Concealed Information Testing with Rapid Serial Visual Presentation, Pupillometry, and Face Stimuli presentation.

Linda Masarà

Department of Behavioral and Social Sciences, University of Groningen

PSB3-BT.2021-2022.3: Bachelor's Thesis

Robbert van der Mijn and Pieter de Vries

E-mail: l.masara@student.rug.nl

July 4, 2022

Author Note

This paper was written as part of the course "Bachelor Thesis" of the Bachelor Program of Psychology, in collaboration with Konstantinos Mazarakis, Rienk Wagenaar Alie de Jong and Nouska van Waveren Hogervorstunder.

Abstract

Concealing information is a common practice that humans apply to many different contexts. However, because in situations such as police investigations discovering the truth becomes a concern of societal importance, methods for the detection of concealed information are being studied. One of these methods, rapid serial visual presentation (RSVP), has already been validated by previous research to be resistant to countermeasures by presenting stimuli on the fringe of awareness, In the current study, we investigate the use of face stimuli in RSVP for concealed information testing with pupillometry. 53 participants were asked to indicate if they saw a specific target face after each RSVP trial. During some of the RSVP streams, a probe (the face of Obama). The aim of our study was to find out whether there would be an increase in pupil size when Obama's face was shown, indicating an automatic response elicited by subliminal salience. We found significant differences in pupil sizes between the familiar target stimuli and irrelevant stimuli but no significant pupil size increase when Obama's face was shown. While pupillometry seems to be a valid instrument to measure familiarity with a stimulus, further research is needed to validate it as a useful instrument for the detection of subliminally salient stimuli.

Keywords: Concealed information testing; RSVP; pupillometry; face stimuli

Concealed Information Testing with Rapid Serial Visual Presentation, Pupillometry, and Face Stimuli presentation.

Regardless of the type of instrument being used, body responses that occur with deception are different from the ones occurring with true answers given to similar stimuli (Block, et al., 1952). Lying has been observed to be a natural and consistent component of human behavior and the frequent use of deceptive messages has made the ability to differentiate between liars and truth-tellers a matter of great importance (Vrij et al., 2010; Ben-Shakhar & Elaad, 2003). However, as shown by previous studies, not even professionals such as psychiatrists and police officers could detect deceit better than college students. (Ekman & O'Sullivan, 1991). Because in certain criminal justice situations successfully detecting lies becomes a deal of great societal relevance and the ability of humans to do so is generally poor, novel techniques have been developed.

Forensic methods for the detection of concealed information rely on the idea that there are certain behavioral and physiological responses associated with lying that are different from those associated with telling the truth. Various studies have investigated behavioral changes displayed by lying individuals based on the assumption that, due to nervousness, they will show unconscious behaviors such as changes in body posture or increased/decreased movement. (Vrij & Mann, 2001). However, since behavior varies considerably across individuals, lie detection cannot solely be based on behavioral responses (Grubin & Madsen, 2005; Bles & Haynes, 2008; Brinke et al., 2014).

In the past few years, physiological approaches measuring changes in arousal and brain activity have been applied to detect concealed knowledge by trying to determine whether people are lying in response to questions they are asked (Bowman et al., 2014). The polygraph test is an example of a physiological approach that assesses the arousal level by looking at the responses of the autonomic nervous system, such as changes in blood pressure and breathing rate occurring due to emotions of tension or anxiety (Keeler, 1933).

Most of the time the polygraph is used in combination with the control question test (CQT) in which a set of relevant and control questions are proposed to the examinee. This test relies on the assumption that greater arousal in response to certain questions indicates lying (Rosenfeld et al., 2004; Matsuda et al., 2019). Despite the promising findings of these CQT techniques, results are often confounded by countermeasures applied by examinees that alter the measurement (Rosenfeld et al., 2004). More recently, neuroimaging techniques such as electroencephalogram (EEG) measures have been applied in CQT to measure stimulus-evoked brain waves, making use of P300 event-related potentials (ERP) as markers for the recognition of concealed information (Rosenfeld et al., 2004).

A more controlled test for discriminating between individuals involved in criminal activity and innocent suspects is the concealed information test (CIT). The idea behind CIT is that involuntary bodily reactions are elicited in response to crime-relevant stimuli that are hidden in the mind of the examinee (Labkovsky & Rosenfeld, 2011). Via CIT it is possible to measure bodily responses of individuals to presented crime-relevant stimuli, by comparing the responses with the ones elicited by neutral alternatives (Ben-Shakhar & Elaad, 2003).

A new CIT method that is being used involves *rapid serial visual presentation* (RSVP), in which around 10 stimuli per second are presented to the examinee, each one taking the place of the previous one at the same location. The rapid presentation puts the stimuli perception at the fringe of awareness, leaving the participants with too little time to exert behavioral control over their responses and reducing the effectiveness of countermeasures (Bowman et al., 2013; Chen et al., 2021). Bowman and colleagues applied the RSVP method to CIT with EEG measurements in a fake name search task. Participants were instructed to look for a previously chosen name as if it was their real name (target stimulus) and not to show any reaction when their real name was presented (the probe stimulus). The researchers found that, regardless of the effort of hiding their responses, the probe generated significant P300 ERPs. These findings suggest that RSVP measures are resistant to countermeasures and for this reason, they can be considered an effective method for CIT (Bowman et al., 2013). Despite these promising results, EEG measurements are not always compatible with an application outside professional laboratories. Hence, it is important to evaluate the effectiveness of the RSVP method with other measurement types that are suitable for widespread application (Chen et al., 2021).

One possibility is offered by *pupillometry*, the measure of pupil response. Previous studies have found mechanisms linking pupil dilation to P300 potentials which suggest that pupil responses and RSVP might be a promising measure in concealed information testing. In short, the processing of task-relevant events and significant stimuli evokes both pupil dilation and P3 ERPs, which are both related to the activation of phasic responses in the locus coeruleus-norepinephrine (LC-NE) system (Nieuwenhuis et al., 2011). For this reason, Chen et al. (2021) replicated the study of Bowman et al. (2013) by using pupillometry instead of EEG measurements and found that, on the group level, pupil dilation is, indeed, effective in RSVP-based CIT. These findings provide evidence for pupil response to be a promising measure to detect concealed information with the RSVP method (Chen et al., 2021).

The present study will focus on pupil response in an RSVP-based CIT task. The aim is to replicate the methodology and stimuli-modality-presentation of Chen et al. (2021), using faces as stimuli instead of names. Humans were found to be particularly able to process faces, probably because of adaptive mechanisms dedicated to the processing of such complex figures (Damasio et al., 1990). Furthermore, previous research supported that the role played by attention is not crucial in face processing, especially with highly familiar faces (e.g., famous faces. Jackson & Reimond, 2006). Asulfyani et al. (2017) conducted a study of similar nature to ours. They used face stimuli in RSVP and by using EEG measurements they found that participants processed famous faces differently compared to unfamiliar faces, adding evidence to the conclusion that RSVP could be used to infer recognition of a familiar face.

In our study, participants will be instructed to look for a face that will be made familiar before the experiment (*the target*), while a series of other faces will be presented in the RSVP stream, including a famous face (*the probe*) and other irrelevant distractor faces. After each RSVP presentation, participants will be asked whether they have seen the target or not, and during the entire task, their pupil sizes will be recorded to investigate whether pupil responses to the probe and irrelevant stimuli differ significantly.

Our work is a subliminal salience study as the aim is to determine whether pupillometry measurements are useful tools to determine whether the participants' perceptual system is able to pick up on certain salient stimuli (e.g. famous faces) without directly searching for them in the RSVP. In other words, it aims at investigating whether pupil responses in RSVP are valid measurements for detecting a suspect's familiarity with certain stimuli (e.g. incriminating information).

In line with EEG research which already demonstrated that salient faces are processed differently from unfamiliar faces (Alsufyani et al., 2017) and with the familiarity pupil effect shown by Chen and colleagues (2021), we expect pupil size to increase when the target face is presented. Furthermore, we also hypothesize that with pupillometry we will be able to detect subliminal salience. When the probe famous face is presented, even if participants are not actively looking for that face nor are aware of its presence, we expect to see a change in the size of the pupil.

Methods

Participants

53 English-speaking individuals took part in the experiment. All of them were first-year Psychology students (37 female and 15 male) at the University of Groningen in the age group of 18-24 (*M* = 19.62, *SD* = 1.25). Participants had normal or corrected to normal vision. Prior to the experiment, participants were instructed to avoid wearing dark eye make-up.

Ethics

This study was conducted following the guidelines of the World Medical Association Declaration of Helsinki (2013) and approved by the ethics committee of the Psychology Department of the University of Groningen (approval number: PSY-2122-S-0168). Informed consent was obtained digitally from all participants before participation, and they were allowed to take a picture of this screen. Oral debriefing was provided to all participants after participation.

Apparatus

Participants were instructed to place their heads on a chin rest with an adjustable height. The distance between them and a 27'' LCD Liyama PL2773H monitor was 60cm. The display resolution was 1280x720 pixels and had a refresh rate of 1000 Hz. RSVP was presented with OpenSesame (Mathôt, S., et al. 2012) running on Windows. Participants used a QWERTY keyboard to indicate their responses. The size of participants' pupils was recorded in arbitrary units by an EyeLink 1000 (SR Research, Canada) during each trial using PyGaze (Dalmaijer, E., et al. 2014). Analyses were performed in JASP (2022) and RStudio (2022).

Stimuli

We selected faces for the experiment from the *10K faces database* (Bainbridge et al., 2013). More specifically, the total number of faces included in the RSVP was 1127. Prior to the experiment one of these faces was randomly selected for each participant to be the target stimulus. As the probe

Figure 1a-b

Examples of stimuli

Note. Figure 1a shows the face of Barack Obama (Souza, 2012). Figure 1b shows a face used as a stimulus (Bainbridge, 2013).

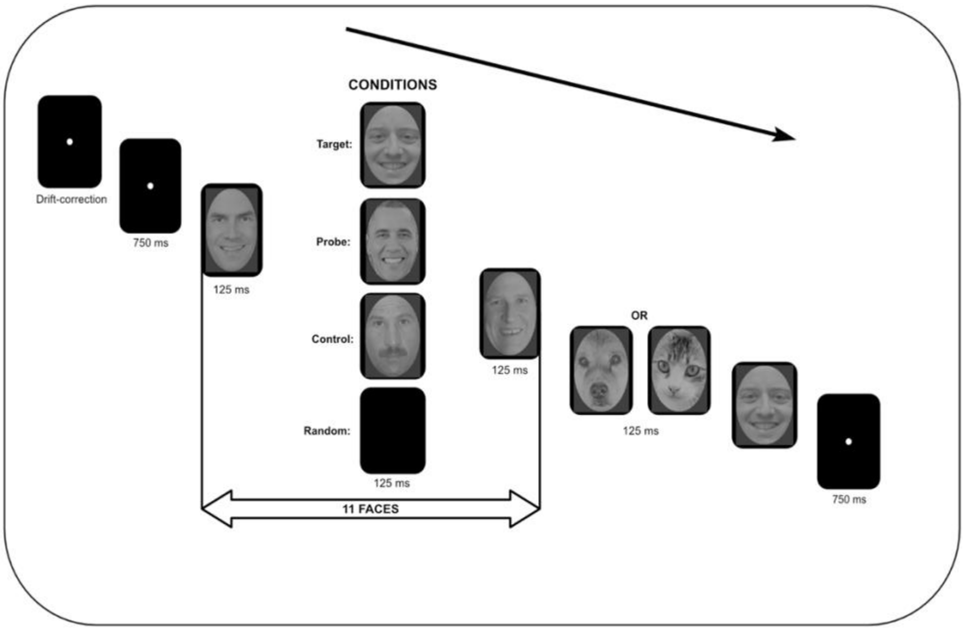
stimulus, a photo of Barack Obama (Figure 1a, Souza, 2012) was placed in the stream. Irrelevant distractor faces in each trial were selected randomly from our overall list of faces. A control face was also randomly selected for each participant before the experiment to check whether there would be no difference with the no target condition. This was done to ensure that a random face – equal in presentation frequency to the probe and target – would not become familiar and elicit a pupil dilation. Pictures were all monochrome and did not represent any body part of the person besides the face. All faces were presented in the center of a gray-colored screen inside a fixed oval shape (140 x 200 pixels), as shown in Figure 1a-b. The visual angle for each picture was 11.42° in height and 6.82° in width. Using custom Matlab scripts the photo of Obama was processed to appear similar to the unfamiliar faces in, for instance, contrast and brightness.

Procedure

After inspecting the pupil traces, we decided on a cut-off value of five or more removed trials in the processed data due to blinks. We calculated it by removing any participant who had exceeded three median absolute deviations from the median (*Mdn* = 1, *MAD* = 1). As a result, the data of 10 out of 53 participants were deleted, leading to a final inclusion of 43 participants.

Figure 2

Visual Representation of a Singular Trial Sequence



Prior to the start of the experiment, participants were shown the target face and were required to finish ten practice trials to get familiar with the task. The experiment consisted of three consecutive blocks and each block of trials consisted of 32 trials resulting in a total of 96 trials. As shown in Figure 2, before each trial a fixation dot was shown for approximately 500-1000 milliseconds to capture the attention of the participant and to enable a baseline pupil size. The participants were then shown 11 faces concluding the trial with either a cat or a dog in an RSVP stream, each for 125 ms. During the stream, four conditions were possible. Participants were either shown the target face, the probe, a control face, or no target face, which was presented randomly on position 5, 6, or 7 for every trial. After the sequence of faces, either a dog or a cat was shown to keep the participant's attention fixed throughout the entire stream. Finally, another fixation dot was presented. Overall, the RSVP trial duration was 3000 ms counting from the first to the last fixation dot.

Tasks

After each RSVP, the participants were initially asked to indicate whether a picture of a dog or a cat was shown. This was done to ensure the participant remained focused during the whole trial. They were instructed to press '*m*' to indicate if they saw a dog and press '*c*' if they saw a cat. After this first task, the target face was shown once more and the participants were asked *"Did you see this face?*". When they did not, they had to press '*c*' and when they did, they had to press '*m*'. The order of response buttons was counterbalanced over participants. After each response, the participant was shown whether they were correct or incorrect. The answer to the first question would either lead to an increase or decrease of the total score by five points. The answer to the second task, of whether they had seen the target face, resulted in either 10 point increase or 10 point decrease.

At the end of the experiment, the participants were given two questions. These were *"Did you notice the face of a famous person was shown sometimes?"* and *"If you had to guess which famous person we showed, who would it be?",* respectively. The latter, which was an open question, was added to ensure that participants who selected '*yes*' on the former question did, in fact, see Obama.

Design

In the RSVP sequence, 11 faces were shown. There were four conditions, one where the target face was present, one where the probe (Obama) was present, a condition where no target was present, and a condition where a control face was presented. This control face was randomly selected from the database and served as a baseline that was similar in presentation frequency to the target and the probe. Each condition was shown either on position 5, 6, or 7. The four conditions in combination with three possible positions, and two possible animals resulted in 24 different combinations. Each possible combination was presented four times which resulted in 96 trials.

Data processing and analysis

To determine to what extent participants were able to sustain their attention during the trials, we first analyzed the accuracy of the responses to the question of whether they saw a dog or a cat at the end of each trial. We then baselined the pupil sizes by taking the average size from the first three samples in each trial after the T1 presentation and subtracting this baseline value from all other samples in that trial. The window that has been chosen for the analysis is based on the study by Göl, Jansen, and Rasztar (2022), where it was found that the biggest difference in pupil response occurs between 640 ms and 920 ms after the T1 presentation.

As an exploratory analysis, we used a Shapiro-Wilk test for the normality assumption. Afterward, we used two nonparametric Mann-Whitney U tests on the group level to check for differences that may exist in the pupil size. In more detail, we used the means of the baseline-corrected pupil size [during the analysis window] as a dependent measure and condition as a fixed effect, to find if a difference exists between the pupil sizes when comparing target with no target, probe with no target and control with no target.

Results

In our study, we had two predictions. First, we checked whether the pupil size would be larger after the target face was presented, in comparison to when no target was presented. This would indicate that the task-relevant stimulus, the face that the participant had to actively look for, elicited a reaction that can be detected by using pupillometry in an RSVP. The second prediction was related to the detection of the famous face. If the pupil size for the probe would be larger than in the no target condition, then it would entail that the task-irrelevant, familiar face (even if subliminal) had elicited a physiological reaction. Supported by results of the experiment, this prediction could provide support for the use of pupil size and RSVP to detect subliminal salience or even concealed information.

Task Performance

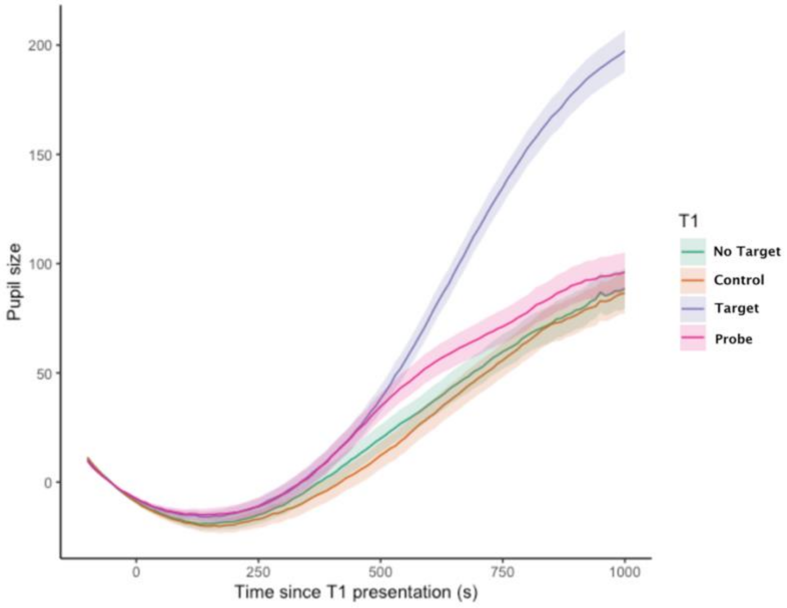
On average participants were able to respond well above the guessing rate in both tasks. Regarding the first task (question: "*Did you see a cat or a dog?*") participants were able to indicate with an accuracy of 99% whether a picture of a cat or a dog was shown at the end of the RSVP. When we inspected the performance on the second task (question: "*Did you see this face?*" accompanied with the target face). Participants responded correctly to the presence of the target in 57% of the cases and to the absence in 95% percent of the cases. To the two end questions *"Did you notice the face of a famous person was shown sometimes?*" and *"If you had to guess which famous person we showed, who would it be?*", of the 43 participants 38 gave an answer (79%) and of these 38 participants 45% indicated that they saw Obama.

Pupil traces

Pupil traces in response to the presentation of T1 are shown in Figure 3. Visual inspection of the pupil size after the presentation of the target face shows a difference starting from approximately 500ms and showing an upward trend until 1000ms. Likewise, the pupil response to the presentation of the probe diverged from no target after approximately 500 ms. This effect, however, is not as large

Figure 3

Pupil trace visualization for the four conditions



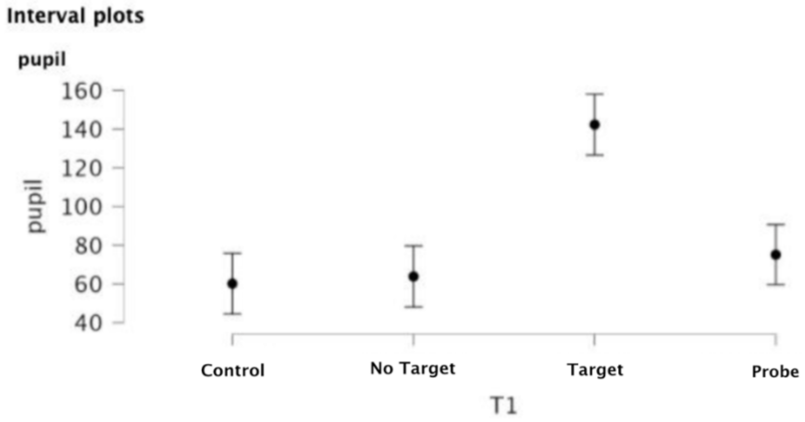
compared to the target condition. Inspecting control and no target condition does not show a big difference.

The mean values of participant's pupil sizes in the four different conditions were measured in the time window of 640 to 920 ms. The four conditions and their respective mean values, control (M = 60.15, SD = 253.28), no target (M = 63.88, SD = 255.76), target (M = 142.18, SD = 254.12), probe (M = 75.12, SD = 251.48) can be seen in Figure 4.

As the normality assumptions were violated, for the group level analysis two Mann-Whitney U tests were conducted to investigate differences in pupil size between the target, the probe, and the no target condition. According to our first hypothesis, pupil sizes in the target condition would be larger compared to the no target condition. A significant difference between target – no target supported our hypothesis (U = 412934, p < .001, d = -0.179). For the comparison probe – no target a non-significant p-value was found (U = 490915, p = .246, d = -0.030). This result does not support our hypothesis that in the trials in which the probe was presented, a larger pupil than in the no target condition would be observed. Lastly, we tested whether there was a significant difference between the control and the no target condition. A significant difference between these two conditions would

Figure 4

Mean pupil size during the 640 - 920 ms window over four conditions



Note. Estimated mean pupil size during the window (640 - 920ms) for control, no target, target, and probe conditions.

indicate that participants became familiar with the control condition as well. This difference, however, was non-significant (U = 496888, p = .584, d = -0.014).

Discussion

Previous studies have investigated the validity of Rapid Serial Visual Presentation (RSVP) and pupillometry to detect deception in Concealed Information Testing (CIT) (Bowman et al., 2013; Chen et al., 2021). Specifically, Chen et al. (2021) validated pupillometry on the group level as an effective method to recognize different reactions to stimuli in an RSVP setting by using names as stimuli. These promising results led to the development of the current study which aims at replicating their findings by adding a new element to this field of research: the use of pupillometry and RSVP in combination with face stimuli for the detection of subliminal salience. Our findings generally showed that pupil response in RSVP can be considered a valid measurement to detect familiarity with a salient stimulus. On the other hand, tests for the detection of subliminal salience did not give significant results in the context of pupil response. Hence, while pupillometry seems to be a valid instrument to measure familiarity with a stimulus that participants are searching for, it cannot yet be considered valid for the detection of subliminally salient stimuli.

In the first part of our analysis, we focused on the extent to which participants were able to identify the occurrence of the target stimulus, a face made familiar to the participant at the beginning of the experiment on which they were instructed to focus their attention. The response accuracy of 57%, only slightly above the 50% chance level, is substantially lower compared to the one registered by previous pupillometry research using name stimuli (Chen et al., 2021). A possible explanation could be found in the fact that, in an RSVP context, faces are more complex stimuli compared to names as they contain more factors changing from face to face that need to be processed (Damasio et al., 1990).

We compared then the means of the pupil sizes of the four conditions in the window 640 - 920 ms (target, probe, no target, and control). Even though the accuracy for the detection of the target was low, the overall mean difference in pupil size for this condition was greater than all the other ones, from which we can infer that either the effect was driven by the participants' high accuracy rate or that, even when participants were not consciously perceiving the target face, an autonomic response was generated for familiarity with that face. Specifically, in line with previous research (Chen et al., 2021), our results showed a significant difference between target and no target conditions, supporting our first hypothesis that pupil size would increase at target face presentation. On the other hand, while the means of pupil size were visibly different (Figure 3), the statistics on the comparison of the pupil sizes on the probe and no target conditions showed non-significant differences. Our second hypothesis was not supported by this study; hence we can conclude that the change in pupil size when the probe was shown was not big enough to validate pupillometry in RSVP as a trustworthy method to detect subliminal salience.

The present study is characterized by some limitations. As only 45% of participants were able to identify the face of Obama, we can suppose that this face was not familiar enough to all participants to be recognized. In fact, for the face to break through into awareness and to detect subliminal salience by looking at pupil size differences, the face needs to be familiar enough to generate an autonomic response. It could be that by using a face that is more familiar to the participants (e.g., the face of their mother) the detection rate could increase and the statistics might become significant. Furthermore, considering the low accuracy rate, we consider that the task was too difficult and the speed of the RSVP too high. However it is complicated to suggest a different alternative: for stimuli to break through into awareness and generate a response, the RSVP needs to be fast.

Future research should focus on finding the right balance for stimuli rate presentation, i.e., to find the right spot in which it is not too slow or too fast so that stimuli are consciously seen but still perceived and processed. Furthermore, since with our analysis we were able to observe a trend of not significant but existing differences between the four groups, additional attention should be focused on this research topic to find significant differences. As Chen et al. (2021) suggested, a change in the research design could be a solution. Such changes could include adjusting the rate of presentation and the type of stimuli presented.

Despite the listed limitations, the findings of the present study are consistent with previous research conducted on the topic. Even though further investigation is needed to validate the utility of this method and increase the sensitivity of the tasks, the use of pupil measurements in the RSVP context seems to be a promising means for detecting concealed information.

References

Alsufyani et al., (2018). Breakthrough Percepts of Famous Faces. *Psychopathology*. https://doi.org/10.1111/psyp.13279

Ben-Shakhar, G., & Elaad, E. (2003). The validity of psychophysiological detection of information with the Guilty Knowledge Test: a meta-analytic review. *The Journal of applied psychology*, *88 1*, 131-51. https://doi.org/10.1037/0021-9010.88.1.131

Block, J. D., Rouke, F. L., Salpeter, M. M., Tobach, E., Kubis, J. F. & Welch L. (1952). An Attempt at Reversal of the Truth-lie Relationship as Measured by the Psychogalvanic Response. *The Journal of Psychology, 34:1*, 55-66. https://doi.org/10.1080/00223980.1952.9916105

Bowman H., Filetti M., Janssen D., Su L., Alsufyani A., et al. (2013). Subliminal Salience Search Illustrated: EEG Identity and Deception Detection on the Fringe of Awareness. https://doi.org/10.1371/journal.pone.0054258

Brinke, L., Stimson, D., & Carney, D. R. (2014). Some Evidence for Unconscious Lie Detection. *Psychological Science*, *25*(5), 1098–1105. https://doi.org/10.1177/0956797614524421

Chen, I., Karabay, A., Mathôt, S., Bowman, H., & Akyürek E. G., (2021). Concealed identity information detection with pupillometry in rapid serial visual presentation**.** https://doi.org/10.1101/2021.06.18.448944

Damasio, A. R., Tranel, D., & Damasio, H. (1990). Face agnosia and the neural substrates of memory. Annual review of neuroscience, 13, 89–109. https://doi.org/10.1146/annurev.ne.13.030190.000513

Ekman, P., & O'Sullivan, M. (1991). Who Can Catch a Liar? American Psychologist, 46, 913-920. http://dx.doi.org/10.1037/0003-066X.46.9.913

Jackson, M. C. & Raymond, J. E. (2006). The Role of Attentions and Familiarity in Face Identification. *Perception & Psychophysics, 68 (4)*, 543-557. https://doi.org/10.3758/BF03208757

Keeler, L. (1933). Scientific methods of crime detection with the polygraph. Kansas Bar Ass. J., 1933, 2, 22-31.

Mart Bles & John-Dylan Haynes (2008) Detecting concealed information using brain-imaging technology, Neurocase, 14:1, 82-92. https://doi.org/10.1080/13554790801992784

Matsuda, I., Ogawa, T., & Tsuneoka, M. (2019). Broadening the Use of the Concealed Information Test in the Field. *Frontiers in psychiatry*, *10*, 24. https://doi.org/10.3389/fpsyt.2019.00024

Rosenfeld, J. P., Soskins, M., Bosh, G., & Ryan, A. (2004). Simple, effective countermeasures to P300-based tests of detection of concealed information. *Psychophysiology*, *41*(2), 205–219. https://doi.org/10.1111/j.1469-8986.2004.00158.

Vrij, A., Granhag, P. A., & Porter, S. (2010). Pitfalls and Opportunities in Nonverbal and Verbal Lie Detection. *Psychological Science in the Public Interest*, *11*(3), 89–121. https://doi.org/10.1177/1529100610390861