

TECHNICAL NOTE

**ID-CHECK:**

**ONLINE CONCEALED INFORMATION TEST REVEALS TRUE IDENTITY**

*In press: Journal of Forensic Science*

Bruno Verschuere<sup>1, 2, 3\*</sup> & Bennett Kleinberg<sup>1</sup>

Affiliation:

1. Department of Clinical Psychology, University of Amsterdam, The Netherlands
2. Department of Psychology, Ghent University, Ghent, Belgium
3. Department of Clinical Psychological Science, Maastricht University, Maastricht, The Netherlands

\* Corresponding author: Bruno Verschuere, Department of Clinical Psychology, University of Amsterdam, Weesperplein 4, 1018 XA, Amsterdam, Netherlands. E-mail: [b.j.verschuere@uva.nl](mailto:b.j.verschuere@uva.nl). Phone: +3125256799

Highest academic degrees:

B. Verschuere: Ph.D; B. Kleinberg: B.Sc.

Abstract (150 words)

Background: The Internet has already changed people's lives considerably and is likely to drastically change forensic research. We developed a web-based test to reveal concealed autobiographical information. Initial studies identified a number of conditions that affect diagnostic efficiency. By combining these moderators, the present study investigated the full potential of the online ID check.

Methods and results: Participants ( $n = 101$ ) tried to hide their identity and claimed a false identity in a reaction time-based Concealed Information Test. Half of the participants were presented with personal details (e.g., first name, last name, birthday), whereas the others only saw irrelevant details. Results showed that participants' true identity could be detected with high accuracy (AUC = .98; overall accuracy: 86 - 94%).

Conclusions: Online memory detection can reliably and validly detect whether someone is hiding their true identity. This suggests that online memory detection might become a valuable tool for forensic applications.

Keywords

Forensic science; deception; memory detection; lie detection; polygraph; concealed information test; reaction times; guilty knowledge

**Running head**

ONLINE ID CHECK

Google, selfies, Facebook, Twitter, and e-mail. In the past two decades, the Internet had changed our lives considerably. It also changes forensic science. Investigators can quickly exchange information and large databases become readily available. Albeit sometimes assisted by technology, contemporary lie detection techniques are near exclusively based upon face-to-face interviewing. In the present study, we explore the accuracy of a web-based test that aims to unmask one's true identity.

Claiming a false identity is not uncommon in the forensic context. Consider the following example. Habteab Berhe Temanu entered the US as an Ethiopian refugee in 2002. In 2011, Homeland Security was informed that someone recognized Habteab Berhe Temanu as being a criminal of war. The informant told Homeland Security that the man's real identity was Keefelegn Alemu Worku, who had worked as a prison guard during the 'Red Terror' that had cost the lives of ten thousands of Ethiopian citizens in the 1970ies. In such a case, a well-researched technique called the Concealed Information Test (CIT; (1); for a review see (2)) can be used to reveal the concealed true identity. The CIT consists of a series of multiple-choice questions, such as 'Is your name: Habteab? [buffer] Louam? [control] Merille? [control] *Keefelegn*? [critical] Semere? [control] Ras? [control], see Footnote1'. Rather than relying on the suspect's overt answer, autonomic (e.g., skin conductance recorded with a polygraph), electrophysiological (e.g., the P300 brain wave from the EEG), or neural (e.g., the BOLD response obtained with fMRI) responses to the alternatives are being measured (3). Stronger physiological responses to the critical identity details related to Keefelegn Alemu Worku than that to the control items, provide an indication that the examinee may actually be Keefelegn Alemu Worku

While physiological responses can have high validity to determine one's true identity, their administration is technically challenging. Because of its ease of application, there has

been renewed attention for short computerized tasks relying on reaction times (RTs). Seymour et al. (4) were the first to show that RTs can provide a quick yet accurate index of concealed information, finding that concealed knowledge is reacted slower upon than to similar, yet irrelevant details. This RT-based test can also be applied as an ID-check, revealing the true identity with accuracy similar to that of autonomic nervous system measures (5). We have modernized this RT-based ID-check to a web-based version (6). In a first set of studies (6)(7), participants were asked to conceal their true identity when taking the online concealed autobiographical information test. These studies identified important factors that affect the test's accuracy, with better detection (a) for highly salient details (i.e., details that are of high personal significance to the examinee) than for low salient details, (b) when randomly presenting all alternatives to all questions (e.g., regarding one's first name, last name) as compared to sequentially presenting them question-by-question, and (c) having a sufficiently long test (i.e., at least 240 trials, (8)). Here we combined these moderators to investigate the full potential of the online ID check.

## Method

The study was approved by the ethical committee of the Department of Psychology of the University of Amsterdam (2014-CP-3389).

### *Participants*

One-hundred and one undergraduate students participated in this study. Nearly all had Dutch as mother tongue (96%). We applied the same exclusion criteria as Kleinberg and Verschuere (6, 7), and excluded data from participants (1) with double occurring IP addresses (2 exclusions; leaving  $n = 99$ ), (2) with an error rate of 50% or more on any of the three item types (26 exclusions). The final sample consisted of 73 participants that either were (knowledgeable condition;  $n = 42$ ;  $M_{\text{age}} = 19.81$  years,  $SD_{\text{age}} = 1.81$ ; 76% females) or were

not (naïve condition;  $n = 31$ ;  $M_{\text{age}} = 19.35$  years,  $SD_{\text{age}} = 1.02$ ; 87% females) presented with their own personal details in the online test. The conditions did not differ in gender,  $X^2(1) = 1.37$ ,  $p = .244$ , or age,  $t(71) = 1.26$ ,  $p = .213$ ,  $d_{\text{between}} = 0.30$ .

### *Procedure*

Participants were recruited through a dedicated university website that provides course credits to first year bachelor students from psychology, psychobiology and communication sciences for participation. Participants took the test at their own time, on their own computer. In specific, participants accessed the website [http://www.lieresearch.com/?page\\_id=689](http://www.lieresearch.com/?page_id=689), and, after providing informed consent, provided demographic (i.e., gender, age, mother tongue) as well as identity-related details (i.e., first name, last name, course, birthday, and country of origin) that were to be used as critical details in the test. We also asked to indicate one other significant first name, last name, course, birthday, and country of origin from a list of possible control items. We assured that these items were not be used as control items, as to avoid that the control items were of significance to the participant.

In the next step, participants were informed that their task was to hide their own identity and adopt a false ID (e.g., First name: Lisa; Last name: Jansen; Course: Criminology; Birthday: 19 May; Country of origin: Ghana). They were required to rehearse and recall their false identity until they did so without error. They were informed that they would do a memory detection test where they should deny recognition of all items but those pertaining to their false identity. The memory detection test began with a three-step practice procedure, followed by the full test. This procedure was similar to that used by Kleinberg and Verschuere (6).

*Concealed Information Test.* Words were rapidly flashed on the computer screen one by one (for 1500ms or until button press; inter-stimulus interval varied randomly between 250-500-750ms), and asked to answer as fast as possible YES or NO to the question ‘Is this you?’. Thus, participants were requested to answer YES only to their false identity, and NO to all other items, including their own identity. The task was practiced in 3 practice phases, and participants had to repeat each practice phase until they met the following criteria: a target error rate below 50%, a mean RT of less than 800ms, and less than 20% of their trials being below 150ms. The test consisted of 600 trials. For all participants, there were 100 items related to their false identity, and 400 irrelevant items (of the same category, but unrelated to either the false or real identity). The remaining 100 items were either the participant’s own personal details (knowledgeable condition), or also irrelevant items (naive condition). A TOO SLOW message appeared when not meeting the 800ms response deadline. And a WRONG message appeared for behavioural errors.

*Saliency ratings.* After the CIT, participants rated the 5 item categories (first name, last name, course, birthday and country of origin) and 5 other categories (e.g., favourite animal) on their personal relevance using a 9-point Likert scale (1 = not relevant at all, 9 = absolutely relevant)

*End.* Finally, all participants received their test results (based upon their Cohen’s  $d_{CIT}$ ; see Results Section), were debriefed, were thanked for participation and exited the task.

## **Results**

The data are publically available on the Open Science Framework on [osf.io/cg5es](https://osf.io/cg5es)

### *Manipulation check*

Participants judged the 5 categories used in the test to be of ‘(some) personal’ significance ( $M = 6.80$ ,  $SD = 1.93$ ).

### *Group analyses*

We excluded trials (1) related to the false identity, (2) with incorrect responses, and (3) with RTs smaller than 150ms and larger than 800ms. The remaining data points were subjected to a 2 (naïve versus knowledgeable) x 2 (personal details versus irrelevant) mixed ANOVA. The ANOVA showed that the significant main effects subsumed under the predicted 2-way interaction,  $F(1, 71) = 111.07, p < .001, f = 1.25$ , see Figure 1. This interaction indicated that the personal-irrelevant difference is larger in the knowledgeable condition than in the naïve condition,  $t(66.46) = 11.42, p < .001, d_{\text{between}} = 2.50$  ( $d_{\text{between}}$  and  $d_{\text{within}}$  refer to Cohen's  $d$  effect size estimate for between-subjects and within-subjects comparisons, respectively, see (9)). The personal-irrelevant difference was significant and large in the knowledge condition,  $t(41) = 13.40, p < .001, d_{\text{within}} = 2.05$ , but not the naïve condition,  $t(30) = 0.31, p = .76, d_{\text{within}} = -0.08$ .

FIGURE 1 AROUND HERE

### *Individual Classification*

Following the procedure used in Noordraven and Verschuere (10), we calculated Receiver Operating Characteristics analyses (ROC) to examine how well the personal-irrelevant RT difference  $d_{\text{CIT}}$  allowed to discriminate knowledgeable from naïve participants. We defined  $d_{\text{CIT}}$  as  $(M_{RT(\text{probes})} - M_{RT(\text{irrelevants})}) / SD_{RT(\text{irrelevants})}$ , so that positive values are indicative of recognition (e.g. (6)(7)). The ROC analysis plots sensitivity against the false positive rate across all possible cut-off points. The corresponding area under the curve (AUC) provides an index of diagnostic efficiency with an AUC value of .5 indicating that the test

performs at chance level and higher values are indicative of higher diagnostic power, with 1 indicating perfect performance. The area under the curve was .98 (95% CI: .95 – 1; Figure2).

#### FIGURE2 AROUND HERE

The ROC curve displays the balance between sensitivity and specificity for all possible cut-off points. Sensitivity and specificity for any single cut-off point can be inferred from the ROC curve. To illustrate how the ROC translates into hit rates, Table1 displays the hit rates for the cut-off points examined by Noordraven and Verschuere (10). In addition, Table1 displays the hit rates for the “optimal” cut-off point based upon Youden’s *J* statistic that is derived by calculating the criterion value ( $d_{CIT}$ ) where the distance to the ROC identity line (i.e. random classification) is maximal (11)(12). Because such optimal cut-off point capitalizes on chance, we split our sample semi-randomly in half, with one half (‘model-building sample’) used to calculate Youden’s *J* statistic, and the other half (‘validation sample’) for cross-validation (Footnote2). In the model-building sample, Youden’s index learned that  $d_{CIT} = 0.29$  was the cut-off point that, if used as criterion to indicate recognition or no recognition, allowed to correctly classify 19 out of 21 knowledgeable participants (sensitivity: 90%; false negatives:  $n = 2$  or 10%), and to correctly classify 15 out of 15 naïve participants (specificity: 100%; false positives:  $n = 0$ ), resulting in an overall accurate classification of 34 of the 36 participants (94%). In the cross-validation sample, with  $d_{CIT} = 0.29$  as a cut-off, 16 out of 21 knowledgeable participants (sensitivity: 76%; false negatives:  $n = 5$  or 24%), and 16 out of 16 naïve participants (specificity: 100%; false positives:  $n = 0$ ) were classified correctly, resulting in an overall accurate classification of 32 of the 37 participants (86%).

#### TABLE1 AROUND HERE

## **Discussion**

In the present study, we examined whether we could detect the true identity of participants claiming a false ID. Rather than relying on face-to-face interviewing or physiological measures, we used an online test that tracked participant's RTs. By combining moderators that were identified during the development of this test (6)(7), we examined its full diagnostic potential. The results showed that the online test was highly accurate (AUC: .98, overall hit rate: 86 - 94%).

Several aspects need to be taken into account when interpreting the high accuracy. First, the accuracy in the present study may represent the upper limit of the diagnostic efficiency that can be obtained with online testing because for forensic applications it is difficult to establish a sufficient number of highly salient details (13). When the suspect was told of being accused to be 'Kefelegn Alemu Worku', the name itself can no longer be used in the online ID-check. Leakage of that information makes it salient and recognizable, also to the innocent suspect (14). Thus, in case the suspect has been explicitly informed on the presumed false identity, the test can only rely upon personal details (e.g., last residential address, name of pet or close relatives) that have not been leaked and the challenge is to assure that these items are of sufficiently high saliency to evoke a marked response. Second, to assure data quality, we set strict exclusion criteria, implying that no judgment was made for a substantial part of our sample (28%). These exclusion criteria are not carved in stone and it will be important to establish criteria that provide a good balance between maximal inclusions and high quality data. Third, online testing in the field will require verification of who takes the test. Depending on the application this may be accomplished by a password, a webcam, iris-scan, or a collaborator that is physically present with the examinee. There may be merit to online testing, even when a collaborator is required to control who is taking the

test. When a local police officer is with the examinee, the forensic expert can develop, administer, analyse and report upon the test without the need to be physically present, thereby saving time and money.

While diagnostic efficiency needs to be established under more realistic conditions, the current findings indicate that there is promise to the online detection of concealed autobiographical information. Under specific conditions, the online ID-check can reach high accuracy and may help in identity verification.

## References

- (1) Lykken DT. The GSR in the detection of guilt. *J Appl Psychol* 1959; **43**: 385–388.
- (2) Verschuere B, Ben-Shakhar G, Meijer E, editors. Memory detection: Theory and application of the Concealed Information Test. Cambridge: Cambridge University Press; 2011.
- (3) Meijer E, Selle NK, Elber L, Ben-Shakhar G. Memory detection with the Concealed Information Test: a meta-analysis of skin conductance, respiration, heart rate, and P300 data. *Psychophysiology* 2014; **51**: 879-904.
- (4) Seymour TL, Seifert CM, Shafto MG, Mosmann AL. Using response time measures to assess “guilty knowledge”. *J Appl Psychol* 2000; **85**: 30-37.
- (5) Verschuere, B, Crombez G, Degrootte T, Rosseel, Y. Detecting concealed information with reaction times: Validity and comparison with the polygraph. *Appl Cogn Psychol* 2010; **24**: 991-1002.
- (6) Kleinberg B, Verschuere B. Memory detection 2.0: The first web-based memory detection test. *PLoS ONE* 2015.
- (7) Verschuere B, Kleinberg B, Theocharidou K. RT-based memory detection: Item saliency effects in the one probe and multiple probe protocol. *J Appl Res Mem Cogn* 2015; **4**(1): 59-65.
- (8) Kleinberg B, Verschuere B. *Online reaction time-based memory detection: Test length and motivation to avoid detection*. University of Amsterdam, Amsterdam: Unpublished manuscript 2015.
- (9) Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Front Psychol* 2013; **4**: 863.

- (10) Noordraven E, Verschuere, B. Predicting the sensitivity of the Reaction Time-based Concealed Information Test. *Appl Cogn Psychol* 2013; **27**: 328-335.
- (11) Youden W. Index for rating diagnostic tests. *Cancer* 1950; **3**: 32-50.
- (12) Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J, Müller M. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 2011; **12**.
- (13) Podlesney JA. A paucity of operable case facts restricts applicability of the guilty knowledge technique in FBI criminal polygraph examinations. *Forensic Science Communications* 2003, **5**: Retrieved 27 November, 2014 from <http://www.fbi.gov/about-us/lab/forensic-science-communications/fsc/july2003/index.htm/podlesny.htm>
- (14) Bradley MT, & Warfield JF. Innocence, information, and the Guilty Knowledge Test in the detection of deception. *Psychophysiology* 1984; **21**: 683–689.



Additional information and reprint requests

Bruno Verschuere, Department of Clinical Psychology, University of Amsterdam,

Weesperplein 4, 1018 XA, Amsterdam, Netherlands. E-mail: [b.j.verschuere@uva.nl](mailto:b.j.verschuere@uva.nl). Phone:

+3125256799

## Footnotes

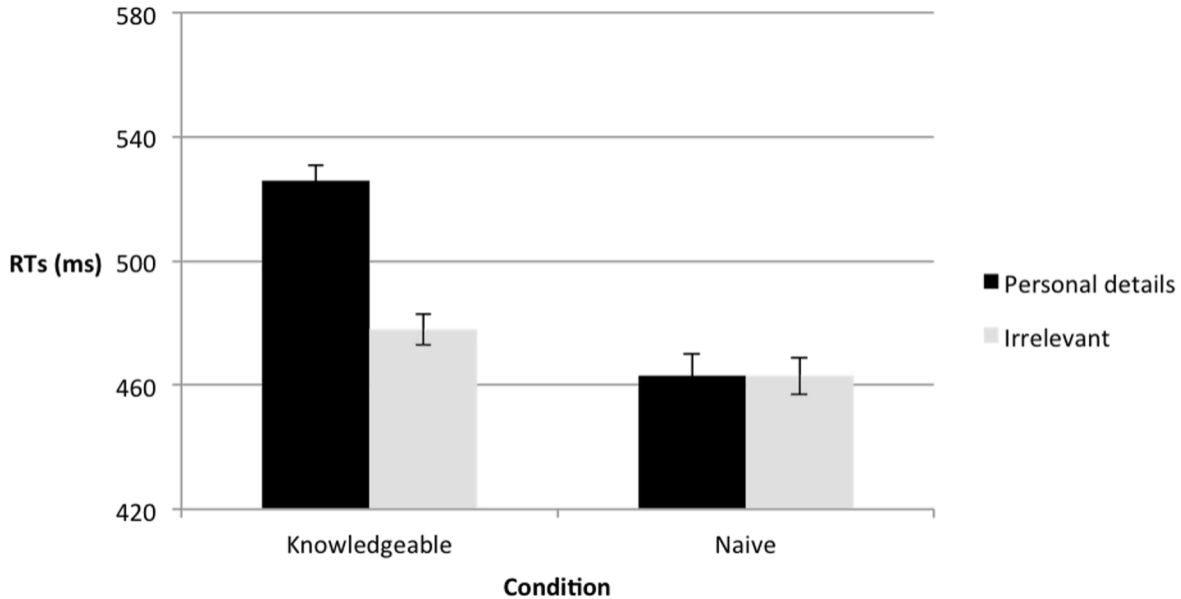
1. The suspect's false name (Kefelegn) is also likely to elicit a strong physiological response, and is therefore presented first, as a buffer item, and excluded from the analysis that is based upon the comparison of the critical versus the control items.
2. Randomisation was as follows: half of the participants in each condition were semi-randomly allocated either a 1 or a 2. Of the resulting sub-samples, the first was used as model-building sample ( $n = 36$ ) and the second as validation sample ( $n = 37$ ).

Table 1. Diagnostic efficiency (sensitivity and specificity) of Online Concealed Information

Test for specific cut-off points

Cut-off ( $d_{CT}$ )	0.00		0.20		<b>0.29*</b>		0.50		0.80	
	Sens.	Spec.	Sens.	Spec.	<b>Sens.</b>	<b>Spec.</b>	Sens.	Spec.	Sens.	Spec.
Model-building sample	100%	67%	95%	93%	<b>90%</b>	<b>100%</b>	43%	100%	10%	100%
Validation sample	100%	56%	81%	94%	<b>76%</b>	<b>100%</b>	43%	100%	24%	100%

Note. \* = Youden's  $J$ , based upon model-building sample; Sens. = Sensitivity; Spec. = Specificity.



ROC curve for criterion  $d_{CIT}$

