Journal of Applied Research in Memory and Cognition 7 (2018) 574–586

Contents lists available at ScienceDirect

Journal of Applied Research in Memory and Cognition

journal homepage: www.elsevier.com/locate/jarmac

The Role of Working Memory Capacity and Cognitive Load in Producing Lies for Autobiographical Information

Ted Maldonado, Frank M. Marchak, Danielle M. Anderson, Keith A. Hutchison*

Montana State University, USA

This study examined effects of working memory capacity (WMC) and cognitive load on one's ability to convincingly tell lies. Recent research suggests imposing a cognitive load improves lie detection. We hypothesized that the cognitive demands of lying, particularly under high load, would lead to poorer performance among low-WMC individuals. In our study, young adults gave truthful or deceptive responses to autobiographical questions when under either high or low cognitive load. Detectors attempted to guess speaker veracity either in person or when later watching a video recording. When under high load, low-WMC individuals' lies were more easily detected and they also had greater difficulty remembering the truth. High-WMC individuals were unaffected by load. These results suggest that imposing cognitive load during interrogations might selectively affect low-WMC individuals, allowing greater lie detection, but also creating wrongful convictions based on changing testimony over time.

Keywords: Working memory capacity, Deception, Cognitive load

General Audience Summary

Telling a lie requires mental effort. For instance, people need to suppress aspects of the truth while quickly generating lies, monitor their behavior to appear genuine, monitor the other person to gauge belief in the lie, and align the lie with shared event knowledge, among other things. Whereas some individuals have the mental capacity to handle the extra effort required to tell a convincing lie, other individuals do not and, as a result, are more often caught in their attempt to lie. The current study examines working memory capacity (WMC) as a hypothesized pre-requisite for telling convincing lies. WMC describes one's ability to hold information in the focus of attention while simultaneously tuning out distractions that might cause forgetting. We hypothesized that low-WMC individuals should have more difficulty managing the extra effort involved in lying when also engaged in another effortful activity. To test this, we had younger adults give truthful and deceptive responses either to a partner or a camera while remembering the location of dots in a matrix. If a partner was present, the partner would decide whether the speaker was telling a truth or lying in real time. If the camera was present, a detector later judged these responses. We found that, when the dot pattern was made more difficult to remember, low-WMC individuals indeed were caught lying more often than high-WMC individuals. This suggests that increased cognitive load, or the amount of mental effort used during a task, impairs low WMC individuals' ability to tell convincing lies. Therefore, adding an additional cognitive load to a low-WMC suspect during interrogation might help catch the suspect in a lie. We suggest researchers conduct additional research to replicate these findings under more realistic interrogation settings.

Author Note

Ted Maldonado, Frank M. Marchak, and Danielle M. Anderson, Department of Psychology, Montana State University, Bozeman, Montana, United States.

We thank Bethany Johanessen, Taylor Merchant, Cara Stettler, and Max Schneider for running participants in this study and Russell Hoeldke, Hilary Maldonado, and Chris Callahan for serving as detectors for Experiment 1b.

* Correspondence concerning this article should be addressed to Keith A. Hutchison, Department of Psychology, Montana State University, P.O. Box 173440, Bozeman, MT 59717-3440, USA. Contact: khutch@montana.edu

Traditionally, lie detection procedures relied on arousal (Vrij, Fisher, Mann, & Leal, 2006) with the assumption that increased arousal stems primarily from peoples' fear of being caught. However, after examining several aspects of arousal-based lie detection, the United States National Research Council (2003) concluded that arousal-based detection identifies changes in physiology, regardless of the cause. Thus, there are recent calls for researchers to look for alternative methods in lie detection, such as specific cognitive determinants of lying (Bond, 2012; National Research Council, 2003; Sporer, 2016; Vrij et al., 2006; Vrij, Granhag, & Mann, 2010).

In line with this call, deception researchers (Vrij, 2008) have capitalized on the assumption that lying is cognitively demanding. Such studies typically use a cognitive load approach (CLA) by imposing additional cognitive demands during interviews, which should interfere with the presumed cognitive resources required for successful lying (see Blandón-Gitlin, Fenn, Masip, & Yoo, 2014; Vrij, 2008, for reviews). Examples of CLA techniques include requiring suspects to tell their story in reverse order (Vrij & Granhag, 2012), strategically revealing evidence toward the end of questioning to force suspects to reconstruct their story to align with new evidence (Hartwig, Granhag, Strömwall, & Vrij, 2005), and placing suspects under time pressure (Walczyk, Mahoney, Doverspike, & Griffith-Ross, 2009). In all cases, the use of the CLA resulted in enhanced lie detection, with a recent plea for better scientific understanding of the nature of such load effects (Blandón-Gitlin et al., 2014; Bond, 2012; Vrij et al., 2006, 2010).

Additional evidence that lying is cognitively demanding comes from neuroimaging studies showing that deception activates areas long known to underlie working memory and executive function (Christ, Van Essen, Watson, Brubaker, & McDermott, 2009; Sip, Roepstorff, McGregor, & Frith, 2008; Vartanian et al., 2013). For instance, Christ et al's (2009) meta-analysis of neuroimaging studies investigating deception showed that, among other areas, the anterior cingulate cortex, inferior frontal gyrus, and the dorsal-lateral prefrontal cortex were typically active during instances of deception. Similarly, Vantanian et al. (2013; see also Vartanian, Kwantes, and Mandel, 2012) gave subjects a Sternberg memory-scanning task (Sternberg, 1966) with a high or low cognitive load, operationalized as a 6-item or 4-item search set, and instructed them to respond truthfully or deceptively regarding whether the memory probe item was in the search set. They found greater right inferior frontal activation, indicative of inhibition, when lying under high load relative to low load. Presumably, the higher load demands required greater effort in suppressing truthful responses.

Taken together with the CLA studies, these data suggest that both individual (e.g., frontal function) and situational (e.g., load) factors influence deception ability, and likely interact in an over-additive fashion, with those lowest in frontal function showing the greatest load-based impairment. One method for investigating this hypothesis is to examine working memory capacity (WMC). As described by Engle (2002), WMC involves both short-term storage and attentional control, such that WMC involves maintaining important information (including task goals) in primary memory while also suppressing task-inappropriate thoughts or behaviors. Kane and Engle (2002) argued that the cognitive performance of individuals low in WMC mimics that of patients with prefrontal cortex lesions. Specifically, both groups perform poorly on verbal fluency tasks and tasks requiring inhibition of prepotent responses. Particularly relevant for our current project, past research shows that low-WMC individuals are impaired, relative to high-WMC individuals, at multitasking (Hambrick, Oswald, Darowski, Rench, & Brou, 2010; Redick et al., 2016), inhibiting dominant, yet inappropriate responses (Kane, Bleckley, Conway, & Engle, 2001; Unsworth, Schrock, & Engle, 2004), conflict monitoring and adjusting behaviors online (Miller, Watson, & Strayer, 2012; Weldon, Mushlin, Kim, & Sohn, 2013), and maintaining context and task goals (Hutchison, 2011; Richmond, Redick, & Braver, 2015). Taken together, low-WMC individuals should struggle in performing all components of lying while under a cognitive load, including quickly generating lies while suppressing the truth, maintaining a "tangled web" of information, regulating behavioral cues, monitoring the story for errors, monitoring the detector for skepticism cues, and recovering from detected inconsistencies to realign their story with the truth.

To use the CLA to examine deception, researchers typically employ face-to-face interviews (Hartwig, Granhag, Strömwall, & Vrij, 2004) or video recorded interviews (Vrij and Mann, 2001a, 2001b) to create more realistic scenarios. Presumably, detection should be greater in face-to-face interviews, because of the additional stress and cognitive load imposed by the physical presence and proximity of an interrogator (Latane, 1981). This additional stress and cognitive load imposed by face-to-face interviews should also increase WMC's contribution during lying (Lane & Vieira, 2012). However, despite these assumptions, previous research findings are inconsistent regarding the effect of different formats on deception ability. Vrij and Mann (2001a, 2001b) conducted two studies in which individuals making veracity judgments via video had a detection rate of 64% (Vrij & Mann, 2001a), but when veracity judgments were made face-to-face, deception detection was exactly at chance levels (Vrij & Mann, 2001b). Thus, in these two experiments, one might conclude that deception detection via video tape could actually lead to higher rates of accurate detection, as compared to making these judgments in real life. However, Hartwig et al. (2004) directly examined this issue by having police detect lies either via a face-to-face scenario or via video recording and found no difference in lie detection accuracy across formats. As noted above, these results are counter-intuitive, as a proximate interrogator should increase a liar's level of stress. Thus, it is also important to understand whether the hypothesized relation between WMC and deception could manifest differently in face-to-face versus videotaped formats. Critically, these differences might help uncover which components of WMC are used and what purpose they serve during lie detection.

The current study represents a first step in investigating the hypothesis that lower WMC individuals will be particularly impaired at producing convincing lies while under cognitive load. As indicated above, we believe there are multiple potential paradigms in which to investigate this relationship. However, we decided to start with simple 1–2 word autobiographical responses to specifically investigate the hypothesis that low-WMC individuals would have a difficult time suppressing truthful responses while quickly generating lies. Because the relation between WMC and deception might depend upon format, we investigated this hypothesis in both face-to-face (Experiment 1a) and videotaped (Experiment 1b) formats.

Method

Participants

We conducted Experiments 1a and 1b concurrently through an online signup in which participants receive partial course credit in an Introduction to Psychology course. Our goal was to achieve 64 participants (32 dyads) in Experiment 1a and approximately the same number of participants in Experiment 1b. We advertised for 2 available slots per session. If both participants showed up, we ran Experiment 1a and if only one participant showed up, we ran Experiment 1b. In Experiment 1a, we could not use a dyad's data if either participants' data were invalid. We therefore needed to run 47 dyads in order to achieve our target of 32 dyads, because we eliminated data due to computer error ($n = 7$ dyads), experimenter error ($n = 1$ dyads), inability to comply with instruction ($n = 4$ dyads), or too many errors on one of the span tasks ($n = 4$). Thus, we analyzed data for 64 participants (32 dyads) in Experiment 1a. None of the dyad participants reported knowing each other prior to participating in the experiment. In Experiment 1b, 59 participants completed the experiment, but we eliminated the data from two participants due to experimenter error ($n = 1$) and inability to comply with instruction ($n = 1$). Thus, we analyzed data for 57 participants in Experiment 1b.

Design and Procedure

Both Experiments 1a and 1b included WM load (high vs. low) and response type (truth vs. lie) measured within groups and WMC measured continuously between groups. When dyads (Experiment 1a) or single participants (Experiment 1b) arrived, they first signed a consent form and then received 64 autobiographical questions to answer truthfully. For dyads, the experimenter immediately engaged participants when they entered the lab so there was no opportunity for participants to talk to each other prior to the experiment. After truthfully answering the initial 64 questions, participants received the Foster et al. (2015) shortened OSPAN task (i.e., Block 1) and the Oswald, McAbee, Redick, and Hambrick (2015) shortened RSPAN task. We used the shortened versions to complete the entire experiment in 1 h and because both versions validly capture WMC (Foster et al., 2015; Oswald et al., 2015). The OSPAN task presented participants with a simple math problem (e.g., $4 \times 5 + 2 = 22$) and instructed them to respond "yes" or "no" via mouse press, depending on whether the answer to the math problem was correct. Following the math problem, participants saw a single letter (e.g., L) to retain in memory. After viewing between 2 and 7 sets of math/letter pairs, participants were instructed to recall the letters in the correct order of presentation. Their OSPAN score was determined by the number of the correctly recalled letters for sets in which all letters were recalled in the correct order. The RSPAN task followed the same procedure as the OSPAN task except, instead of solving math equations, participants were presented with a sentence and instructed to respond "yes" or "no" by mouse press indicating whether or not the sentence was sensible. Also, rather than viewing 2–7 sets of items once each, participants viewed sets of 4, 5, or 6 items twice each.

Upon completion of the span tasks, in Experiment 1a, participants received a packet with instructions for the memory load, the speaker task, and the detector task. The experimenter randomly assigned participants to be either the speaker or detector for the first block of trials (see Figure 1). The experimenter then asked the speaker to memorize a 4×4 matrix containing four dots for four seconds and to remember the location of the dots for later recall. The dots were either in a straight line (low load; see Packet A in Figure 1) or scattered throughout the matrix (high load; see Packet B in Figure 1). The experimenter then read aloud eight questions answered earlier from the 64-item questionnaire in a pseudo-randomized order, such that all participants saw the same randomized order of questions. The speakers answered the eight questions based on a random order of four truthful and four deceptive responses as requested by their packet. The detector attempted to determine if the speaker was lying or telling the truth and recorded their response in their packet. The experimenter wrote down the speaker's responses to ensure the speaker complied with the instruction to tell a truth or a lie. After answering all eight questions, the speaker filled in the blank matrix with the dots held in memory. Once completed, roles were reversed for the next block of eight questions. The goal for the speaker was to be convincing in all responses. The goal for the detector was to correctly discern whether the speaker was lying or telling the truth for each response. This process repeated three more times. At the end of the experiment, participants completed an exit survey containing questions about their perceived ability to produce lies, detect lies, and any exposure to their partner prior to the experiment.

The procedures used in Experiment 1b were like those used in Experiment 1a, with a few key differences. In Experiment 1b, instead of speaking to a partner, the participant spoke to a camera. A Sony digital camera captured participant responses and all video recordings were stored on a password-protected computer. Additionally, because the detector was replaced with a camera, participants did not switch roles, which altered the packet slightly such that a speaker did not have to detect (see Figure 1). Further, three young adults, a 24-year-old female, a 25-year-old male, and a 35-year old male volunteered to be detectors. We chose three detectors because doing so allowed us to examine a matrix of interrater reliabilities across both high and low load conditions and should provide a more reliable measure of detection than relying on a single detector. We relied on these particular three detectors because none of them were involved in data collection, aware of speaker WMC or load conditions, or informed of the purpose of the study. Additionally, they were willing to volunteer several hours of their time to watch hours of video and make veracity judgments for 59 participants. Further, they were not lie detection experts and thus would make natural

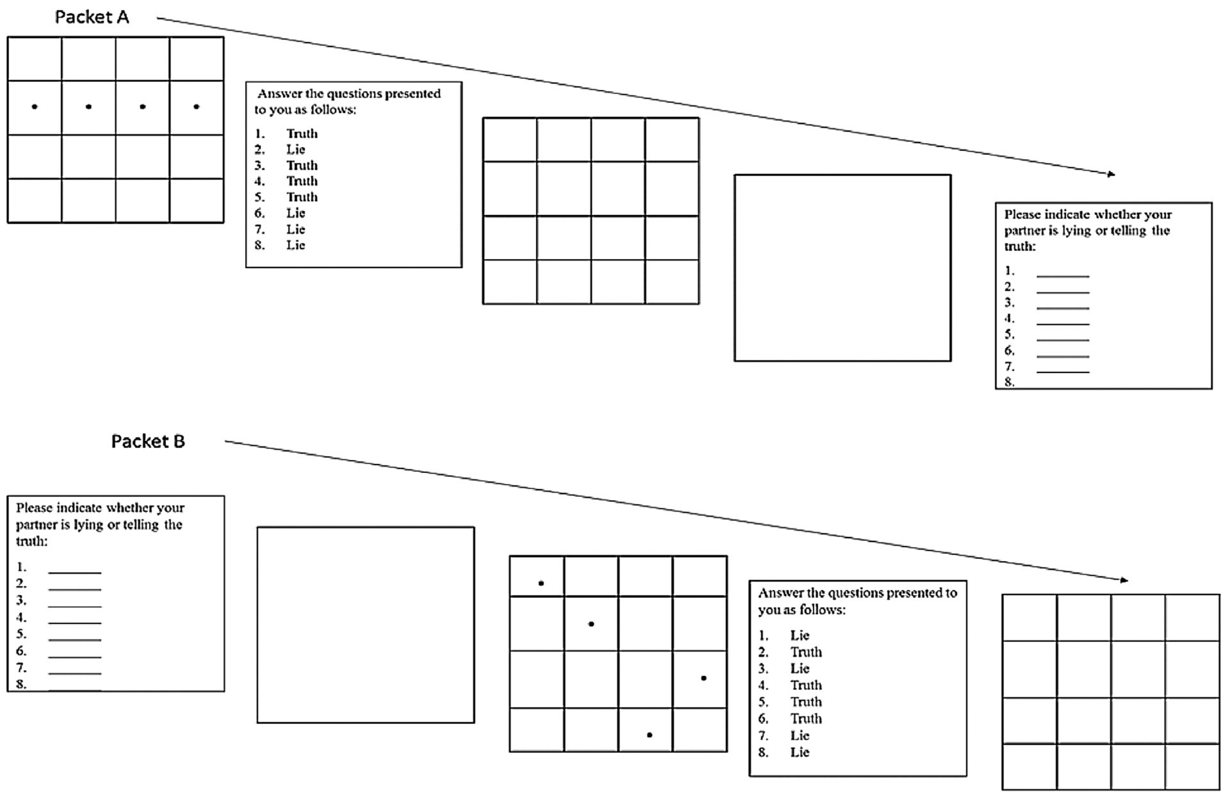

Figure 1. Sample trial under low memory load (top) and high memory load (bottom) conditions in Experiment 1a. In Experiment 1b, participants do not engage in the detection task and only see the first 3 (top) or last 3 (bottom) pages depending upon load.

judgments, not based on a specified framework. The detector's goal was simply to determine if the speaker was lying or telling the truth on each trial.

Materials

The materials described below are for Experiment 1a. The materials used in Experiment 1b were identical, except that participants played the role of the speaker only, thus eliminating any need for a detector and corresponding materials.

Questionnaire. Participants answered 64 questions requiring one-word or two-word truthful answers prior to the experiment. Questions included "What was the name of the first elementary school you attended?" and "In what city or town does your nearest sibling live?" Naturally, some people might not remember the name of the elementary school they attended or might be an only child and would respond "I do not remember" or "I am an only child," respectively. For other questions, participants might respond "I do not know" because they simply do not have an answer. Participants were told these are acceptable responses and would be counted as truthful responses. If they were later prompted to tell the truth, they would simply need to repeat this response. If later prompted to lie for these questions, they would have to come up with a convincing answer. Participants gave "I don't know," "I don't remember," or blank responses to 2.3% of questions for Experiment 1a and 2.4% of the time for Experiment 1b. These responses were most common for question 5 ("what was the last name of the teacher who gave you your first failing grade?"; 12% of participants), and question 45 ("If you could be any fictional character, who would it be?"; 18% of participants). All significant effects reported below remained unchanged when such responses were removed. We chose questions with 1–2 word answers because they allow for intrusion of a correct response during later lies, control for complexity of answer, and would allow us to examine response latency in future studies.

Packet. Participants each received a 20-page packet that first clearly explained the procedure. The next two pages contained the 64 initial questions. If the participant received Packet A, the first page of their packet was a 4×4 matrix that contained four solid dots. A high load matrix contained four dots randomly dispersed within the matrix and a low load matrix contained the four dots in a straight vertical or horizontal line (see Figure 1). On the next page, the participant saw the words *Lie* and *Truth* four times each in a random order. The following page contained a blank 4×4 matrix in which the participant filled in dots from the matrix retrieved from memory during that block of questions. The last page the participant saw during a trial was a sheet with eight blank lines used to write down whether the participant's partner was lying or telling the truth.

Participants who received packet B detected lies first. In this case, the first page contained a sheet of eight blank lines to write down whether the participant's partner was lying or telling the truth. The participant then saw the 4×4 dot matrix and was asked to remember it for later recall. On the next page, the participant saw the words *Lie* and *Truth* four times each in random order indicating how to respond to the experimenter's questions. The last page contained a blank matrix the participant filled in from memory. There were four versions of both packets A and B to counterbalance questions across response (lie vs. truth) and load (high vs. low) conditions.

This sequence occurred four times. Thus, participants completed four blocks in which they answered eight questions for a total of 32 responses across the entire experiment. During each block, participants were placed under a cognitive load, twice under high load and twice under low load, which alternated across each block. Figure 1 depicts the progression of one block.

Exit survey. The last page of the packet contained an exit survey with the questions "Did you know the other participant prior to today?" (Experiment 1a only); "What factors did you use to determine if the participant was lying?"; "Do you believe you are a good liar? Why?"; and "Do you think you are good at catching lies? Why?" These questions were used to ensure participants did not know each other prior to the experiment and to serve as a qualitative report on how well participants believed they could lie and discern lies.

Results

Data Scoring

To ensure speakers complied with the instructions regarding giving truthful versus dishonest responses, we compared participants' responses during the task to their initial 64 truthful answers given at the beginning of the experiment. We marked the response as incorrect if their response (lie or truth) did not match the instructions. In addition to compliance analyses, we scored the proportion of correct dot memory responses as a manipulation check for our high and low WM load conditions. Finally, we examined detectors' ability to discriminate deceptive from truthful responses using a signal detection analysis (Macmillan & Creelman, 2004) with d' (standardized hit scores minus standardized false alarm scores divided by the square root of 2) and C (the negated sum of standardized hit and false alarm scores divided by the square root of 2). We used Macmillan and Creelman's (2004, p. 21) correction by changing 0 or 1 values to 0.05 and .95, respectively. In all analyses reported below, significant effects contain a two-tailed p-value $< .05$ and partial eta-square (η_p^2) as our measure of effect size. Although Experiments 1a and 1b were conducted at the same time, we first present results from each study separately before examining the combined data. No effects involving gender approached significance.

Experiment 1a

The purpose of this study was to examine the role of WMC in deception when one must lie in a face-to-face situation under cognitive load. The top half of Table 1 shows the descriptive statistics for Experiment 1a. Most measures appear normally distributed except for low load dot performance and compliance, due to ceiling effects. These ceiling effects might lead to null correlations involving these measures. Additionally, reliability appears high across all measures except for d'. We calculated d' reliability by computing odd-even correlations, with Spearman–Brown correction, on valid trials within each Participant × Load condition. This reliability for d' was low across the 64 different detectors. We were unable to calculate reliability for dot performance, because 100% of low-load participants and 94% of high load participants received perfect scores on either the first or the second matrix.

Compliance

To examine instruction compliance in Experiment 1a (see top of Table 1), we first conducted a 2 × 2 ANOVA with load (high vs. low) and response type (truth vs. lie) manipulated within subjects. In this ANOVA, there was a main effect of response type, F $(1, 63) = 26.95$, $p < .01$, $\eta_p^2 = .30$, but not load, F $(1, 63) = 1.57$, $p = .22$, $\eta_p^2 = .03$. Specifically, participants made more errors when trying to tell the truth ($M = 18\%$, $SE = 1.7\%$) than when lying ($M = 7.7\%$, $SE = 1.1\%$). The Load × Response Type interaction was not significant, F $(1, 63) = .83$, $p = .37$, $\eta_p^2 = .01$.

The correlations between the complex span tasks, compliance, and lie discriminability (d') under high and low cognitive load conditions are shown at the top of Table 2. We created a WMC composite score (average of z-transformed RSPAN and OPSAN scores) as a covariate in an ANCOVA to examine potential interactions between load, response type, and WMC. There was a main effect of WMC, F $(1, 62) = 5.10$, $p = .03$, $\eta_p^2 = .08$, such that higher WMC individuals showed fewer errors overall in complying with instructions ($r = -.276$, $p = .027$, for the correlation between WMC composite and overall compliance errors). No other effects involving WMC reached significance (all $ps > .17$, $\eta_p^2 < .031$).

Although interactions involving the WMC composite were not significant, exploration of the correlation matrix in Table 2 showed that RSPAN alone might relate to performance, whereas OSPAN did not correlate with performance under any conditions. We therefore decided to repeat the analyses above using only RSPAN for exploratory purposes, with an appreciation and caution that such post-hoc exploration of data will inflate the alpha level, increasing the probability of Type 1 error. When included as a covariate, the main effect of RSPAN was significant, F $(1, 62) = 7.36$, $p < .01$, $\eta_p^2 = .11$, showing overall greater compliance for those higher in reading span. However, this was qualified by a significant Response Type × RSPAN interaction, F $(1, 62) = 4.36$, $p = .04$, $\eta_p^2 = 0.7$. Follow-up analyses revealed that truth errors significantly decreased as RSPAN increased ($r = -.35$, $p = .005$), but lie errors did not ($r = .074$, $p = .56$). No other effects involving RSPAN approached significance (all $ps > .45$, $\eta_p^2 < .01$). When OSPAN was included as a covariate, no effects involving OSPAN approached significance (all $ps > .29$, $\eta_p^2 < .02$).

Lie Detection

To examine lie detection, we conducted an ANCOVA with load (high vs. low) manipulated within subjects and WMC composite score as a covariate. In this ANCOVA, no effects approached significance (all $ps > .20$, $\eta_p^2 < .03$). In an exploratory ANCOVA with just RSPAN, there was a marginal Load × RSPAN interaction, F $(1, 62) = 2.86$, $p = .096$,

Table 1
Descriptive Statistics for Experiment 1a and 1b

	Mean	Standard Deviation	Skewness	Kurtosis	Internal Consistency
Experiment 1					
HL dots	.818	.185	−.717	−.531	–
LL dots	.992	.063	−8.00	64.000	–
HL *d′*	.544	.698	.110	.365	.148[c]
LL *d′*	.503	.700	.595	1.307	.090[c]
Overall *d′*	.523	.512	.727	−.051	.049[c]
RSPAN	13.578	8.986	.149	−.939	.760[a]
OSPAN	17.391	5.362	−.290	−.648	.690[b]
WMC span composite	.000	.833	.265	−.612	–
HL lie compliance	.940	.102	−1.768	2.537	.601[c]
HL true compliance	.824	.147	−.833	.501	.683[c]
LL lie compliance	.906	.128	−1.536	2.030	.693[c]
LL true compliance	.816	.181	−1.362	1.966	.747[c]
Overall compliance	.872	.082	−.968	.926	.811[c]
Experiment 2					
HL dots	.829	.194	−.920	.133	–
LL dots	.991	.066	−7.550	57.000	–
HL *d′*	.327	.491	.142	−.097	.530[c]
LL *d′*	.291	.358	.060	−.512	.108[c]
Overall *d′*	.309	.336	.040	.319	.553[c]
RSPAN	13.175	8.773	.116	−.902	.760[a]
OSPAN	18.263	5.259	−1.096	1.654	.690[b]
WMC span composite	.000	.893	−.337	−.493	–
HL lie compliance	.965	.094	−4.745	28.179	.628[c]
HL true compliance	.912	.120	−2.010	5.998	.623[c]
LL lie compliance	.950	.091	−2.646	9.764	.496[c]
LL true compliance	.877	.150	−1.120	.294	.753[c]
Overall compliance	.926	.072	−1.557	3.167	.778[c]

[a] Oswald et al. (2015).
[b] Foster et al. (2015).
[c] Based on odd–even split-half correlation with Spearman–Brown correction.

$\eta_p^2 = .044$, such that lie detection under high load decreased as RSPAN increased ($r = -.294$, $p = .019$); however, lie detection under low load was unrelated to RPSAN ($r = -.008$, $p = .951$). There were again no effects when OSPAN was entered as a covariate (all $ps > .29$, $\eta_p^2 < .02$).

We performed a similar analysis on participants' criterion for responding (C), indicating how liberal or conservative detectors were in guessing lies. However, no effects were significant (all $ps > .13$, $\eta_p^2 < .04$).

Manipulation Check

As a manipulation check, we compared the proportion of dots correctly recalled in the high-load versus low-load tasks. As expected, memory performance dropped from low ($M = .99$, $SE = .01$) to high load ($M = .82$, $SE = .02$), $t(64) = 7.042$, $p < .001$. Additionally, speakers' dot performance was unrelated to their probability of lies being detected (d') under either high load ($r = .078$, $p = .54$) or low load ($r = .043$, $p = .74$). When we included the WMC composite score in the analysis, there was a marginal main effect of WMC, F (1, 62) = 3.03, $p = .087$, $\eta_p^2 = .047$, with a trend for higher dot performance among higher WMC individuals. This WMC effect did not interact with load, $F < 1$, $\eta_p^2 < .01$, and the effect of WMC was not significant when we considered either RSPAN ($p = .29$, $\eta_p^2 < .02$) or OSPAN ($p = .22$, $\eta_p^2 < .03$) alone. However, as noted above and shown in Table 1, the ceiling effects present in low load dot performance mean one needs to be careful interpreting null correlations between this measure and other variables.

Experiment 1b

The purpose of this study was to examine the role of WMC in deception when one must lie to a video recorder under cognitive load. The bottom half of Table 1 contains the descriptive statistics for Experiment 1b. Again, most measures appear normally distributed except for low load dot performance and compliance, due to ceiling effects. Internal consistency is sufficiently high across all measures except for d' under low load.

Compliance

As in Experiment 1a, to examine compliance, we first conducted a 2 × 2 ANOVA with WM load (high vs. low) and response type (truth vs. lie) manipulated within subjects. This ANOVA showed main effects of both response type, F (1, 56) = 14.00, $p < .01$, $\eta_p^2 = .200$, and load, F (1, 56) = 4.20, $p = .045$, $\eta_p^2 = .07$. Specifically, participants made more errors responding when under low load ($M = .087$, $SE = .012$) than when under high load ($M = .061$, $SE = .01$), and when telling

Table 2
Correlation Matrix for Experiment 1a, Experiment 1b, and the Combined Experiments

Experiment 1a	1	2	3	4	5	6	7	8	9	10
RSPAN	–									
OSPAN	.388*	–								
SPAN composite	.833*	.833*	–							
HL lie compliance	.082	.075	.094	–						
HL true compliance	.273*	.050	.194	.040	–					
LL lie compliance	.034	.068	.061	.109	−.007	–				
LL true compliance	.296*	.110	.243	−.074	.361*	.113	–			
HL *d′*	−.294*	−.055	−.209	.204	−.103	−.046	−.299*	–		
LL *d′*	−.008	.024	.009	.107	−.008	.111	−.111	.074	–	
Overall *d′*	−.206	−.022	−.136	.212	−.076	.044	−.280*	.732*	.733*	–
Experiment 1b	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
RSPAN	–									
OSPAN	.594*	–								
SPAN composite	.893*	.893*	–							
HL lie compliance	−.296*	−.012	−.173	–						
HL true compliance	.087	−.056	.017	−.006	–					
LL lie compliance	.008	−.037	−.016	.541*	.276*	–				
LL true compliance	.022	−.183	−.090	−.034	.258	.151	–			
HL *d′*	−.310*	−.130	−.246	.074	−.004	−.085	.086	–		
LL *d′*	.051	.127	.100	−.028	−.182	−.207	.225	.239	–	
Overall *d′*	−.199	−.027	−.127	.039	−.099	−.172	.183	.857*	.706*	–
Combined	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
RSPAN	–									
OSPAN	.480*	–								
SPAN composite	.861*	.858*	–							
HL lie compliance	−.088	.048	−.030	–						
HL true compliance	.178	.035	.110	.061	–					
LL lie compliance	.019	.044	.029	.283*	.143	–				
LL true compliance	.173	.003	.094	−.033	.357*	.155	–			
HL *d′*	−.287*	−.098	−.216*	.129	−.120	−.090	−.193*	–		
LL *d′*	.014	.035	.036	.039	−.108	−.003	−.051	.142	–	
Overall *d′*	−.188*	−.045	−.125	.113	−.151	−.064	−.165	.776*	.734*	–

Note. HL = high load, LL = low load.
* $p < .05$ (two-tailed).

the truth ($M = .105$, $SE = .014$) than when lying ($M = .043$, $SE = .011$). The Load × Response Type interaction was not significant, F (1, 56) = .61, $p = .44$, $\eta_p^2 = .01$.

We next included a WMC composite score as a covariate in an ANCOVA to examine potential interactions between these variables and WMC. No effects involving WMC reached significance (all ps > .24, $\eta_p^2 < .03$). As in Experiment 1a, we then individually included RSAPN and OSPAN scores as covariates in an ANCOVA. When added, no effects involving RSPAN (all ps > .17, $\eta_p^2 < .06$) or OSPAN approached significance (all ps > .32, $\eta_p^2 < .02$).

Lie Detection

Inter-rater reliability. Because Experiment 1b included the same three observers detecting all participants, we first examined inter-rater reliability. Significant inter-rater reliability in lie detection would indicate stable individual differences in deception ability, such that a bad liar should be consistently detected across multiple observers and a good liar should consistently fool multiple observers. On the other hand, if detectors each relied on different information to make their judgments, or if observers were randomly guessing, then inter-rater reliability should be zero. Table 3 shows the interrater reliabilities between our three observers in lie detection as a function of speaker load. When speakers were under high load, inter-rater reliabilities were moderate, indicating stable individual differences in deception ability. However, under low load, inter-rater reliability was lower and not significant in 2/3 cases. An additional intraclass correlation (Shrout & Fleiss, 1979) analysis demonstrated that, as a group, raters had an intraclass correlation of .463, indicating moderate consistency across detectors.

Detection performance. To examine lie detection, we conducted an ANCOVA with WM load (high vs. low) manipulated within subjects and WMC composite score as a covariate, as in Experiment 1a. This ANCOVA showed a significant Load × WMC interaction [$F(1, 55) = 5.19$, $p = .03$, $\eta_p^2 = .09$]. Lie detection under high load marginally decreased as WMC increased ($r = -246$, $p = .065$), but lie detection under low load was unrelated to WMC ($r = .100$, $p = .459$). No other effects approached significance (all ps > .34, $\eta_p^2 < .02$).

Table 3
Inter-Observer Reliabilities Between Detectors 1, 2, and 3 in Lie Detection of Speakers Under High Load (HL) or Low Load (LL)

	HL_1	HL_2	HL_3
HL_1	–	.56*	.30*
HL_2		–	.37*
HL_3			–
	LL_1	LL_2	LL_3
LL_1	–	.00	.14
LL_2		–	.31*
LL_3			–

* $p < .05$ (two-tailed).

We again examined each WMC measure separately. When adding RSPAN as a covariate, we found a Load × RSPAN interaction, $F(1, 55) = 6.23$, $p = .02$, $\eta_p^2 = .10$, such that lie detection under high load significantly decreased as RSPAN increased ($r = -.310$, $p = .019$), but lie detection under low load was unrelated to RSPAN ($r = .051$, $p = .706$). When adding OSPAN as a covariate, no effects reached significance (all $ps > .11$, $\eta_p^2 < .001$). When we performed a similar analysis on criterion, there was a marginal overall effect of RSPAN, $F(1, 55) = 3.55$, $p = .065$, $\eta_p^2 = .06$, such that observers tended to be more liberal in guessing lies for those lower in RSPAN. No other effects on criterion were significant ($ps > .18$, $\eta_p^2 < .04$).

Manipulation Check

We again compared the proportion of dots correctly recalled in the high-load versus low-load tasks. As expected, memory performance dropped from low ($M = .99$, $SE = .01$) to high load ($M = .83$, $SE = .03$), $t(64) = 6.872$, $p < .001$. Additionally, speakers' dot performance was unrelated to their probability of lies being detected (d') under either high load ($r = .042$, $p = .75$) or low load ($r = -.042$, $p = .76$). No other effects approached significance ($ps > .51$ $\eta_p^2 < .01$), nor was WMC related to dot performance ($ps > .74$, $\eta_p^2 < .001$). Again, this null effect of WMC might have been due to ceiling effects, particularly in the low load condition.

Experiment 1a and 1b Combined

The purposes for combining data were twofold. First, we are able to increase our statistical power to detect any subtle effects not captured when experiments were examined in isolation. Second, a combined data analysis allows us to compare performance differences between face-to-face and videotaped formats.

Compliance

To examine instruction compliance across Experiments 1a and 1b, we first conducted a 2 × 2 × 2 mixed ANOVA with format (face-to-face vs. videotaped) manipulated between subjects and load (high vs. low) and response type (truth vs. lie) manipulated within subjects. In this initial ANOVA, there were main effects of response type, $F(1, 119) = 39.66$, $p < .01$, $\eta_p^2 = .25$, format, $F(1, 119) = 15.03$, $p < .01$, $\eta_p^2 = .11$, and load, $F(1, 119) = 4.80$, $p = .03$, $\eta_p^2 = .04$. Specifically, participants made more errors when responding truthfully ($M = .14$, $SE = .01$) than when lying ($M = .06$, $SE = .01$), when responding face-to-face ($M = .13$, $SE = .01$) than when being videotaped ($M = .07$, $SE = .01$), and when responding under low load ($M = .11$, $SE = .01$) than high load ($M = .09$, $SE = .01$). No other effects approached significance (all $ps > .12$, $\eta_p^2 < .02$).

We next sequentially included RSPAN and OSPAN as covariates in an ANCOVA to examine potential interactions between these variables and WMC. When RSPAN was included as a covariate, the main effect of RSPAN was marginal, $F(1, 118) = 3.29$, $p = .07$, $\eta_p^2 = .03$. However, as shown in Figure 2, this was qualified by a significant Response Type × RSPAN interaction, $F(1, 118) = 5.79$, $p = .02$, $\eta_p^2 = .05$. Follow up analyses revealed that truth errors significantly decreased across increasing RSPAN ($r = -.21$, $p = .02$), but lie errors did not ($r = .04$, $p = .68$). No other effects involving RSPAN approached significance (all $ps > .45$, $\eta_p^2 < .01$). Unlike RSPAN, when OSPAN was included as a covariate, no effects involving OSPAN approached significance (all $ps > .69$, $\eta_p^2 < .01$).

Lie Detection

To examine lie detection, we first conducted a 2 × 2 mixed ANOVA on detector d' with format (face-to-face vs. videotaped) manipulated between subjects and load (high vs. low) manipulated within subjects. In this initial ANOVA, only the main effect of format was significant, $F(1, 119) = 7.20$, $p < .01$, $\eta_p^2 = .06$. Specifically, lie detection (d') was greater in the face-to-face ($M = .52$, $SE = .06$) than videotaped ($M = .31$, $SE = .06$) format, although both were significantly above zero, $t(63) = 8.17$, $p < .01$, and $t(56) = 6.94$, $p < .01$, in face-to-face and videotaped formats, respectively. No other effects approached significance ($ps > .58$, $\eta_p^2 < .01$).

To examine WMC, we next conducted an ANCOVA with load (high vs. low) manipulated within subjects, format (face-to-face vs. videotaped) examined between experiments, and WMC composite score as a covariate. There was a significant Load × WMC interaction, $F(1, 118) = 4.79$, $p = .03$, $\eta_p^2 = .04$. Lie detection under high load decreased as WMC increased ($r = -.216$, $p = .017$), but lie detection under low load was unrelated to WMC ($r = .036$, $p = .693$). No other effects approached significance (all $ps > .34$, $\eta_p^2 < .02$).

As before, we next sequentially included RSPAN and OSPAN as covariates in an ANCOVA to examine potential interactions between these variables and WMC. When RSPAN was included as a covariate, the main effect of RSPAN was significant, $F(1, 118) = 4.89$, $p < .01$, $\eta_p^2 = .04$, such that lie detection decreased as the speakers' RSPAN scores increased. However, this was qualified by a significant Load × RSPAN crossover interaction, $F(1, 118) = 7.06$, $p < .01$, $\eta_p^2 = .06$, shown in Figure 3. Follow up correlations revealed that lie detection decreased across increasing RSPAN when speakers were under high load ($r = -.29$, $p < .01$), but not when they were under low load ($r = .01$, $p = .88$). This Load × RSPAN interaction did not depend upon format ($p > .33$, $\eta_p^2 < .02$). Further, the crossover

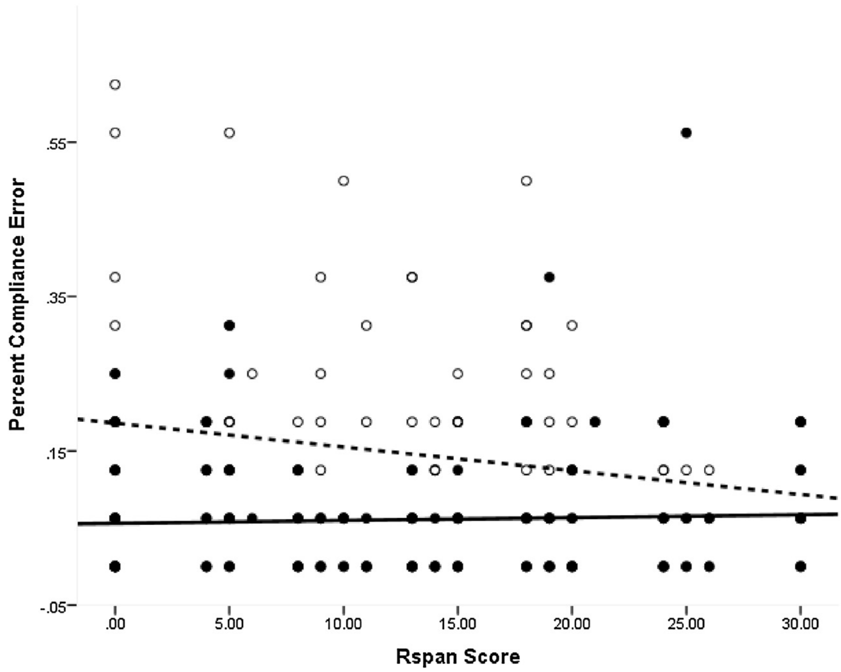

Figure 2. Percent compliance errors collapsed across Experiments 1a and 1b for truthful (dotted line, open circles) and deceptive (solid line, filled circles) responses as a function of WMC (RSPAN).

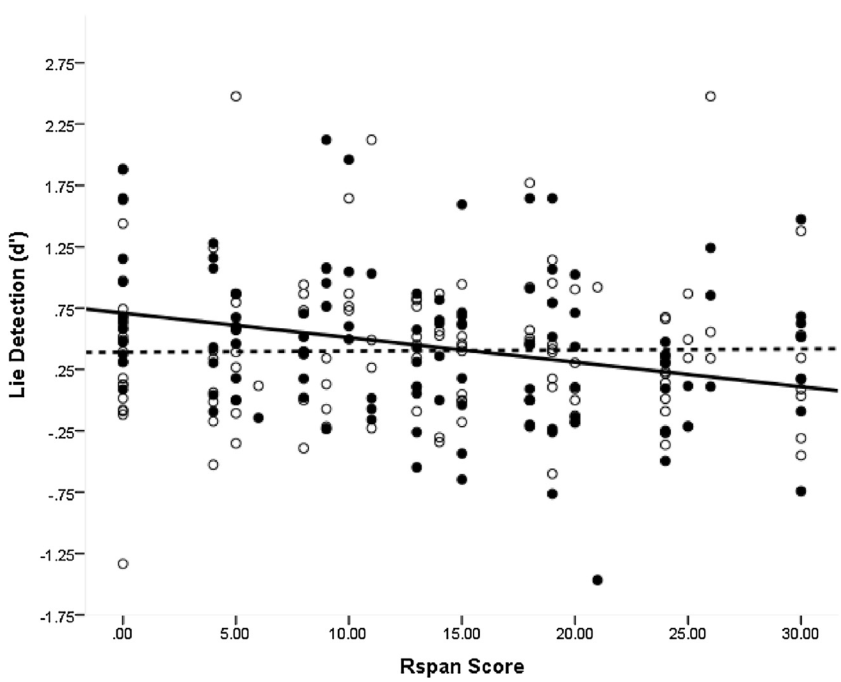

Figure 3. Lie detection when speakers were under high load (solid line, filled circles) or low load (dotted line, open circles) as a function of WMC (RSPAN).

interaction suggests the possibility that not only were low spans more easily detected under high load, but high spans were actually *less* easily detected under high load than low load. To test this, we examined only those who were at least one standard deviation above (high span) or below (low span) the mean RSPAN score. Whereas low spans indeed were more easily detected under high load ($M = .78$, $SE = .14$) than low load ($M = .34$, $SE = .14$), $t(44) = 2.30$, $p = .03$, there was no difference in high spans' detectability between high ($M = .29$, $SE = .13$) and low load ($M = .39$, $SE = .15$), $t(44) = -0.50$, $p = .62$. Finally, when OSPAN was instead included as a covariate, no effects involving OSPAN approached significance (all ps > .22, $\eta_p^2 <$ .02).

As noted above, the pattern for RSPAN mimicked that observed for the overall WMC composite. Indeed, we believe the WMC × Load interaction was completely driven by RSPAN.

Specifically, regressing the load effect (high load d' − low load d') on RSPAN alone in step 1 produced a significant R^2 ($R^2 = .056$, $p = .009$), but adding OSPAN in step 2 resulted in no change ($\Delta R^2 = .000$, $p = .925$). In contrast, regressing the load effect on OSPAN in step 1 resulted in a null effect ($R^2 = .012$, $p = .227$), but adding RSPAN in step 2 led to a significant increase in predictability ($\Delta R^2 = .044$, $p = .020$). Therefore, RSPAN alone produced the significant interaction observed in the overall composite.

We next examined potential criterion differences as a function of WM load, format, and WMC. There was a trend towards more conservative responding among the 3 observers viewing videotapes in Experiment 1b ($M = .47$, $SE = .04$) than among the 64 participants who took turns as both speakers and detectors in Experiment 1a ($M = .34$, $SE = .06$): $F(1, 118) = 3.04$, $p = .08$, $\eta_p^2 = .03$, for the difference in overall criterion across Experiments 1a and 1b. No other effects approached significance (all ps > .42, $\eta_p^2 < .01$).

Exit Surveys

Examining our exploratory exit survey, we did not find a relationship between self-reported lying ability and performance under either high load ($r = .02$, $p = .91$ and $r = -.08$, $p = .54$, for Experiments 1a and 1b, respectively) or low load ($r = .07$, $p = .57$ and $r = .09$, $p = .53$). However, we did find a relationship between self-reported detection ability and actual detection under high load ($r = .28$, $p = .03$). The highest detection scores were from those who credited their ability to practice or focusing on verbal cues such as reaction time and hesitation, whereas those who focused on body language had the lowest detection. While exploratory, this is still in line with Hartwig et al.'s (2004) suggestion that verbal cues might be more useful, particularly in a face-to-face scenario. This finding is also consistent with previous findings that demonstrate no consistent nonverbal deception cue (i.e., something akin to Pinocchio's nose), despite a widespread belief in such cues (see DePaulo et al., 2003 for a review).

Discussion

In the present study, we examined WMC's role in deception when under cognitive load using two format types. Previous research is not clear on what characteristics comprise a good liar, although CLA research does provide evidence to suggest WMC might play a role in deception. As predicted, when under high load, individuals low in WMC had more lies detected than high-WMC individuals. In fact, although we predicted both a main effect of cognitive load and an interaction with WMC, there surprisingly was no main effect of load. Specifically, imposing a cognitive load selectively impaired low-WMC individuals' ability to lie convincingly whereas high-WMC individuals were equally, if not more, convincing under the higher cognitive load. This null load effect suggests a problem with the CLA in that it might not be effective overall, just among those low in WMC.

Surprisingly, participants made overall greater errors when under low load than under high load. Intuitively, increased load should make it harder to retrieve the initial responses to the 64-item questionnaire from memory. However, past research shows little to no effect of divided attention during retrieval (see Craik, 2001, for a review) so perhaps the higher load did not impair participants' ability to retrieve prior responses from memory. Relatedly, an increase in load should make it harder to purposefully generate lies under deception instructions because the increased load needs to contend with the other processes necessary to tell a convincing lie. Although the overall increase in errors under low load runs counter to any a priori hypothesis, perhaps the higher load might have caused participants to simply produce or reject the first response to come to mind under truth and lie conditions, respectively, rather than spending time trying to recollect which option they chose when filling out the initial questionnaire. If so, one would probably see overall faster response latencies under high load than low load. We recommend that future research not only test whether this effect replicates, but also incorporate reaction time measures to test whether it corresponds to faster responding in the high load condition.

Another surprising finding was that participants made more errors under truth instructions than lie instructions. This finding exposes an important memory component to deception in that forgetting one's initial response would likely be mistaken as evidence of deceitfulness during an interrogation. Indeed, detecting inconsistency is often used as evidence for deception (for a review of testimonial consistency, memory accuracy, and deception, see Fisher, Vrij, & Leins, 2013). That this interacted with RSPAN further supports that this non-compliance is due to forgetting of original responses rather than to participants changing their mind regarding their answer. Shipstead, Lindsey, Marshall, and Engle (2014) suggested that primary memory, secondary memory, and attentional control compose a multi-mechanism model that explains individual differences in WMC. Under Shipstead et al.'s framework, participants in the current paradigm would need primary memory and attentional control to simultaneously handle the load matrix and focus on appropriately answering the questions. However, participants would also need secondary memory to retrieve their prior responses. These truth errors presumably reflect such a deficiency in retrieving information from secondary memory. The interaction with WMC is particularly concerning, because it could lead to low-WMC individuals experiencing more false-positive arrests when questioned repeatedly, due to their poor memory for actual events and resulting inconsistencies in their story.

Format also affected lie detection, such that subjects made more errors and told less convincing lies in the face-to-face format versus the videotaped format. This is consistent with our hypothesis that the face-to-face format is more stressful than the videotaped format. This additional stress could potentially cause both increased lie detection and increased forgetting of truthful responses. However, although increased stress should predict that low-WMC individuals would be especially vulnerable to effects of format, we did not obtain a WMC × Format interaction.

Consistent with our effect of format, police officers strongly believe it is easier to detect lies in-person than when watching videotapes (Strömwall & Granhag, 2003); however, in practice, police officers who interrogated individuals face-to-face were no more accurate in detecting deception, and often perceive individuals as more truthful (truth bias), than police officers who were observing video (Hartwig et al., 2004). Hartwig et al. suggested that officers might rely more on verbal cues when detecting face-to-face, even though police officers believe nonverbal cues are more reliable than verbal cues (Strömwall & Granhag, 2003). In regard to truth bias, our criterion data showed the opposite trend in truth bias described by Hartwig et al., such that when participants detected face-to-face they were less likely to label responses truthful than the three detectors who detected via video. Perhaps the lack of emotion, arousal, or personal consequences associated with the current autobiographical questions could account for the difference from past studies. Relatedly, potential differences between our three video detectors in Experiment 1b and the 64 undergraduate face-to-face detectors in Experiment 1a might also drive the difference in criterion observed across formats. However, we should also note that this difference did not reach statistical significance and both groups did show an overall truth bias, as reflected by criterion scores above zero. Future research should keep the detectors constant between video and face-to face formats and examine differences in criterion as a result of format.

One null effect is especially noteworthy. When we examined RSPAN and OSPAN individually, we only saw effects of RSPAN, such that the working memory task involving simultaneous sentence processing was related to deceptive behavior but the working memory task involving simultaneous math computations was not. While this result might be surprising, it is somewhat consistent with a past study by Alloway, McCallum, Alloway, and Hoicka (2015). Specifically, Alloway et al. gave 6–7-year-old children verbal and spatial complex span tasks and a trivia game. While playing the trivia game, the children had an opportunity to peek at the correct answer on the back of a card. Researchers asked follow-up questions to try to detect "peekers." Good liars, defined as not being fooled by the follow up questions, had better verbal WMC scores, but equal spatial WMC scores. Similarly, in our study, the WMC task involving concurrent sentence processing was more strongly related to the ability to deceive, which corresponds with overcoming verbal distractors (truthful responses) when lying. This suggests a specific role in processing and manipulating the linguistic components of successful lying. However, we must be careful not to exaggerate this finding. For one, the differential predictability between WMC tasks was not predicted a priori. Moreover, it is possible the differential predictability in the current study was artifactually caused by another factor, such as greater variability in distribution of RSPAN scores than OSPAN scores. Thus, this differential pattern needs to be further replicated before drawing any firm conclusions. Nonetheless, we believe researchers should further examine the contributions of verbal-specific versus domain-general WMC mechanisms in influencing one's ability to lie.

The application of this knowledge could span over a broad literature, but seems particularly useful to the CLA to detect deception. Historically, deception detection has been at chance levels, even for individuals expected to make veracity judgments in their vocation, such as judges, law enforcement personnel, and psychiatrists (Ekman & O'Sullivan, 1991; Ekman, O'Sullivan, & Frank, 1999). Thus, current CLA tactics might benefit from the current findings, such as simply knowing that imposing a load selectively impacts low WMC individuals and not only affects their ability to lie, but sometimes causes them to forget the truth. This information can help improve interrogations and the chances of catching liars while protecting the innocent. Further, these results could be parlayed with recent research conducted by Van't Veer, Stel, and van Beest (2014), who had participants report the results of three die roles under high (remember a string of eight letters) and low load (remember a string of two letters). Critically, participants had the opportunity to earn more money if they lied about one of the dice rolls. The authors found that participants lied more under low load than high load, which suggests having a limited cognitive capacity makes one susceptible to being honest whereas more cognitive capacity provides one with a greater opportunity to tell a lie. Additionally, this pattern was not seen on rolls in which participants were not paid, suggesting lying might only be done to fulfill self-served interests. Thus, current CLA tactics might also consider that simply being placed under load might reduce suspects' propensity to lie. However, to the extent that their reduction in propensity to lie under high load was due to a reduction in cognitive capacity, we believe van't Veer et al.'s effect of load would be greater for low-WMC individuals. Future research might examine the differential impact of load and WMC on not only people's ability to lie, but also their propensity to do so when given a choice. Taken together, understanding how cognitive capacity relates to the tendency, motivation, and ability to lie under different cognitive load conditions could be very useful to interrogators, as could the caution regarding low-WMC individuals' inability to remember the truth under high load conditions.

Limits to Generalizability

There are several factors that limit the generalizability of the current study. First, we sampled our participants from a population of 18–22-year-old college students taking introductory psychology. It is unclear whether these findings would apply to other populations. In addition, we chose simple 1–2 word autobiographical answers to provide more stringent control over the complexity of required responses and truthful intrusions. However, effects obtained using such responses might not generalize to more realistic scenarios involving retelling emotional sequences of events while simultaneously being pressed for details, or even directly accused, by interrogators. We suggest researchers test whether indeed this pattern will replicate under more ecologically valid situations involving crime scene events and accusations of guilt. However, given that WMC is a multifaceted construct consisting of correlated attentional control, primary memory, and secondary memory components (Shipstead et al., 2014; Unsworth, Fukuda, Awh, & Vogel, 2014), we believe such realistic scenarios will actually increase the contribution of WMC. Specifically, longer stories and events require greater primary and secondary memory demands, emotion regulation requires greater attentional control, and maintaining one's story while dealing with interruptions strains all three components. We therefore predict these modifications should increase, rather than decrease, WMC effect sizes beyond those obtained in our current study using simple 1–2 word autobiographical responses of little emotional consequence.

Conclusion

Adding a cognitive load to increase lie detection appears to only affect low-WMC individuals. Moreover, such approaches might also increase false positives among these individuals, due to altered story details across repeated retellings. High-WMC individuals, on the other hand, appeared immune to such CLA effects in the current study. We should note, however, that some research (e.g., Beilock & Carr, 2005; Conway, Tuholski, Shisler, & Engle, 1999; Rosen & Engle, 1997) suggests that cognitive load can sometimes primarily affect high spans' performance, essentially making their performance mimic low spans' performance. It is not clear why cognitive load sometimes affects low spans more than high spans or vice versa, but one possibility is that the affected group depends upon the nature and difficulty of the primary task. In our study, telling a truth or lie is relatively simple and easily completed by all participants. It was not until we added a cognitive load that we saw performance differences. In contrast, in the studies above, they found performance differences in the low load condition that disappeared under load. Because both WMC and deception are multifaceted, there are many potential ways in which WMC relates to deception and we recommend that future researchers include other paradigms involving maintaining an elaborative story, participating in crime scenes, or dealing with emotional accusations.

Conflict of Interest Statement

The authors declare no conflict of interest.

Author Contributions

Ted Maldonado conducted these experiments as part of his master's thesis at Montana State University. His contributions included writing drafts of all sections and initial analyses. Frank Marchak was a committee member and also contributed insight to the design of the project and edited initial drafts. Danielle Anderson was an undergraduate in the Attention and Memory lab who also wrote sections of the paper, presented a poster of these data at an MSU conference, and edited initial drafts. Keith Hutchison is the principle investigator in the Attention and Memory lab, was Ted's thesis advisor, conducted several analyses, edited all sections, and handled most of the revisions.

References

Alloway, T. P., McCallum, F., Alloway, R. G., & Hoicka, E. (2015). Liar, liar, working memory on fire: Investigating the role of working memory in childhood verbal deception. *Journal of Experimental Child Psychology*, *137*, 30–38.

Beilock, S. L., & Carr, T. H. (2005). When high-powered people fail: Working memory and "choking under pressure" in math. *Psychological Science*, *16*(2), 101–105.

Blandón-Gitlin, I., Fenn, E., Masip, J., & Yoo, A. H. (2014). Cognitive-load approaches to detect deception: Searching for cognitive mechanisms. *Trends in Cognitive Sciences*, *18*(9), 441–444.

Bond, G. D. (2012). Focus on basic cognitive mechanisms and strategies in deception research (and remand custody of 'wizards' to Harry Potter movies). *Journal of Applied Research in Memory and Cognition*, *1*, 128–130.

Christ, S. E., Van Essen, D. C., Watson, J. M., Brubaker, L. E., & McDermott, K. B. (2009). The contributions of prefrontal cortex and executive control to deception: Evidence from activation likelihood estimate meta-analyses. *Cerebral Cortex*, *19*(7), 1557–1566.

Conway, A. R., Tuholski, S. W., Shisler, R. J., & Engle, R. W. (1999). The effect of memory load on negative priming: An individual differences investigation. *Memory & Cognition*, *27*(6), 1042–1050.

Craik, F. I. (2001). *Effects of dividing attention on encoding and retrieval processes*.

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin*, *129*(1), 74.

Ekman, P., & O'Sullivan, M. (1991). Who can catch a liar? *American Psychologist*, *46*, 913–920.

Ekman, P., O'Sullivan, M., & Frank, M. G. (1999). A few can catch a liar. *Psychological Science*, *10*(3), 263–266.

Engle, R. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, 19–23.

Fisher, R. P., Vrij, A., & Leins, D. A. (2013). Does testimonial inconsistency indicate memory inaccuracy and deception? Beliefs, empirical research, and theory. In B. Cooper, D. Griesel, & M. Ternes (Eds.), *Applied issues in investigative interviewing, eyewitness memory, and credibility assessment*. New York, NY: Springer.

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*(2), 226–236.

Hambrick, D. Z., Oswald, F. L., Darowski, E. S., Rench, T. A., & Brou, R. (2010). Predictors of multitasking performance in a synthetic work paradigm. *Applied Cognitive Psychology*, *24*(8), 1149–1167.

Hartwig, M., Granhag, P. A., Strömwall, L. A., & Vrij, A. (2004). Police officers' lie detection accuracy: Interrogating freely versus observing video. *Police Quarterly*, *7*(4), 429–456.

Hartwig, M., Granhag, P. A., Strömwall, L. A., & Vrij, A. (2005). Detecting deception via strategic disclosure of evidence. *Law and Human Behavior*, *29*(4), 469.

Hutchison, K. A. (2011). The interactive effects of list-based control, item-based control, and working memory capacity on Stroop performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 851–860.

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working memory capacity, executive attention, and general fluid intelligence: An individual differences perspective. *Psychonomic Bulletin & Review*, *9*, 637–671.

Kane, M. J., Bleckley, M. K., Conway, A. R. A., & Engle, R. W. (2001). A controlled-attention view of WM capacity. *Journal of Experimental Psychology: General*, *130*, 169–183.

Lane, S. M., & Vieira, K. M. (2012). Steering a new course for deception detection research. *Journal of Applied Research in Memory and Cognition*, *1*(2), 136–138.

Latane, B. (1981). The psychology of social impact. *American Psychologist*, *36*(4), 343.

Macmillan, N. A., & Creelman, C. D. (2004). *Detection theory: A user's guide*. Psychology Press.

Miller, A. E., Watson, J. M., & Strayer, D. L. (2012). Individual differences in working memory capacity predict action monitoring and the error-related negativity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(3), 757.

National Research Council. (2003). *The polygraph and lie detection. Committee to review the scientific evidence on the polygraph*. Washington, DC: The National Academic Press.

Oswald, F. L., McAbee, S. T., Redick, T. S., & Hambrick, D. Z. (2015). The development of a short domain-general measure of working memory capacity. *Behavior Research Methods*, *47*(4), 1343–1355.

Redick, T. S., Shipstead, Z., Meier, M. E., Montroy, J. J., Hicks, K. L., Unsworth, N., . . . & Engle, R. W. (2016). Cognitive predictors of a common multitasking ability: Contributions from working memory, attention control, and fluid intelligence. *Journal of Experimental Psychology: General*, *145*(11), 1473.

Richmond, L. L., Redick, T. S., & Braver, T. S. (2015). Remembering to prepare: The benefits (and costs) of high working memory capacity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*, 1764–1777.

Rosen, V. M., & Engle, R. W. (1997). The role of working memory capacity in retrieval. *Journal of Experimental Psychology: General*, *126*(3), 211.

Shipstead, Z., Lindsey, D. R., Marshall, R. L., & Engle, R. W. (2014). The mechanisms of working memory capacity: Primary memory, secondary memory, and attention control. *Journal of Memory and Language*, *72*, 116–141.

Shrout, P., & Fleiss, J. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*(2), 420–428.

Sip, K. E., Roepstorff, A., McGregor, W., & Frith, C. D. (2008). Detecting deception: The scope and limits. *Trends in Cognitive Sciences*, *12*(2), 48–53.

Sporer, S. L. (2016). Deception and cognitive load: Expanding our horizon with a working memory model. *Frontiers in Psychology*, *7*, 420.

Sternberg, S. (1966). High-speed scanning in human memory. *Science*, *153*(3736), 652–654.

Strömwall, L. A., & Granhag, P. A. (2003). How to detect deception? Arresting the beliefs of police officers, prosecutors, and judges. *Psychology, Crime, & Law*, *9*, 19–36.

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.

Unsworth, N., Schrock, J. C., & Engle, R. W. (2004). Working memory capacity and the antisaccade task: Individual differences in voluntary saccade control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(6), 1302.

Van't Veer, A. E., Stel, M., & van Beest, I. (2014). Limited capacity to lie: Cognitive load interferes with being dishonest. *Judgment and Decision Making*, *9*, 199–206.

Vartanian, O., Kwantes, P., & Mandel, D. R. (2012). Lying in the scanner: Localized inhibition predicts lying skill. *Neuroscience Letters*, *529*(1), 18–22.

Vartanian, O., Kwantes, P. J., Mandel, D. R., Bouak, F., Nakashima, A., Smith, I., & Lam, Q. (2013). Right inferior frontal gyrus activation as a neural marker of successful lying. *Frontiers in Human Neuroscience*, *7*.

Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities* (2nd ed.). Chichester: John Wiley.

Vrij, A., Fisher, R., Mann, S., & Leal, S. (2006). Detecting deception by manipulating cognitive load. *Trends in Cognitive Sciences*, *10*, 141–142.

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*(2), 110–117.

Vrij, A., Granhag, P. A., & Mann, S. (2010). Good liars. *The Journal of Psychiatry & Law*, *38*(1–2), 77–98.

Vrij, A., & Mann, S. (2001a). Telling and detecting lies in a high-stake situation: The case of a convicted murderer. *Applied Cognitive Psychology*, *15*, 187–203.

Vrij, A., & Mann, S. (2001b). Who killed my relative? Police officers' ability to detect real life high-stake lies. *Psychology, Crime, & Law*, *7*, 119–132.

Walczyk, J. J., Mahoney, K. T., Doverspike, D., & Griffith-Ross, D. A. (2009). Cognitive lie detection: Response time and consistency of answers as cues to deception. *Journal of Business and Psychology*, *24*(1), 33–49.

Weldon, R. B., Mushlin, H., Kim, B., & Sohn, M. H. (2013). The effect of working memory capacity on conflict monitoring. *Acta Psychologica*, *142*(1), 6–14.

Received 5 June 2017;
received in revised form 16 May 2018;
accepted 23 May 2018
Available online 30 June 2018