeed, we observed equally strong recognition memory ERP waveforms in both the lie old and truth old conditions, suggesting that the GKT is great at what it does, but it is unsuitable for truth-based deception.

Our behavioral results showed that deceptive responses had slower RT and lower accuracy than truth responses. Interestingly, this RT difference between lies and truth was correlated with participants' verbal WM capacity such that participants with high WM capacity could lie with very little RT cost. For these participants, we argue it is essential to look into their frontal EEG signals, in which deceptive responses had a larger negative-going waveform during the preparation stage, and a larger positive-going waveform during the execution stage. Importantly, these frontal amplitudes during the execution stage were negatively correlated with participants'

RT cost for deceptive responses. In other words, high WM capacity individuals who can verbalize a deceptive response efficiently are also the ones who showed larger frontal amplitude during the execution stage, suggesting a possible signature of WM usage that can potentially be used to detect truth-based deception.

The idea of combining WM load and deception is not new as previous studies have successfully induced cognitive load via dual tasking to reveal behavioral cues to deception (for a review, please see Walczyk et al., 2013). The novelty and uniqueness of the present paradigm are (1) passive EEG measures that tracks each participant's WM usage instead of actively inducing WM load, thus providing neurophysiological indices for individual differences in lying efficiency, and (2) the well-defined boundary between the preparation and execution stages, which we discuss separately in more details below.

Recognition memory effect

Although not a primary focus of the present study, our experimental design did include the component of recognition memory, and thus an old–new contrast can be conducted. To this end, we observed significant old–new ERP difference between the truth old and truth new condition, as well as between the lie old and the lie new (or truth new) condition. This difference was most robust at P3, which is similar to the left parietal old–new ERP effect in memory-retrieval literature (Curran, 2000). Previous studies have suggested that left parietal old–new effect is an electrophysiological index of recognition and is a positive-going ERP component that appears around 400 to 500 ms after stimulus onset (Rugg & Curran, 2007), induced by correctly recognized old items (Curran, 2000). This positivity is greater over the left hemisphere, and is greater for hits and misses than new items (Paller & Kutas, 1992; Rugg & Nagy, 1989; Rugg et al., 1998). Furthermore, this marker of recognition memory is recollection-based (Curran, 2004; Herron & Rugg, 2003), and therefore can be observed even when no deceptive response is required (Gamer et al., 2009). This can be useful in detecting simple-denial type of deception.

However, further contrasts in our data showed that this component was no different between the lie old and truth old conditions, suggesting it is indeed a recognition memory effect regardless of deceptive intent. This observation is similar to previous findings by Johnson et al. (2003), who found intact parietal memory effect regardless of participants' intention to deceive or tell the truth, suggesting a recognition memory-based ERP measure that is similar to the rationale of GKT. Interestingly, its amplitude also did not correlate with any of the WM measures, which implies that our task was only tapping into the WM, but not episodic memory, aspect of deception—a direction that is worth pursuing for future studies. Therefore, although recognition memory paradigms are great for denial type of lies, fabrication (involving no recognition) or truth-based lies (recognition is inevitable) may require other types of measures, such as WM usage.

Preparation stage

In this experiment, we hypothesized that participants would show pronounced ERP differences during the preparation stage due to the demand of keeping several plausible answers in their WM. Indeed, our preparatory ERP results showed a larger frontal amplitude in the lie condition than in the truth condition. This difference is significant from 500 to 900 ms after cue onset, although such difference did not correlate with participants' WM capacity like the data from the execution stage did, which we discuss in more details below.

The present results are consistent with Walczyk et al.'s (2003) activation–decision–construction model, in which the construction component involves long-term memory and attention during the preparation stage. In this model, deception requires retrieving semantic and episodic information from long-term memory to generate a plausible lie with credible details, whereas truth-telling simply requires long-term memory retrieval without the additional manipulation or reassociation between different memory information (which would take place in WM). Therefore, according to Walczyk et al.'s (2003) model, preparation processes for deception and truth-telling would be quite different, which is translated to difference in WM load (i.e., recalling more words in test phase) in the present paradigm.

Previous neuroimaging studies have associated elevated prefrontal activities with task preparation and information maintenance for responding to upcoming stimuli (Ito et al., 2012; Sakai & Passingham, 2006). Here, the observation of frontal negative wave during preparation is also consistent with findings of ERP negative slow wave, an index of WM operations (Mecklinger & Pfeifer, 1996; Ruchkin, Canoune, Johnson, & Ritter, 1995). Ruchkin et al. used a delayed match-to-sample task to investigate encoding and retention processes in WM. They found a negative slow wave throughout the interval in between the memory frame and the test frame. Furthermore, this negative slow wave activity varied with memory load. Mecklinger and Pfeifer used an object memory task with three levels of memory load and found that amplitude of the negative slow wave varied as a function of object memory load near the frontal and central regions during memory retention. Therefore, in the context of our study, frontal negative amplitude may be an electrophysiological index of WM traces during the preparation stage of deception.

Previous studies have repeatedly shown a larger CNV component around the 1,000-ms window, before participants had to provide their deceptive responses. The meaning of CNV has long been interpreted as a reflection of anticipation and motivation, and in the context of deception it has been interpreted as anticipation and response preparation. This was true for a variety of different tasks, including scenarios in which participants had to deny recognition of familiar target faces (Fang et al., 2003) or knowledge of a mock crime (Suchotzki et al., 2015), identify fake banknotes as genuine (Sun et al., 2011), or rate an unattractive face as attractive (Dong et al., 2010). In the present task, however, we did not observe similar component during the preparation stage, in which the participants have yet to give their response. To account for this inconsistency, we will note three important differences. First, CNV is usually observed around the 1,000-ms time window, but our preparation window was too long (for WM rehearsal purposes) for a conventional CNV interval, which may have buried the CNV component due to WM demand. Second, and perhaps more important, CNV has been

suggested as a combination of anticipatory processes coupled with motor preparation (Rohrbaugh & Gaillard, 1983), including the activation of the hand-motor area (Verleger et al., 2000). As such, other researchers have argued that the anticipatory effect should be examined before participants can see the critical target stimulus. Indeed, when Suchotzki et al. used such design, they only observed a marginally significant CNV related to deception, with a smaller effect size. In this light, our design is quite similar to that of Suchotzki et al. because our participants did not know whether a deceptive (lie old condition) or truthful (lie new condition) response should be performed during the preparation stage (see also Panasiti et al., 2014, for a card game task that did not show any effect in CNV). Therefore, there was no clear response anticipation during our preparation window. And finally, our study utilized a vocal response that did not require activation of the hand-motor area, thus it is reasonable that the centroparietal component of CNV would be absent from the present findings.

One remaining question, however, is that if the observed frontal amplitude between the 500 to 900-ms window is indicative of the WM processes underlying deception preparation, why does it not correlate with participant's WM capacity and lying efficiency? This lack of correlation suggests that participants were not exhausting their WM capacity to prepare as many old words as possible. Rather, they maintained just enough words in WM, below their capacity limit, during the preparation stage for selection during the execution stage. Therefore, in any given trial, their WM load would be kept at a minimal but sufficient set size, as opposed to being maxed out. This strategy becomes gradually harder to implement toward the end of the block (when all the well-remembered words have been used), as evidenced by the decreasing accuracy from Trials 4 to 6 due to WM exhaustion. When this happens, the correlation between one's accuracy and WM capacity emerges (Fig. 5). Therefore, the strategy of keeping WM load low in the first half of the block can possibly explain why the ERP waveform itself during the preparation stage was not correlated with participants' end-result behavioral performance.

These positive findings in frontal ERP during the preparation stage mark an interesting contrast to one important study by Ito et al. (2012), who used a modified recognition memory task and fMRI to investigate processes during the preparation period. They used three types of cues in their task: truth, lie, and uncertain. And the authors found a significant activation in the left dorsolateral prefrontal cortex during the preparation stage that followed an informative (i.e., truth and lie) cue over noninformative (i.e., uncertain) cue. However, preparatory processes that followed truth and lie cues were no different in their BOLD fMRI signals. Although these results may appear to conflict with the data presented here, it is important to note the differences in task design, and precisely what processes are being probed when using the term "preparation." In

Ito et al.'s (2012) study, participants did not have to search their memory after seeing an informative cue, and they had to make an old–new response upon seeing the target picture, which is similar to a GKT design. In contrast, the present study required participants to prepare several seen words in advance. Therefore, our task design focused mostly on WM usage during the preparation stage, whereas Ito et al.'s (2012) design prompted deceptive intention and early initiation of inhibitory processes (of the truth).

Our results from the preparation time window are also consistent with one recent fMRI study by Ofen, Whitfield-Gabrieli, Chai, Schwarzlose, and Gabrieli (2017), who found increased superior parietal and dorsolateral cortex activity during the preparation stage. Importantly, dorsolateral cortical activities were negatively correlated with participants' deception cost, which is similar to the correlation we have reported here between deception cost and frontal amplitude, albeit at a later time window (i.e., executive time period). Therefore, not only can fMRI or ERP measures be potentially useful for detecting deception, the distribution of such neural activities also seem to correlate directly with participants' individual differences in their abilities to lie efficiently.

Execution stage

During the execution stage of verbalizing a lie, we observed larger positive-going amplitude near the frontal and central sites in the 300- to 550-ms window. Most importantly, such amplitude was positively correlated with participants' WM capacity that was independently measured via the operation span task, and it also negatively correlated with participants' deception cost in RT. In other words, participants with higher WM capacity also showed higher frontal-central ERP amplitude, and lower lie–truth RT cost.

The predictive nature of the ERP amplitudes during the execution stage but not the preparation stage is somewhat surprising. The frontal-central amplitude during the execution stage was correlated with participants' WM capacity and deception cost in RT, yet no correlation was observed between these measures and preparatory frontal amplitude. Earlier we mentioned how participants could maintain just enough words in WM during the preparation stage for selection later in the execution stage. Here, during the execution stage, participants' selection of deceptive words must be rapidly completed as the target word appears, which involves WM selection. Furthermore, in some trials, it is possible that the target word coincides with the ones participants have prepared for verbalizing. In these scenarios, participants had to suppress the pre-prepared old word and quickly switch to another backup old word for lying. Furthermore, in a small portion (<10%) of the trials, upon seeing a new word, participants had to give up the rehearsed words and instead say the word on the screen as is (lie new condition). Therefore, in addition to the selection

processes, suppression and switch processes could also have taken place rapidly in order to minimize cost in RT. Consequently, although the WM demand may not be very different between the preparation and the execution stages, the time pressure (and thus the increased WM or mental effort) is nevertheless higher during the execution stage. Indeed, in the last decade the WM literature has begun to highlight the importance of WM selection and inhibition processes in addition to just the maintenance of WM content (e.g., Tseng et al., 2012; Vogel, McCollough, & Machizawa, 2005; Xu, 2010). As such, a person can have large WM capacity but still suffer from poor WM performance due to inapt suppression of task-irrelevant WM content (Vogel et al., 2005). These findings and the multiprocess view of WM integrity shed light on why ERP correlation actually occurred during the execution and not the preparation stage, and also explains why participants' WM capacity can be used to predict their RT cost during this stage. In this light, our preparation stage is cognitively "purer" in the sense that only WM rehearsal is taking place at that time, and the execution stage is more akin to the WM retrieval stage that other researchers have used in order to investigate the processes of WM selection (and suppression of other words) (e.g., Shipstead, & Engle, 2013). This task design was done intentionally to mimic the many decisions a liar has to make upon hearing the investigator's questions (e.g., should the interrogatee use script A or script B, or maybe tell the truth, etc.), and is where we observed the most robust effect in frontal amplitude, reaction time, and their correlations with participants' verbal WM operation span performance. These results from the execution stage (and the absence of any correlation during the preparation stage) suggest additional neural mechanisms beyond memory maintenance that are at play during the execution stage, and supports the idea that inhibitory processes, whether in WM or other cognitive domains, can be crucial to successful deception (e.g., Debey, Verschuere, & Crombez, 2012; Nunez, Casey, Egner, Hare, & Hirsch, 2005; Walczyk et al., 2003).

Our results here also complement previous findings by Johnson, Bamhardt, and Zhu (2003, 2004, 2005) well. In their series of studies, Johnson and colleagues used a recognition memory-based deception task, instead of a recall-based task in the present study. They found slower RT and lower accuracy, coupled with increased medial frontal negativities (Johnson et al., 2004) and decreased P300 amplitude (Johnson et al., 2003), which they interpreted as a reflection of ongoing monitoring processes that is specific to deception. Specifically, Johnson et al. (2004) also identified the anterior cingulate cortex as a possible dipole for the increased medial frontal negativities, and such activity remained unaffected even if participants went through repeated practice (Johnson et al., 2005). This would be consistent with our effects from the frontal regions. Therefore, although our tasks used a recall memory task (as opposed to the recognition memory task by

Johnson et al.), and that our task was not designed to optimize the observation of medial frontal negativities, the present study is nevertheless similar to the Johnson et al. studies in the sense that the frontal region and its executive functions are heavily involved.

The findings here provide further insight to previous studies that did not differentiate between the preparation and execution stage of deception. One important study by Tardif et al. (2000) tested between the control and a malingering group using an old–new memory paradigm. They observed larger left frontal amplitude between old and new words in the 0- to 450-ms window after stimuli onset in the malingering group than the control group. This difference, however, was not associated with a well-defined separation between the preparation and execution stages. Comparing their time windows with our 500- to 900-ms window and 300- to 550-ms window from the preparation and execution stages, respectively, it seems that our execution processing timeframe fits better with their time series. Also, in our results, the frontal-central topography and the 300- to 550-ms latency are quite similar to the typical P3a component. The P3a component is suggested to represent frontal attentional processing that is engaged in evaluating current stimuli (Hagen, Gatherwright, Lopez, & Polich, 2006; Polich, 2007), which is quite consistent with our experimental setup during the execution stage, in which participants saw the target word. One previous study, by Hagen et al., found that when task difficulty increased, the P3a amplitude increased as well, which is also consistent with the WM and word-switch demand that we discussed earlier.

The promising finding from the present study is perhaps the correlation between this possible P3a component and participants' deception efficiency (i.e., lower RT cost). This implies that, in those high performers for whom the behavioral cues of deception may not be immediately obvious, they nonetheless are the ones with the largest EEG signature of deception. Indeed, one previous study by Yang et al. (2005) reported that pathological liars had higher verbal intelligence and increased prefrontal white matter, compared with normal and antisocial controls. Therefore, people with better verbal intelligence, such as our participants who scored high on the operation span task, may be harder to probe behaviorally. For these high-performers, we argue that execution-stage ERP amplitudes are especially important for developing potential indicators of deception. Though, it is important to bear in mind that countermeasures, as well as repeated rehearsal of lies in advance, can possibly compromise the integrity of the present ERP findings (Foerster et al., 2017). Since most perpetrators prepare and practice sharing alibis in advance so that they are internally and externally consistent, multimodal, and delivered with minimal WM load. Therefore, future research needs to consider a wider range of variables and methods in order to improve the EEG markers accordingly.

Author note This study was sponsored by a Taiwan Ministry of Education “Academic Strategic Alliance: Taiwan and Oxford University” project grant, as well as by grants from the Ministry of Science and Technology (MOST), Taiwan (MOST 104-2420-H-038-001-MY3; MOST 104-2410-H-038-013-MY3; MOST 106-2410-H-038-005-MY2) and from Taipei Medical University (TMU 104-AE1-B07; 105 TMU-SHH-20).

References

- Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, 4, 417–423. [https://doi.org/10.1016/S1364-6613\(00\)01538-2](https://doi.org/10.1016/S1364-6613(00)01538-2)
- Browndyke, J.N., Paskavitz, J., Sweet, L.H., Cohen, R.A., Tucker, K.A., Welsh-Bohmer, K.A., Burke, J.R., & Schmechel, D.E. (2008). Neuroanatomical correlates of malingered memory impairment: event-related fMRI of deception on a recognition memory task. *Brain Injury*, 22(6), 481–489.
- Browndyke, J. N., Paskavitz, J., Sweet, L. H., Cohen, R. A., Tucker, K. A., Welsh-Bohmer, K. A., ... Schmechel, D. E. (2010). *Brain Injury*, 22, 481–489.
- Chadwick, M. J., Hassabis, D., Weiskopf, N., & Maguire, E. A. (2010). Decoding individual episodic memory traces in the human hippocampus. *Current Biology*, 20, 544–547.
- Chen, N.-F., Lo, C.-M., Liu, T.-L., & Cheng, S.-K. (2016). Source memory performance is modulated by transcranial direct current stimulation over the left posterior parietal cortex. *NeuroImage*, 139, 462–469.
- Curran, T. (2000). Brain potentials of recollection and familiarity. *Memory & Cognition*, 28, 923–938. <https://doi.org/10.3758/BF03209340>
- Curran, T. (2004). Effects of attention and confidence on the hypothesized ERP correlates of recollection and familiarity. *Neuropsychologia*, 42, 1088–1106.
- Debey, E., Verschuere, B., & Crombez, G. (2012). Lying and executive control: An experimental investigation using ego depletion and goal neglect. *Acta Psychologica*, 140, 133–141. <https://doi.org/10.1016/j.actpsy.2012.03.004>
- DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin*, 129, 74–118. <https://doi.org/10.1037/0033-2909.129.1.74>
- Dong, G., Wu, H., & Lu, Q. (2010). Attempting to hide our real thoughts: Electrophysiological evidence from truthful and deceptive responses during evaluation. *Neuroscience Letters*, 479, 1–5. <https://doi.org/10.1016/j.neulet.2010.05.014>
- Fang, F., Liu, Y., & Shen, Z. (2003). Lie detection with contingent negative variation. *International Journal of Psychophysiology*, 50, 247–255.
- Farwell, L. A., & Donchin, E. (1991). The truth will out: Interrogative polygraphy (“lie detection”) with even-related brain potentials. *Psychophysiology*, 28, 531–547.
- Foerster, A., Wirth, R., Herbold, O., Kunde, W., & Pfister, R. (2017). Lying upside-down: Alibis reverse cognitive burdens of dishonesty. *Journal of Experimental Psychology: Applied*, 23(3), 301.
- Foster, J. L., Shipstead, Z., Harrison, T. L., Hicks, K. L., Redick, T. S., & Engle, R. W. (2015). Shortened complex span tasks can reliably measure working memory capacity. *Memory & Cognition* 43, 226–236. <https://doi.org/10.3758/s13421-014-0461-7>
- Gamer, M., Bauermann, T., Stoeter, P., & Vossel, G. (2007). Covariations among fMRI, skin conductance, and behavioral data during processing of concealed information. *Human Brain Mapping*, 28, 1287–1301.
- Gamer, M., Klimecki, O., Bauermann, T., Stoeter, P., & Vossel, G. (2009). fMRI-activation patterns in the detection of concealed information rely on memory-related effects. *Social cognitive and affective neuroscience*, 7(5), 506–515.
- Granhag, P. A., & Hartwig, M. (2008). A new theoretical perspective on deception detection: On the psychology of instrumental mind-reading. *Psychology, Crime & Law*, 14, 189–200. <https://doi.org/10.1080/10683160701645181>
- Gratton, G., Coles, M. G. H., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, 55, 468–484.
- Hagen, G. F., Gatherwright, J. R., Lopez, B. A., & Polich, J. (2006). P3a from visual stimuli: Task difficulty effects. *International Journal of Psychophysiology*, 59, 8–14.
- Herron, J. E., & Rugg, M. D. (2003). Strategic influences on recollection in the exclusion task: Electrophysiological evidence. *Psychonomic Bulletin & Review*, 10(3), 703–710.
- Huang, C.-R., Ahrens, K., & Chen, K.-J. (1998). A data-driven approach to the mental lexicon: Two studies on Chinese corpus linguistics. *Bulletin of the Institute of Mathematics Academia Sinica, New Series*, 69, 151–180.
- Ito, A., Abe, N., Fujii, T., Hayashi, A., Ueno, A., Mugikura, S., ... Mori, E. (2012). The contribution of the dorsolateral prefrontal cortex to the preparation for deception and truth-telling. *Brain Research*, 1464, 43–52.
- Johnson, R., Barnhardt, J., & Zhu, J. (2003). The deceptive response: Effects of response conflict and strategic monitoring on the late positive component and episodic memory-related brain activity. *Biological Psychology*, 64, 217–253. <https://doi.org/10.1016/j.biopsycho.2003.07.006>
- Johnson, R., Barnhardt, J., & Zhu, J. (2004). The contribution of executive processes to deceptive responding. *Neuropsychologia*, 42, 878–901. <https://doi.org/10.1016/j.neuropsychologia.2003.12.005>
- Johnson, R., Barnhardt, J., & Zhu, J. (2005). Differential effects of practice on the executive processes used for truthful and deceptive responses: An event-related brain potential study. *Cognitive Brain Research*, 24, 386–404. <https://doi.org/10.1016/j.cogbrainres.2005.02.011>
- Leins, D. A., Fisher, R. P., & Ross, S. J. (2013). Exploring liars’ strategies for creating deceptive reports. *Legal and Criminological Psychology*, 18, 141–151.
- Mecklinger, A., & Pfeifer, E. (1996). Event-related potentials reveal topographical and temporal distinct neuronal activation patterns for spatial and object working memory. *Cognitive Brain Research*, 4, 211–224.
- Nunez, J. M., Casey, B. J., Egner, T., Hare, T., & Hirsch, J. (2005). Intentional false responding shares neural substrates with response conflict and cognitive control. *NeuroImage*, 25, 267–277.
- Ofen, N., Whitfield-Gabrieli, S., Chai, X. J., Schwarzlose, R. F., & Gabrieli, J. D. E. (2017). Neural correlates of deception: Lying about past events and personal beliefs. *Social Cognitive and Affective Neuroscience*, 12, 116–127.
- Paller, K. A., & Kutas, M. (1992). Brain potentials during memory retrieval provide neurophysiological support for the distinction between conscious recollection and priming. *Journal of Cognitive Neuroscience*, 4, 375–392. <https://doi.org/10.1162/jocn.1992.4.4.375>
- Panasiti, M. S., Pavone, E. F., Mancini, A., Merla, A., Grisoni, L., & Aglioti, S. M. (2014). The motor cost of telling lies: Electrocutaneous signatures and personality foundations of spontaneous deception. *Social Neuroscience*, 9, 573–589. <https://doi.org/10.1080/17470919.2014.934394>
- Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, 118, 2128–2148. <https://doi.org/10.1016/j.clinph.2007.04.019>
- Rohrbaugh, J. W., & Gaillard, A. W. K. (1983). 13 sensory and motor aspects of the contingent negative variation. *Advances in*

- Psychology*, 10, 269–310. [https://doi.org/10.1016/S0166-4115\(08\)62044-0](https://doi.org/10.1016/S0166-4115(08)62044-0)
- Rosenfeld, J. P., Angell, A., Johnson, M., & Qian, J. (1991). An ERP-based control-question lie detector analog: Algorithms for discriminating effects within individual waveforms. *Psychophysiology*, 28, 320–336.
- Rosenfeld, J. P., Labkovsky, E., Winograd, M., Lui, M. A., 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, 906–919.
- Rosenfeld, J. P., Nasman, V. T., Whalen, R., Cantwell, B., & Mazzeri, L. (1987). Late vertex positivity in event-related potentials as a guilty knowledge indicator: A new method of lie detection. *International Journal of Neuroscience*, 34, 125–129.
- Ruchkin, D. S., Canoune, H. L., Johnson, R., Jr., & Ritter, W. (1995). Working memory and preparation elicit different patterns of slow wave event-related brain potentials. *Psychophysiology*, 32, 399–410.
- Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences*, 11, 251–257.
- Rugg, M. D., Mark, R. E., Walla, P., Schloerscheidt, A. M., Birch, S., & Allan, K. (1998). Dissociation of the neural correlates of implicit and explicit memory. *Nature*, 392, 595–598.
- Rugg, M. D., & Nagy, M. E. (1989). Event-related potentials and recognition memory for words. *Electroencephalography and Clinical Neurophysiology*, 72, 395–406.
- Sakai, K., & Passingham, R. E. (2006). Prefrontal set activity predicts rule-specific neural processing during subsequent cognitive performance. *Journal of Neuroscience*, 26, 1211–1218.
- Seymour, T. L., Seifert, C. M., Shafto, M. G., & Mosmann, A. L. (2000). Using response time measures to assess “guilty knowledge”. *Journal of Applied Psychology*, 85, 30–37.
- Shipstead, Z., & Engle, R. W. (2013). Interference within the focus of attention: Working memory tasks reflect more than temporary maintenance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39, 277–289. <https://doi.org/10.1037/a0028467>
- Suchotzki, K., Crombez, G., Smulders, F. T. Y., Meijer, E., & Verschuere, B. (2015). The cognitive mechanisms underlying deception: An event-related potential study. *International Journal of Psychophysiology*, 95, 395–405. <https://doi.org/10.1016/j.ijpsycho.2015.01.010>
- Sun, S.-Y., Mai, X., Liu, C., Liu, J.-Y., & Luo, Y.-J. (2011). The processes leading to deception: ERP spatiotemporal principal component analysis and source analysis. *Social Neuroscience*, 6, 348–359. <https://doi.org/10.1080/17470919.2010.544135>
- Talwar, V., & Lee, K. (2002). Emergence of white-lie telling in children between 3 and 7 years of age. *Merrill-Palmer Quarterly*, 48, 160–181.
- Tardif, H. P., Barry, R. J., Fox, A. M., & Johnstone, S. J. (2000). Detection of feigned recognition memory impairment using the old/new effect of the event-related potential. *International Journal of Psychophysiology*, 36, 1–9.
- Tseng, P., Hsu, T. Y., Chang, C. F., Tzeng, O. J., Hung, D. L., Muggleton, N. G., ... & Juan, C. H. (2012). Unleashing potential: transcranial direct current stimulation over the right posterior parietal cortex improves change detection in low-performing individuals. *Journal of Neuroscience*, 32(31), 10554–10561.
- Unsworth, N., Fukuda, K., Awh, E., & Vogel, E. K. (2014). Working memory and fluid intelligence: Capacity, attention control, and secondary memory retrieval. *Cognitive Psychology*, 71, 1–26.
- van Hooff, J. C., Brunia, C. H. M., & Allen, J. J. B. (1996). Event-related potentials as indirect measures of recognition memory. *International Journal of Psychophysiology*, 21, 15–31.
- Verleger, R., Vollmer, C., Wauschkuhn, B., van der Lubbe, R. H., & Wascher, E. (2000). Dimensional overlap between arrows as cueing stimuli and responses?: Evidence from contra-ipsilateral differences in EEG potentials. *Cognitive Brain Research*, 10, 99–109.
- Verschuere, B., Suchotzki, K., & Debey, E. (2015). Detecting deception through reaction times. In P. A. Granhag, A. Vrij, & B. Verschuere (Eds.), *Deception detection: Current challenges and new approaches* (pp. 269–291). Oxford, UK: Wiley.
- Vogel, E. K., McCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature*, 438, 500–503. <https://doi.org/10.1038/nature04171>
- Vrij, A., Fisher, R., Mann, S., & Leal, S. (2006). Detecting deception by manipulating cognitive load. *Trends in Cognitive Sciences*, 10, 141–142.
- Vrij, A., Fisher, R., Mann, S., & Leal, S. (2008). A cognitive load approach to lie detection. *Journal of Investigative Psychology and Offender Profiling*, 5, 39–43.
- Vrij, A., & Granhag, P. A. (2012). Eliciting cues to deception and truth: What matters are the questions asked. *Journal of Applied Research in Memory and Cognition*, 1, 110–117. <https://doi.org/10.1016/j.jarmac.2012.02.004>
- Vrij, A., Mann, S., Fisher, R., Leal, S., Milne, B., & Bull, R. (2008). Increasing cognitive load to facilitate lie detection: The benefit of recalling an event in reverse order. *Law and Human Behavior*, 32, 253–265. <https://doi.org/10.1007/s10979-007-9103-y>
- Vrij, A., Mann, S., Leal, S., & Fisher, R. (2010). “Look into my eyes”: Can an instruction to maintain eye contact facilitate lie detection? *Psychology Crime and Law*, 16, 327–348.
- Walczyk, J. J., Igou, F. P., Dixon, A. P., & Tcholokian, T. (2013). Advancing lie detection by inducing cognitive load on liars: A review of relevant theories and techniques guided by lessons from polygraph-based approaches. *Frontiers in Psychology*, 4, 14. <https://doi.org/10.3389/fpsyg.2013.00014>
- Walczyk, J. J., Roper, K., Seemann, E., & Humphrey, A. M. (2003). Cognitive mechanisms underlying lying to questions: Response time as a cue to deception. *Applied Cognitive Psychology*, 17, 755–774.
- Wu, H., Hu, X., & Fu, G. (2009). Does willingness affect the N2-P3 effect of deceptive and honest responses? *Neuroscience Letters*, 467, 63–66. <https://doi.org/10.1016/j.neulet.2009.10.002>
- Xu, Y. (2010). The neural fate of task-irrelevant features in object-based processing. *Journal of Neuroscience*, 30, 14020–14028. <https://doi.org/10.1523/JNEUROSCI.3011-10.2010>
- Yang, Y., Raine, A., Lencz, T., Bihle, S., Lacasse, L., & Colletti, P. (2005). Prefrontal white matter in pathological liars. *British Journal of Psychiatry*, 187, 320–325.