How to Accurately Detect Autobiographical Events

Giuseppe Sartori, Sara Agosta, Cristina Zogmaister, Santo Davide Ferrara and Umberto Castiello

Psychological Science 2008 19: 772
DOI: 10.1111/j.1467-9280.2008.02156.x

Research Article

How to Accurately Detect Autobiographical Events

Giuseppe Sartori,[1] Sara Agosta,[1] Cristina Zogmaister,[2] Santo Davide Ferrara,[3] and Umberto Castiello[1]

[1]Psychology Department, [2]Developmental and Social Psychology Department, and [3]Environmental Medicine and Public Health Department, University of Padua

ABSTRACT—*We describe a new method, based on indirect measures of implicit autobiographical memory, that allows evaluation of which of two contrasting autobiographical events (e.g., crimes) is true for a given individual. Participants were requested to classify sentences describing possible autobiographical events by pressing one of two response keys. Responses were faster when sentences related to truly autobiographical events shared the same response key with other sentences reporting true events and slower when sentences related to truly autobiographical events shared the same response key with sentences reporting false events. This method has possible application in forensic settings and as a lie-detection technique.*

Autobiographical memory is an individual's ability to remember events he or she has experienced directly. The majority of studies on autobiographical memory have focused on the amount of information that is remembered (e.g., Crovitz & Schiffman, 1974). In addition, indirect methods of evaluating the veracity of reported autobiographical events have proven useful, and the detection of accurate memories may be valuable for detecting lies (e.g., Lykken, 1960).

Detecting lies plays an essential role during crime investigations and criminal trials. Two tests frequently used for lie detection are the Control Question Test (Moore, Petrie, & Braga, 2003) and the Guilty Knowledge Test (GKT; Lykken, 1960, 1998). The former is based on differential patterns of physiological activation (e.g., heart rate) that suspects exhibit when asked direct questions about the crime (e.g., "Did you do it?") versus neutral questions. The latter uses physiological responses to multiple-choice questions, each including a "relevant" answer (e.g., feature of the crime under investigation) and several "control" answers that cannot be discriminated from the relevant answer by an innocent suspect (Lykken, 1998). Typically, guilty suspects exhibit larger physiological responses for relevant than for control alternatives. Recently, researchers have considered how these two tests might be used in conjunction with functional magnetic resonance imaging (fMRI; Ganis, Kosslyn, Stose, Thompson, & Yurgelun-Todd, 2003; Langleben et al., 2005; for an account regarding the use of fMRI for lie detection, see American Academy Symposium, 2007). However, despite such methodological advancements, these methods are still plagued by poor specificity and sensitivity (e.g., Iacono & Lykken, 1999).

The Implicit Association Test (IAT; Greenwald, McGhee, & Schwarz, 1998) could provide an important step forward. For instance, Gray and her colleagues have elegantly illustrated how the IAT can be fruitfully applied in a forensic setting (Gray, Brown, MacCulloch, Smith, & Snowden, 2005; Gray, MacCulloch, Smith, Morris, & Snowden, 2003). They showed that it can correctly identify psychopathic murderers through their implicit beliefs and can also identify pedophilic attitudes.

Another adaptation of the IAT that has the potential to be used in forensic settings is the Timed Antagonistic Response Alethiometer (TARA; Gregg, 2007). By using a task requiring speeded classification of sentences and manipulating response incongruity, the TARA classifies the respondent as a truth teller or a liar.

Here we present a new IAT-based methodology we call the autobiographical IAT (aIAT). The aIAT allows one to evaluate which of two contrasting autobiographical events is true for a given individual. This is accomplished by requiring the respondent to complete two critical blocks of categorization trials, each of which pairs a different potentially autobiographical event with true events. Because pairing of a truly autobiographical event with true events should facilitate responses, the specific pattern of response times (RTs) in the two blocks indicates which autobiographical event is true and which is false. We report six experiments that tested the accuracy of the aIAT in identifying truly autobiographical events.

Address correspondence to Giuseppe Sartori, Via Venezia 8, 35100 Padova, Italy, e-mail: giuseppe.sartori@unipd.it.

TABLE 1

Schematic Description of the Autobiographical Implicit Association Test (aIAT) Used in All Experiments

Response key	Block 1: logical discrimination	Block 2: initial autobiographical discrimination	Block 3: initial double categorization	Block 4: reversed autobiographical discrimination	Block 5: reversed double categorization
"A" key	True sentences	Guilty sentences	True and guilty sentences	Innocent sentences	True and innocent sentences
"L" key	False sentences	Innocent sentences	False and innocent sentences	Guilty sentences	False and guilty sentences

Note. The difference in average response time (RT) between Block 3 and Block 5 was used to identify autobiographical events that were true for the respondents. If RT was faster in Block 3, guilty sentences were true for that respondent; if RT was faster in Block 5, innocent sentences were true for that respondent. Note that the order of Blocks 3 and 5, and of Blocks 2 and 4, was reversed for half of the participants.

GENERAL METHOD

The experimental procedures were approved by the ethics committee of the University of Padua and were in accordance with the Declaration of Helsinki.

Methods were the same for all six experiments, except as specified. The computerized task consisted of five separate blocks of categorization trials (see Table 1). In each trial, a stimulus was presented at the center of a computer monitor, and participants were requested to classify the stimulus as quickly and accurately as possible, by pressing one of two labeled keys. Stimuli were sentences of variable length, each describing a potentially autobiographical fact. In Block 1 (20 trials; logical discrimination), participants classified sentences as true or false. They pressed the "A" key if the sentence was true (e.g., "I'm in front of the computer") and the "L" key if the sentence was false (e.g., "I'm at the beach"; the true and false sentences were the same for all experiments and are listed in Table 2). In Block 2 (20 trials; initial autobiographical discrimination), participants classified sentences along the critical dimension guilty-innocent (see the sections on the individual experiments for explanations of what constituted guilty and innocent sentences). They pressed the "A" button if the sentence was of the guilty type (e.g., "I took cocaine recently") and the "L" button if the sentence was of the innocent type ("I never made use of cocaine"). In Block 3 (60 trials; double categorization), participants pressed the "A" key if the sentence was of either the true or the guilty type and the "L" key if the sentence was of either the false or the innocent type. In Block 4 (40 trials; reversed autobiographical discrimination), participants pressed the "A" key for sentences of the innocent type and the "L" key for sentences of the guilty type. In Block 5 (60 trials; reversed double categorization), participants pressed the "A" key for true and innocent sentences and the "L" key for false and guilty sentences.

Reminder labels in the form of category names remained on the monitor for the entire duration of each block. An error signal appeared after an incorrect response. True-false sentences and guilty-innocent sentences were presented in alternation in Blocks 3 and 5. Half of the participants were administered the blocks in the order just outlined, whereas for the other half, the order of Blocks 3 and 5 was reversed (and the order of Blocks 2 and 4 was reversed accordingly). Preliminary analyses indicated that the order of presentation did not influence the main results and did not interact with the other factors. Therefore, the data were collapsed across the two block orders.

The comparison of interest was between average RT in Block 3 and average RT in Block 5. Both guilty and innocent respondents took part in Experiments 1, 2, and 5. We expected that innocent participants would be faster in the block that associated innocent with true sentences (congruent block), as compared with the block that associated guilty with true sentences (incongruent block), whereas guilty participants would exhibit the opposite pattern. That is, the specific pattern of facilitation was expected to depend on each individual's autobiographical knowledge. No innocent participants were included in Experiments 3, 4, and 6: In these experiments, we expected all participants to be faster in the block of trials that associated guilty with true sentences (congruent block), as compared with the block that associated guilty with false sentences (incongruent block).

DATA ANALYSIS

Two dependent measures were considered: mean RT in the double-categorization blocks and *D* (Greenwald, Nosek, & Banaji, 2003). RTs shorter than 150 ms or longer than 10,000 ms were discarded. Unless specified, data were submitted to an analysis of variance (ANOVA) with group (guilty vs. innocent) as a between-subjects factor and congruency (congruent vs. incongruent) as a within-subjects factor. The *D* index includes a penalty for incorrect responses and expresses the IAT effect (the difference in performance between the two double-categorization blocks) in terms of the standard deviation of the latency measures. We calculated *D* by subtracting corrected (see Greenwald et al., 2003, for the procedure) mean RT for the block associating guilty and true sentences from mean RT in the block associating innocent and true sentences and then dividing this difference by the inclusive standard deviation of the two blocks. Guilty participants were expected to have positive *D* values, whereas innocent participants were expected to have negative *D* values. We classified participants on the basis of *D* scores and determined the number who were classified correctly.

TABLE 2

Sentences Used in Experiments 1 Through 5

Experiment and category	English translation
All experiments	
True	I'm in the basement of the psychology department
	I'm in a little room with a computer
	I'm doing a psychology experiment
	I'm in the psychology laboratory
	I'm in front of the computer
False	I'm climbing a mountain
	I'm at the beach
	I'm eating in a downtown restaurant
	I'm playing football
	I'm in a shop
Experiment 1	
Guilty: 4 of diamonds	I picked card number 4
	I turned over card "four"
	I saw the 4 of diamonds
	I turned over the 4 of diamonds
	I have the 4 of diamonds
Innocent: 7 of clubs	I picked card number 7
	I turned over card "seven"
	I saw the 7 of clubs
	I turned over the 7 of clubs
	I have the 7 of clubs
Experiment 2	
Guilty: "I stole the CD"	I entered the professor's office
	I stole a CD with a copy of the exam
	I stole the exam for clinical neuropsychology
	I entered the office to steal the CD with the exam
	I stole the exam
Innocent: "I did not steal the CD"	I never entered the professor's office to steal the CD
	I have never stolen the CD containing the clinical neuropsychology exam
	I did not steal the exam
	I have never stolen the exam for clinical neuropsychology
	I did not steal the exam for clinical neuropsychology
Experiment 3	
Guilty: "I used cocaine"	I have tried cocaine once
	I took cocaine recently
	I was addicted to cocaine
	I used cocaine
	I was a cocaine abuser
Innocent: "I did not use cocaine"	I never tried cocaine
	I did not take cocaine
	I was never addicted to cocaine
	I never used cocaine
	I was not a cocaine abuser
Guilty: "I used heroin"	I have tried heroin once
	I took heroin recently
	I was addicted to heroin
	I used heroin
	I was a heroin abuser

Table 2. (Contd.)

Experiment and category	English translation
Innocent: "I did not use heroin"	I never tried heroin
	I did not take heroin
	I was never addicted to heroin
	I never used heroin
	I was not a heroin abuser
Experiment 4	
Guilty: "I went to Paris"	Last summer I went to Paris
	I saw the Eiffel Tower
	I visited the Louvre
	I saw the Mona Lisa
	I visited the Arc de Triomphe
Innocent: "I went to London"	Last summer I went to London
	I saw Big Ben
	I had a typical English breakfast
	I visited the Tate Modern Museum
	I visited the British Museum
Experiment 5	
Guilty: "My driving license was suspended because of drunk driving"	I drove after I drank, so my driving license was suspended
	I drove my car while drunk, and they suspended my driving license
	I drove while not sober, and they suspended my driving license
	They suspended my driving license because I was drunk and I was driving
	They suspended my driving license because I was above the alcohol level
Innocent: "My driving license was not suspended because of drunk driving"	My driving license was not suspended because I was drunk
	They did not suspend my driving license because of my alcohol level
	They never suspended my driving license because I was drunk
	My driving license was not suspended because I was above the alcohol level
	They never suspended my driving license because I was above the alcohol threshold

Note. In Experiment 4, the guilty and true statements were customized for each participant; guilty sentences referred to the participant's last vacation, and innocent sentences referred to a vacation the participant did not take. This table presents the sentences used for 1 of the participants.

We conducted a receiver-operating-characteristic (ROC) analysis (Swets, 1988) to determine how well the aIAT discriminated between differing groups. This analysis allowed us to compare the results obtained with the aIAT with those obtained using the GKT and fMRI (Ben-Shakhar & Elaad, 2003; Langleben et al., 2005).

EXPERIMENT 1: PLAYING CARDS

In Experiment 1, we evaluated the accuracy of the aIAT in identifying which of two playing cards participants had selected.

Participants

Thirty-seven students (8 males and 29 females; 19–30 years old, mean age = 23.4) volunteered for this study.

Procedure and Stimuli

In a preliminary consolidation task, participants selected one of two playing cards (the 4 of diamonds or 7 of clubs) and memorized it. Out of the 37 participants, 17 selected the 4 of diamonds, and 20 selected the 7 of clubs. Next, in each trial of the consolidation task, one of eight different playing cards (e.g., 4 of diamonds, 7 of clubs, 3 of hearts, 3 of diamonds) was presented in a central position on the monitor, and participants were asked to press the space bar if that card was the previously chosen card. Each card was presented 5 times, for a total of 40 trials. Error feedback was presented for 400 ms if participants responded incorrectly. After the consolidation task, participants performed the experimental task.

In the experimental task, we arbitrarily defined guilty sentences as sentences referring to the 4 of diamonds, and innocent sentences as sentences referring to the 7 of clubs. Table 2 lists the "4 of diamonds" and "7 of clubs" sentences that were used together with the true and false sentences. Block 3 and Block 5 each included a total of 60 trials (15 true sentences, 15 false sentences, 15 "4 of diamonds" sentences, and 15 "7 of clubs" sentences). Each sentence was displayed until the participant responded.

Results

Figure 1a represents the significant effect of congruency, $F(1, 35) = 37.275$, $p < .001$, $p_{rep} > .99$, $\eta^2 = .516$. Mean latencies were lower for the congruent than for the incongruent blocks

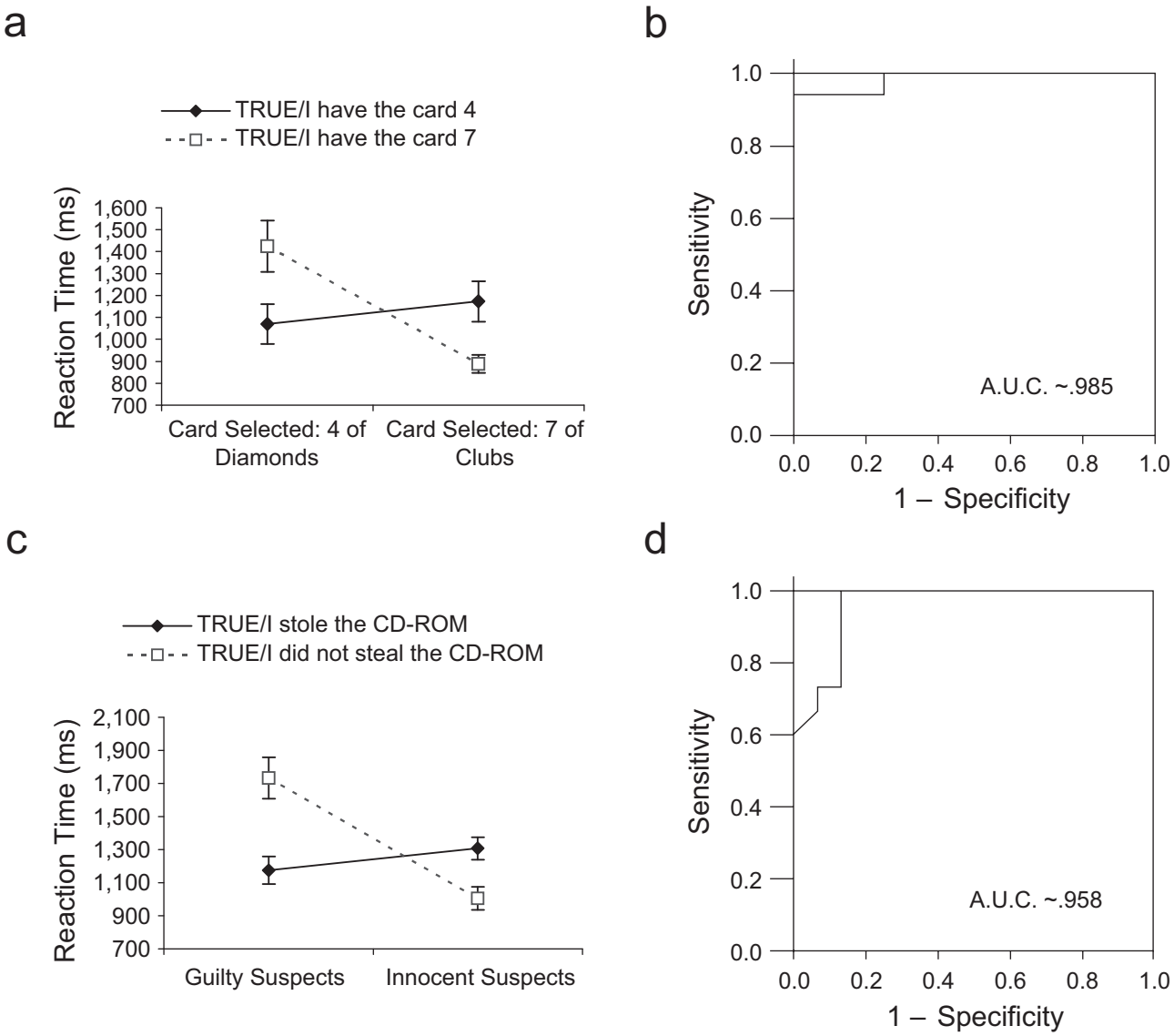

Fig. 1. Results for Experiments 1 and 2: graphical representations of the interaction between group and stimulus pairings and receiver-operating-characteristic (ROC) curves. In the line graphs, mean reaction time is plotted as a function of group. For Experiment 1 (a), results are shown for the critical block associating true sentences with ''4 of diamond'' sentences and for the critical block associating true sentences with ''7 of clubs'' sentences. For Experiment 2 (c), results are shown for the critical block associating true sentences with guilty sentences and for the critical block associating true sentences with innocent sentences. In the ROC curves (b: Experiment 1; d: Experiment 2), ''sensitivity'' refers to the percentage of guilty participants correctly classified as guilty, and ''1 – specificity'' refers to the percentage of innocent participants erroneously classified as guilty. A.U.C. = area under the curve (from the ROC analysis).

(972 vs. 1,288 ms), suggesting a strong association between sentences referring to the selected card and true statements. RTs did not differ between the two groups of participants, that is, those who chose the 4 of diamonds (guilty) and those who chose the 7 of clubs (innocent), $F(1, 35) = 3.696$, $p = .063$, $\eta^2 = .096$. The interaction between congruency and group was not significant, $F(1, 35) = 0.459$, $p = .502$, $\eta^2 = .013$.

To test the accuracy of the instrument, we computed the *D* index (based on the difference between performance in the block associating "7 of clubs" and true sentences and performance in the block associating "4 of diamonds" and true sentences). Higher values of the index pointed to the autobiographical knowledge of having picked the 4 of diamonds, whereas lower values pointed to autobiographical knowledge of the opposite event. The mean *D* index was positive for the group who selected the 4 of diamonds and negative for the group who selected the 7 of clubs (0.62 vs. −0.49). This difference was significant, $F(1, 35) = 82.753$, $p < .001$, $p_{rep} > .99$, $\eta^2 = .70$. The accuracy of the method was confirmed by the ROC analysis (area under the curve, AUC = .985; see Fig. 1b). Classification based on the aIAT was more accurate than classification based on the GKT (Ben-Shakhar & Elaad, 2003; AUC = .80) or fMRI

(Langleben et al., 2005; AUC = .80) for the same playing-card test. The *D* index calculated from aIAT performance accurately classified 35 of 37 participants.

EXPERIMENT 2: MOCK CRIME

In this experiment, participants in the guilty group simulated a theft, whereas those in the innocent group simply read a press report on the same event.

Participants

Thirty students volunteered for the experiment (14 males and 16 females; 23–30 years old, mean age = 25.3). They were randomly assigned to the guilty and innocent groups.

Procedure and Stimuli

Guilty suspects were instructed to enter the office of a teaching assistant and steal a CD containing an upcoming examination. Innocent suspects read a press report on this event. The aIAT procedure was the same as in Experiment 1, except that different guilty (e.g., "I stole the exam") and innocent (e.g., "I did not steal the exam") sentences were used. Table 2 lists all the guilty and innocent sentences used.

Results

Figure 1c shows that RT was faster for the congruent than for the incongruent condition (1,091 vs. 1,520 ms), $F(1, 28) = 43.328$, $p < .001$, $p_{rep} > .99$, $\eta^2 = .607$. RTs for guilty and innocent participants did not differ, $F(1, 28) = 7.523$, $p = .011$, $\eta^2 = .212$. The interaction between congruency and group was not significant, $F(1, 28) = 3.892$, $p = .058$, $\eta^2 = .122$.

Analysis of the *D* index revealed a significant difference between guilty and innocent participants (0.78 vs. −0.85), $F(1, 28) = 68.462$, $p < .001$, $p_{rep} > .99$, $\eta^2 = .710$. All guilty suspects showed a strong association between the guilty and the true sentences and therefore were correctly classified as guilty. Thirteen of the 15 innocent suspects showed a strong association between the innocent and the true sentences. The ROC analysis revealed an AUC of .96 (see Fig. 1d). The aIAT outperformed the GKT (AUC = .87; Ben-Shakhar & Elaad, 2003) in classifying participants.

EXPERIMENT 3: HEROIN AND COCAINE

In this experiment, the aIAT was applied within an ecological setting: the detection of illegal drug usage.

Participants

Fourteen participants (13 males and 1 female; 23–45 years old, mean age = 35.4) with at least 5 years of both heroin and cocaine abuse were tested at a local substance-abuse clinic. Half of the participants were administered a version of the aIAT that investigated their previous use of cocaine, whereas the other half were given a version of the test that investigated their previous use of heroin.

Procedure and Stimuli

The true and false sentences were the same as in Experiments 1 and 2. The guilty sentences referred to past heroin or cocaine usage, whereas the innocent sentences referred to nonusage of heroin and cocaine (see Table 2). In the congruent condition, statements referring to heroin or cocaine usage were paired with true sentences and statements referring to nonusage of heroin or cocaine were paired with false sentences; in the incongruent condition, nonusage of drugs was paired with true statements and usage of drugs was paired with false statements.

Results

An ANOVA with aIAT type (heroin vs. cocaine) as a between-subjects factor and congruency (congruent vs. incongruent) as a within-subjects factor revealed no RT difference between participants responding to the heroin aIAT and those responding to the cocaine-aIAT, $F(1, 12) = 0.205$, $p = .659$, $\eta^2 = .017$. The only significant effect indicated that responses to congruent associations were faster than responses to incongruent associations (1,601 vs. 2,234 ms), $F(1, 12) = 24.389$, $p < .001$, $p_{rep} > .99$, $\eta^2 = .670$. No other effects approached significance. Thirteen of the 14 drug users had a positive *D*. The average *D* was 0.98 for the participants who took the heroin aIAT and 0.40 for participants who took the cocaine aIAT.

EXPERIMENT 4: AUTOBIOGRAPHICAL MEMORY

It might be argued that the sentences used in Experiment 3 did not tap into autobiographical memories, but rather described participants' characteristics. In Experiment 4, to ascertain the accuracy of the aIAT in detecting single autobiographical events limited in time and space, we asked participants to report a personal experience.

Participants

Twenty participants (8 males and 12 females; 19–53 years old, mean age = 32.2) volunteered for this experiment.

Procedure and Stimuli

Our goal was to determine whether the aIAT could correctly identify the actual last vacation (guilty sentences) taken by examinees. Therefore, the critical comparison was between a block in which the last vacation taken was paired with true sentences and a fabricated last vacation (innocent sentences) was paired with false sentences and a block in which the last vacation taken was paired with false sentences and the fabricated last vacation was paired with true sentences. Participants were requested to fill out a questionnaire describing their last vacation and an imagined vacation they never took. Then, for

each participant, we created a personalized aIAT with guilty sentences describing the true vacation and innocent sentences describing the vacation the participant never took (Table 2 lists the sentences for a single participant as an example).

Results

A repeated measures ANOVA with congruence (congruent vs. incongruent) as a within-subjects factor was conducted. Mean RT was faster for the congruent block than for the incongruent block (1,041 vs. 1,260 ms), $F(1, 19) = 40.101$, $p < .001$, $p_{rep} > .99$, $\eta^2 = .679$. For 18 of the 20 participants, we correctly identified the real event on the basis of the double-categorization block in which they were fastest. The average D score was 0.44.

EXPERIMENT 5: SUSPENSION OF DRIVING LICENSE FOR DRUNK DRIVING

A possible problem in the previous experiments is that participants were not exposed to the high level of stress typical of an investigative setting and would not have experienced direct advantages from faking. An important challenge for experimental studies of deception is to use a valid setting comparable to real situations in which subjects may lie or conceal spontaneously. Therefore, we decided to run an experiment in which participants were highly motivated to pass the test. All participants in the guilty group had had their driving license suspended for driving with an excessive blood alcohol level.

Participants

Fifty participants (44 males and 6 females; 18–73 years old, mean age = 35.72) took part in the experiment. The guilty group consisted of 25 participants who had had their driving license suspended for driving while they had a blood alcohol level greater than 0.5 mg/ml (as determined by testing at a hospital after the driver was stopped by a police officer). The innocent group consisted of 25 participants who were matched to the guilty group for age, sex, and education level; their driving-license records indicated no incidents of driving with an excessive blood alcohol level.

Procedure and Stimuli

For the guilty group, the aIAT was included as part of the compulsory medical and psychological assessment required for reinstatement of a driving license. Guilty participants were made to believe that the aIAT outcome would determine whether or not their driving license was reinstated. The true and false sentences were the same as in the four previous experiments. The five guilty sentences described the illegal act, and the five innocent sentences stated that the respondent's license had not been suspended for drunk driving (see Table 2). The guilty participants were expected to show an association between true and guilty sentences (and between false and innocent sentences), whereas the innocent participants were expected to show the reverse pattern.

Results

As Figure 2a shows, both groups (guilty and innocent) had faster RTs in the congruent block than in the incongruent block (1,805 vs. 2,250 ms), $F(1, 48) = 32.029$, $p < .001$, $p_{rep} > .99$, $\eta^2 = .400$. No other effect approached statistical significance.

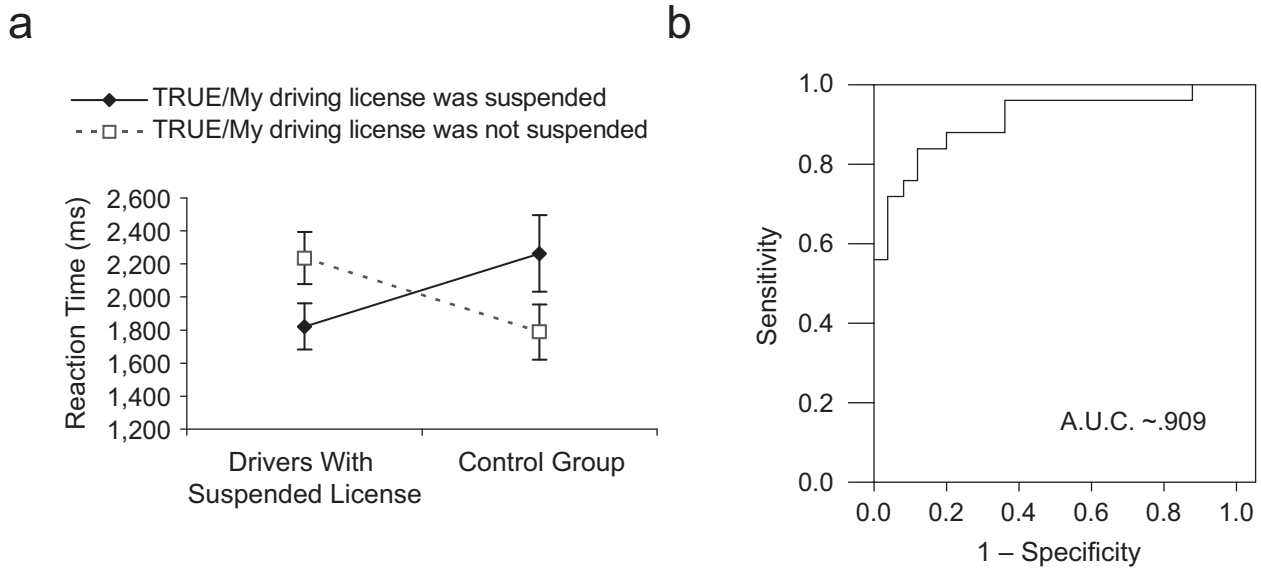

Fig. 2. Results for Experiment 5: (a) graphical representation of the interaction between group and stimulus pairings and (b) receiver-operating-characteristic (ROC) curves. In (a), mean reaction time is plotted as a function of group. Results are shown for the critical block associating guilty sentences with true sentences and for the critical block associating innocent sentences with true sentences. In the ROC curve, "sensitivity" refers to the percentage of guilty participants correctly classified as guilty, and "1 – specificity" refers to the percentage of innocent participants erroneously classified as guilty. A.U.C. = area under the curve (from the ROC analysis).

Analysis of the D index revealed that the difference between guilty and innocent participants was significant, $F(1, 49) = 44.719$, $p < .001$, $p_{rep} > .99$, $\eta^2 = .482$. The average D was positive for the guilty group, but negative for the innocent group (0.39 vs. −0.44). Using D, we correctly classified 44 of the 50 participants (22/25 for the guilty group and 22/25 for the innocent group). Finally, the ROC analysis yielded an AUC of .91 (see Fig. 2b).

EXPERIMENT 6: CRIMINALS

We administered the aIAT to 2 individuals who were found guilty after confessing their crimes and who were classified as insane on the basis of a forensic psychiatric assessment. Both were under medication and were examined in a forensic mental hospital. The first examinee (D.E.) had attempted to kill his two sons. The second examinee (C.S.) was found guilty of killing his mother. For each criminal, we created a personalized aIAT with five guilty sentences describing the crime (e.g., "I attempted to kill my children" or "I killed my mother") and five innocent sentences denying the crime (e.g., "I did not attempt to kill my children" or "I did not kill my mother"). In each case, the congruent block associated guilty and true sentences, and the incongruent block associated innocent and true sentences.

D.E. responded more quickly in the congruent block (4,296 ms[1]) than in the incongruent block (6,733 ms), $t(119) = -3.336$, $p < .001$, $p_{rep} > .99$. This pattern indicates a strong association between the guilty episode and the attribute "true" ($D = 1.0$). Similarly, C.S.'s average RT for the congruent block (1,019 ms) was significantly faster than his average RT for the incongruent block (2,213 ms), $t(119) = -9.611$, $p < .001$, $p_{rep} > .99$. This pattern reveals a strong association between the murder and true sentences ($D = 0.61$).

DISCUSSION

We have reported a novel method that reliably detects concealed autobiographical knowledge and could be used in forensic science. It is important to note that the aIAT uses sentences as stimuli, rather than single words or pictures (as is characteristic of the original IAT). It allows one to investigate autobiographical memory, rather than semantic memory. The results from the experiments reported here provide compelling evidence of the high level of accuracy with which concealed autobiographical knowledge can be detected using this instrument. The accuracy of the aIAT is evident not only at the group level, but also at the individual level. On average, and using a variety of tasks, we were able to classify 91% of the participants correctly. The aIAT also is flexible in that the examinee's knowledge of virtually any type of factual information can be assessed using a verbal format. Like the GKT (Lykken, 1998), the aIAT could be used as a lie-detection technique.

A relevant issue is whether aIAT performance can be faked. Experiment 5 provides persuasive evidence that, even in a naturalistic setting in which respondents would be extremely prone to faking, the aIAT is still able to detect autobiographical events accurately.

In the experiments reported here, participants complied with the instructions they were given. Such compliance is typical of innocent suspects taking a "lie detection" test. However, there are also situations in which guilty suspects agree to take a test that may prove their guilt. Naturally, they are highly motivated to fake the results. Some researchers have reported that IAT measures may be faked by participants who have been instructed to slow down on congruent trials and speed up on incongruent trials (e.g., Fiedler & Bluemke, 2005; Kim, 2003; Steffens, 2004). Whether indirect indices (algorithms) could be developed to detect such countermeasures is an open issue for future research.

Issues concerned with faking are particularly evident in the case of psychophysiological and neuroimaging techniques. First, effective countermeasures to psychophysiological assessment are easy to implement. A polygraph may be faked if the guilty suspect is trained in the use of physical (e.g., biting the tongue or pressing the toes to the floor) and mental (e.g., engaging in mental activities that require effort, such as counting backward) countermeasures (Ben-Shakar & Elaad, 2003; Honts, Raskin, & Kircher, 1994). Second, although the use of fMRI-based techniques has revealed that activity within the frontal lobe is sensitive to the production and complexity of lies (e.g., Ganis et al., 2003), two main problems cast doubt on the validity of this costly and cumbersome technique: The results could perhaps be faked by intentional head movements, which might prevent an exact anatomical localization, and guilty suspects could activate the "deception" frontal network during control questions by covertly engaging in a concurrent cognitive task, so that lies would be confused with truthful responses (such as counting backward; e.g., Cole & Schneider, 2007).

Both the aIAT and the TARA (Gregg, 2007) use response incongruity to identify lies. However, our method differs from the TARA in three important ways. First, the TARA uses only two categories ("true" and "false") instead of the four ("true" and "false"; "guilty" and "innocent") used by the aIAT. Second, in the application phase, the TARA uses only one critical block instead of two critical blocks, one congruent and one incongruent (as in the typical IAT; Greenwald et al., 1998). Third, the TARA discriminates truth from lie on the basis of the absolute level of RT in the critical block: If the average RT is fast, then the respondent is honest; otherwise, the respondent is lying. This procedure therefore requires a comparison with appropriate cutoffs obtained from carefully matched control groups. This requirement may highlight a practical limit of the TARA. Consider the results obtained for D.E. in Experiment 6. This crim-

[1]D.E.'s very slow RTs were presumably due to his medication for neurolepsy. It is unlikely that they were due to an attempt to fake his performance, as he was similarly slow in performing another RT task (the stop-signal task).

inal is a medicated patient with very slow RTs. Testing him with the TARA would require a medicated, age-matched control group. Using a nonmedicated control group with normal RTs would cause D.E. to be misclassified as a liar even if he responds truthfully.

To conclude, the aIAT is an accurate method of detecting concealed knowledge that outperforms currently available lie-detection techniques. It can be used to assess the existence of virtually any kind of autobiographical memory; for example, it can be used to identify malingering in a range of psychiatric and neurological disorders (e.g., depression or whiplash; Sartori, Agosta, & Gnoato, 2007). This method has the potential to provide novel insights in detecting lies and malingering in forensic settings, although (like other techniques) it leaves important neuroethical issues unresolved (Wolpe, Foster, & Langleben, 2005).

REFERENCES

American Academy Symposium. (2007, Spring). Is there science underlying truth detection? *Bulletin of the American Academy of Arts & Sciences*, *LX*, 6–7.

Ben-Shakhar, G., & Elaad, E. (2003). The validity of psychophysiological detection of information with the Guilty Knowledge Test: A meta-analytic review. *Journal of Applied Psychology*, *88*, 131–151.

Cole, M.W., & Schneider, W. (2007). The cognitive control network: Integrated cortical regions with dissociable functions. *NeuroImage*, *37*, 343–360.

Crovitz, H.F., & Schiffman, H. (1974). Frequency of episodic memories as a function of their age. *Bulletin of the Psychonomic Society*, *4*, 517–518.

Fiedler, K., & Bluemke, M. (2005). Faking the IAT: Aided and unaided response control on the Implicit Association Test. *Basic and Applied Social Psychology*, *27*, 307–316.

Ganis, G., Kosslyn, S.M., Stose, S., Thompson, W.L., & Yurgelun-Todd, D.A. (2003). Neural correlates of different types of deception. *Cerebral Cortex*, *13*, 830–836.

Gray, N.S., Brown, A.S., MacCulloch, M.J., Smith, J., & Snowden, R.J. (2005). An implicit test of the associations between children and sex in pedophiles. *Journal of Abnormal Psychology*, *114*, 304–308.

Gray, N.S., MacCulloch, M.J., Smith, J., Morris, M., & Snowden, R.J. (2003). Violence viewed by psychopathic murderers. *Nature*, *423*, 497–498.

Greenwald, A.G., McGhee, D.E., & Schwarz, J.L.K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*, 1464–1480.

Greenwald, A.G., Nosek, B.A., & Banaji, M.R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*, 197–216.

Gregg, A.I. (2007). When vying reveals lying: The Timed Antagonistic Response Alethiometer. *Applied Cognitive Psychology*, *21*, 621–647.

Honts, C.R., Raskin, D.C., & Kircher, J.C. (1994). Mental and physical countermeasures reduce the accuracy of polygraph tests. *Journal of Applied Psychology*, *79*, 252–259.

Iacono, W.G., & Lykken, D.T. (1999). Update: The scientific status of research on polygraph techniques: The case against polygraph tests. In D.L. Faigman, D.H. Kaye, M.J. Saks, & J. Sanders (Eds.), *Modern scientific evidence: The law and science of expert testimony* (pp. 174–184). St. Paul, MN: West Publishing.

Kim, D.Y. (2003). Voluntary controllability of the Implicit Association Test. *Social Psychology Quarterly*, *66*, 83–96.

Langleben, D.D., Loughead, J.W., Bilker, W.B., Ruparel, K., Childress, A.R., Busch, S.I., & Gur, R.C. (2005). Telling truth from lie in individual subjects with fast event-related fMRI. *Human Brain Mapping*, *26*, 262–272.

Lykken, D.T. (1960). The validity of the guilty knowledge technique: The effects of faking. *Journal of Applied Psychology*, *44*, 258–262.

Lykken, D.T. (1998). *A tremor in the blood: The uses and abuses of the lie detector*. Reading, MA: Perseus Books.

Moore, H.M., Petrie, C.V., & Braga, A.A. (2003). *The polygraph and lie detection*. Washington, DC: National Academies Press.

Sartori, G., Agosta, S., & Gnoato, F. (2007, October). *High accuracy detection of malingered whiplash syndrome*. Paper presented at the International Whiplash Trauma Congress, Miami, FL.

Steffens, M.C. (2004). Is the Implicit Association Test immune to faking? *Journal of Experimental Psychology*, *51*, 165–179.

Swets, J.A. (1988). Measuring the accuracy of diagnostic systems. *Science*, *240*, 1285–1293.

Wolpe, P.R., Foster, K.R., & Langleben, D.D. (2005). Emerging neurotechnologies for lie-detection: Promises and perils. *The American Journal of Bioethics*, *5*(2), 39–49.

(Received 7/16/07; Revision accepted 12/16/07)