The world's first ocular-motor deception test.

Converus
August 2019

Converus Inc.
610 S. 850 E., Suite 4
Lehi, UT 84043 USA
+1-801-331-8840
www.converus.com

Introduction

Most theories of deception detection hypothesize that lying is more cognitively demanding than telling the truth.[1] Deceptive individuals use cognitive resources to inhibit the truth, fabricate the lie, and maintain its consistency, coherence, and believability over time. Deceptive individuals may surveil their own behavior and internal state of arousal to monitor whether they are leaking incriminating information, especially during an interrogation or examination[2]. During interrogation, they may also use cognitive resources to observe the behavior of any interviewers for feedback on their own perceived believability. Inhibiting truthful responses, attempting to maintain credibility over time, monitoring the interviewer and self-monitoring for signs of information leakage are cognitive processes that require mental effort. Therefore, it is more difficult to lie than to tell the truth, and lying causes increases in mental effort (cognitive load), which are measurable.

Psychologists have long known there is a correlation between increased cognitive load and certain eye behaviors. For example, pupils dilate commensurate with cognitive workload increases.[3] For instance, an individual's pupils will dilate slightly when attempting to mentally multiply 17 x 2. By contrast, that individual's pupil dilation will be more pronounced when attempting to mentally multiply 17 x 31. In the same way, an individual's pupils dilate slightly when responding truthfully to questions; but when responding deceptively to questioning, the individual's pupils will show greater dilation because of the cognitive load increase required to fabricate a response. These are two simple examples of how an increase in mental exertion will temporarily increase pupil diameter.

Other important ocular-motor indicators of cognitive processes include:

1) Guilty individuals blink less often as they process statements answered deceptively versus those answered truthfully.
2) Guilty participants respond faster, make fewer fixations, and spend less time reading and re-reading statements about their own inappropriate behaviors than when responding to statements about neutral topics or inappropriate behaviors in which they do not engage.
3) Guilty participates show an increase in cognitive load associated with recalling a task and when distinguishing between deceptive and non-deceptive responses. This is more pronounced when responding to complex statements.

Polygraph and Other Lie Detectors

For decades, the de facto standard in lie detection technology has been the polygraph. Invented in the 1920s, polygraph has been the only credibility assessment tool to show accuracy rates of up to about 90% when used in specific event questioning and 81% in screening tests.[4] The traditional polygraph approach to detecting deception for criminal investigations and general screening applications is to base decision on within-subject comparisons of physiological responses to different types of test questions. Polygraph sensors record breathing movement using transducers wrapped around the chest and abdomen, electro dermal activity from electrodes attached to the tips of two fingers, and relative blood pressure changes via a blood pressure cuff on the upper or lower arm.

In polygraph, various theoretical constructs have been proposed to explain the differential physiological responses of truthful and deceptive people to the different types of test questions. Physiological responses are loaded on one type of question or another as a function of truthfulness or deception. The putative physiological underpinnings of the responses include attention, conflict, conditioned response, and fear of detection. No single explanation is sufficient to account for all effects. Despite the lack of consensus about specific mechanisms producing physiological response, it is clear that emotional processes play an important role in polygraph testing, especially in the field where there may be serious consequences to the individual if he or she fails the test.

In recent years, several new cognition-based tests for deception have been developed, all of which are generally based on the notion that lying is cognitively more demanding than telling the truth. In theory, tests based on the

[1] Johnson, Barnhardt, & Zhu, 2005; Kircher, 1981; Vrij, Fisher, Mann, & Leal, 2000.
[2] Kircher, 1981
[3] Kahneman & Beatty, 1966
[4] Meta-Analytic Survey of Criterion Accuracy of Validated Polygraph Techniques, 2012, table 2.

concept of mental workload may be predominantly cognitive. However, these tests likely include an emotional component, just as polygraph techniques include a cognitive component.

Ocular-motor Deception Test

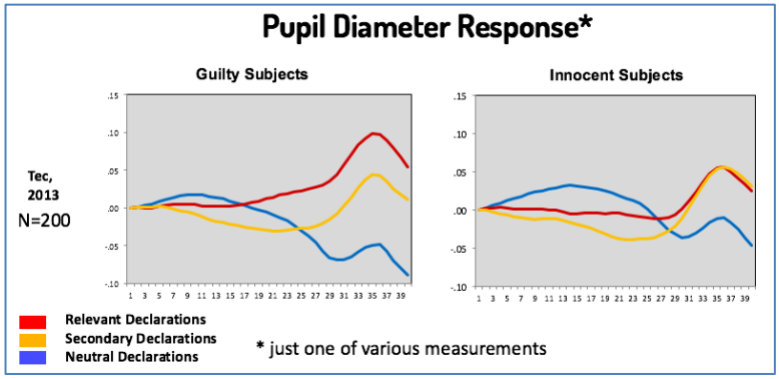
The concept of measuring deception based on ocular-motor (eye) behavior was first conceived in 2003 by two psychologists from the University of Utah, Dr. John Kircher, a psychophysiologist and pioneer in deception detection, and his colleague Dr. Doug Hacker, an educational psychologist with expertise in the psychology of reading. In addition to Kircher and Hacker, three others joined the science team including cognitive scientists Dr. Dan Woltz and Dr. Ann Cooke, as well as well-known polygraph expert Dr. David Raskin. This group researched and tested the concept of an ocular-motor deception test (ODT) for more than a decade.

Kircher, the lead scientist, was co-inventor with Raskin of the computerized polygraph in 1991. Both are recognized experts in deception detection. Kircher has published more than 50 scientific articles and reports related to credibility assessment. He has consulted with, and conducted research on, deception detection for the U.S. Department of Defense, National Science Foundation, CIA, U.S. Secret Service, National Institute of Justice, Department of Homeland Security, National Science Foundation, National Research Council, Royal Canadian Mounted Police, and many police departments.

Initially, the science team looked at existing theories that discussed the affects of deception on eye and reading behavior. Then they devised experiments to evaluate a variety of ocular motor responses in cognitive-based exercises to determine if a correlation existed. After nine years of research, their findings were published in 2012 in a peer-reviewed article entitled, "Lyin' Eyes: Ocular-motor Measures of Reading Reveal Deception."[5] In this seminal study, subjects were randomly assigned to either a "Guilty" group that committed one of two mock crimes or they were assigned to an "Innocent" group that only learned about the crime after-the-fact. Subjects completed a computer-administered questionnaire that used true/false statements to address their possible involvement in the mock crimes. Subsequent experiments also manipulated the participants' incentive to pass the test and changed the difficulty of the true/false statements on the test.

In these experiments, Guilty participants had increased pupillary responses to true/false statements answered deceptively. In addition, Guilty participants spent less time fixating on, reading, and rereading true/false statements than participants responding truthfully. These various ocular motor behaviors were measured and were optimally weighted in a logistic regression analysis. By definition, a logistic regression analysis combines one or more variables in a data set to predict a binary outcome, such as truthful or deceptive.

Findings from these and subsequent studies indicated that discrimination between guilt and innocence was improved when offering greater incentives to pass the test and when using statements with simple syntax. These findings suggested that two cognitive processes are involved in deception: 1) vigilance and 2) strategy; and these processes are reflected in various ocular motor and behavioral measures.



[5] "Lyin' Eyes: Ocular-motor Measure of Reading Reveal Deception," Journal of Experimental Psychology: Applied, 18(3), 301-313. September 2012.

The science team's efforts produced a cognition-based test that uses ocular-motor measures of cognitive effort, including some based on reading behavior. In general, when a person experiences difficulty in reading a word or phrase, their eyes behave in specific ways and eye responses correlate with performance on a wide variety of cognitive tasks. Although early research on eye behavior suggested that emotional factors are relatively unimportant in determining eye responses, there is an association between the eye response and emotional arousal.[6]

Lab and Field Study Results

In 2016, Kircher conducted field studies with the support and assistance of three groups in the Mexican federal government and published new data that showed the mean accuracy of ODT to be 86% for screening tests using the Relevant-Comparison Test (RCT) protocol. The data was published in in December 2016 in "Laboratory and Field Research on the Ocular-motor Deception Test" in the European Polygraph Journal.

In late 2018, Kircher and Raskin reviewed field data on a small sample of tests using the Directed Lie protocol, which is primarily used for diagnostic or single-issue testing. That data showed the mean accuracy of that protocol to be over 90%. More data will be gathered prior to any publication.

In June 2019, Converus announced the availability of its third testing protocol: Multi-Issue Comparison Test (MCT). The MCT protocol was developed to allow screening tests with 3 relevant issues and 1 comparison issue, where all issues are scored independently. This new protocol is the focus of the dissertation work of Andrew Potts, candidate for Ph.D. at the University of Utah. Andrew is working under the direction of Dr. John Kircher. In his initial lab studies, the overall mean accuracy of the MCT is 88%.

EyeDetect

After the science team reached its conclusions and published the peer-reviewed study in 2012, skilled programmers followed the lab-based processes and developed computer-based software to administer tests in an automated, standardized and objective process. They also developed a web-based dashboard and administration tool to summarize test results. With this software in place, credibility assessment testing precision is increased as potential human error is minimized. There is no examiner bias, nor can the examiner affect the outcome of the test.

These software programs were made commercially available in April 2014 under the brand name EyeDetect® — the world's first ocular-motor deception test (ODT). It is the only screening test to have a mean accuracy as high as 85% in determining deceptive and truthful examinees.

Test Format

In polygraph, there are a number of testing techniques that are considered validated. Some examples include AFMGQT, CIT, DLST, Federal ZCT, and the Utah. With EyeDetect, there are three techniques in use today, and more are being developed.

Relevant Comparison Test

The first to be developed was the Relevant Comparison Test (RCT). Coincidentally, it was originally developed for an automated polygraph screening system by Kircher, Raskin, Gardner, Jewell, and Patnaik in 2002. It was designed primarily as a screening test at border entry points. In the European Polygraph Journal of December 2016, Dr. Kircher and Dr. Raskin published the mean accuracy at 86%. (TN .89 and TP .83)

[6] Bradley, Miccoli, Escrig & Lang, Psychophysiology, 2008 July; 45(4): 602-607.

Essentially, in an RCT, two to four issues are presented to the examinee. Each test includes an issue of primary concern (Relevant Question or RQ) of a variety of target behaviors, such as sex crimes, criminal history, drug use, stealing, association with known criminals, divulging confidential information, drug trafficking, falsifying a police application, etc.

When considering the target behaviors to address with the test, it is important to be as specific as possible to eliminate any uncertainty for the examinee.

The following is a list of common relevant issues. For each issue, additional clarifying information should be provided in the pre-test instructions to ensure the examinee understands the issue.

- ***Stealing***
  - From previous employers or from the current employer
  - Examples: money, products, equipment, raw materials, etc.
  - Value of items: any value, $100, $500, etc.
- ***Drug use***
  - Examples: marijuana, cocaine, heroin, amphetamines, steroids, etc.
  - Time frame: 90 days, 12 months, 24 months, as an adult, etc.
- ***Serious crimes***
  - Crimes against a person or property
  - Committed as an adult
  - Whether caught or not caught
  - Examples: burglary, robbery, drug manufacturing, domestic violence, etc.
- ***Criminal ties***
  - Examples: cartels, gangs, organized crime, other delinquents
  - Type of affiliation: support, work with, receive benefits from, etc.
- ***Divulging confidential info***
  - To unauthorized persons
  - Examples: confidential, classified, secret, top secret, etc.
- ***Bribes***:
  - Accepting or asking for bribes
  - Examples: money, gifts, favors, vacations, etc.

In addition to the relevant question, the RCT includes a secondary relevant issue (called the Comparison Question or CQ). This issue must meet the following criteria:

- Must be a crime more serious than the relevant issue
- No cross over with the primary relevant issue (should not be a related topic)
- Must have face validity for the examinee; the examinee must believe the issue is important
- Expected prior probability of guilt of 1-3%

Examples of CQ topics include terrorism, violent crime, identity theft, counterfeiting, arms trafficking, and others. The CQ is used to measure differences in the examinee's reaction to the RQ. As such, an ipsative, or within-subject measurement is derived to quantify the examinee's comparative reactions between the relevant and comparison issues.

During an RCT, the examinee responds to a series of true (T) and false (F) statements regarding the issues, as well as neutral (irrelevant) and alpha-arithmetic statements. A high-precision eye tracker measures and records eye movements and reading behavior. The eye tracker takes up to 60 measurements per second and records the data while software records aspects of the subject's responses to all T/F statements.

At the conclusion of the test, the ocular-motor measures and test question responses are combined by means of a logistic regression equation to compute the probability of deception to the RQ and CQ questions. There are more than 1 million data points gathered during the RCT.

The test format includes a set of pre-test instructions of the topics using an audio-visual presentation and may include mind maps. Mind maps are graphical representations of the test topics.

The pre-test instructions are followed by two short practice sessions that are used to familiarize the examinee with the testing process. For better test results, it is important the examinee become familiar with the testing process during a practice test rather than during the "real" test.

After the pre-test instructions and practice test, the test is administered for about 22 minutes.

During the test, the examinee must read T/F statements on-screen and respond using a computer keyboard or mouse. There are over 300 questions given during the test. Each question is allocated a specific response time.

Questions are asked every 6 to 9 seconds. To extract the best reaction in the event the examinee is deceptive, examinees must respond quickly. Examinees that intentionally delay in responding, those that respond randomly, or those that attempt to use countermeasures are considered non-cooperative and will be given a Not Credible score in the test.

During the RCT, the examinee must confirm or deny participation in the disqualifying behavior(s) no less than 80 times each. The examinee will also respond to 80 irrelevant questions (general knowledge questions) and 48 alpha-arithmetic questions.

These highly repetitive responses, carried out in a minimal time frame, generate a sufficient amount of data to use for comparison, analysis, and score calculation.

At the conclusion of the test, a Converus Credibility Score is calculated in less than 5 minutes. Credible scores range from 50 to 99 and deceptive scores range from 1 to 49. The closer to 99, the higher the probability of correctly classifying a subject as credible. The closer to 1, the higher the probability of correctly classifying a subject as deceptive.

Also, when test data is scored, a summary report is generated and saved in PDF or HTML format. Test results and scores are available from any web browser that has two-level encryption for access. Each test scored requires a test license.

Directed Lie Test

EyeDetect also supports the Directed Lie Comparison (DLC) test protocol, which is a technique originally developed for polygraph. The mean accuracy is 90%, according to field data gathered and analyzed by Kircher and Raskin in December 2018.

In the DLC test, one relevant issue is presented to the examinee. The RQ could cover issues of primary concern such as specific sex crimes, robbery, theft, drug use, compliance with probation rules, etc. The relevant question pertains directly to the matter under investigation for which the examinee is being tested. The relevant question can also address a screening concern.

In the case of the DLC test, the CQ consists of a series of directed lie questions. These are questions about transgressions that most everyone will readily admit. During the test, the examinee is **directed** to **lie** to this question. In the pre-test instructions, they are referred to as questions for which the examinee is "**required to lie**."

As such, a within-subject measurement is derived to quantify the examinee's comparative reactions to the relevant versus comparison issue, just as the RCT.

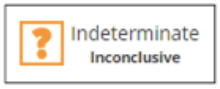
During the DLC test, the examinee will respond to a series of T/F statements regarding the RQ, CQ (directed lie), and some arithmetic statements.

The same eye tracker measures and records eye and reading behavior. At the conclusion of the test, the ocular-motor measures and test question responses are combined by means of a logistic regression equation to compute the probability of deception to the relevant issue.

The test format includes a pre-test explanation of the topics. That preamble is followed by a practice session that is used to familiarize the examinee with the testing process. Then, the test is administered for about 15 minutes. During the test, the examinee must read the T/F statements on-screen and respond using a mouse, computer keyboard or other device. There are 30 versions of the RQ and 30 of the CQ (directed lie) presented. Each question is allocated a specific response time. Examinees must respond quickly — this is to extract the best reaction in the event the examinee is deceptive.

Examinees that intentionally delay in responding, those that respond randomly, or those that attempt to use countermeasures are considered non-cooperative and will be considered having not passed the test.

During the DLC test, the examinee will be asked to confirm or deny participation in the relevant issue 30 times. At the conclusion of the test, a Converus Credibility Score is calculated in less than 5 minutes. In the case of the DLC, an additional test result has been introduced: *inconclusive*. Where the probability of credible or deceptive response is high, a score will be indicated. Where examinee responses are too weak to reliably predict, the outcome of "Inconclusive" will be given.



Credible scores are shown in the range of 60 to 99 and deceptive scores are shown as 1 to 40. The closer to 99, the higher the probability of correctly classifying the subject as credible. The closer to 1, the higher the probability of correctly classifying the person as deceptive.

Also, a summary report is generated and saved in PDF or HTML format. Test results and scores are available from any web browser that has two-level encryption for access. Each test requires a test license.

Our observation is that during diagnostic and screening tests with the DLC test, examinees normally obtain scores on the ends of the spectrum. Most guilty examinees score closer to "1" and most innocent examines score closer to "99". If eye behaviors and other measures are atypical, *Converus would rather indicate "inconclusive" to warrant a post-test interview or follow-on polygraph exam*.

Multi-Issue Comparison Test

This new testing protocol was announced in June 2019. The MCT protocol supports three RQ and 1 CQ in a general screening test wherein each issue is treated independently of the others and is scored separately.

For example, the RQ could address the examinee's participation in target behaviors or activities such as (a) illegal drug use, (b) acts of crime, (c) unreported work-related discipline, as well as a CQ with a lower prior probability of guilt such as (d) terrorism.

Unlike the current RCT, the MCT will question the examinee about each issue separately.

When the test is scored, there will be an overall score, whether Credible or Deceptive, and information and a score will be provided for each issue to indicate how the examinee reacted. (See image.)

This innovative new protocol will allow organizations to screen examinees for up to four target behaviors. The test will take approximately 28 minutes.

In the MCT, up to three RQ and one CQ are presented to the examinee. Any type of target behaviors, such as sexual assault, criminal history, drug use, stealing, association with known criminals, weapons trafficking, falsifying a police application, etc.

When considering the target behaviors to address with the test, it is important to be as specific as possible to eliminate any uncertainty for the examinee.

In addition to the RQ, the MCT includes a secondary relevant issue (CQ). Like the RCT, this issue must meet the following criteria:

- Must be a crime more serious than the relevant issue
- No cross over with the primary relevant issue (should not be a related topic)
- Must have face validity for the examinee; the examinee must believe the issue is important
- Expected prior probability of guilt of 1-3%

During an MCT, the examinee responds to a series of true (T) and false (F) statements regarding the issues, as well simple math statements. A high-precision eye tracker measures and records eye movements and reading behavior. The eye tracker takes up to 60 measurements per second and records the data while software records aspects of the subject's responses to all T/F statements.

At the conclusion of the test, the ocular-motor measures and test question responses are combined by means of a logistic regression equation to compute the probability of deception to the RQ and CQ questions.

The MCT test includes a set of pre-test instructions of the topics using an audio-visual presentation and may include mind maps.

The pre-test instructions are followed by two short practice sessions that are used to familiarize the examinee with the testing process. For better test results, it is important the examinee become familiar with the testing process during a practice test rather than during the "real" test.

After the pre-test instructions and practice test, the test is administered for about 10 minutes.

During the test, the examinee must read T/F statements on-screen and respond using a computer keyboard or mouse. Each question is allocated a specific response time.

Questions are asked every 6 to 9 seconds. To extract the best reaction in the event the examinee is deceptive, examinees must respond quickly. Examinees that intentionally delay in responding, those that respond randomly, or those that attempt to use countermeasures are considered non-cooperative and will be given a Not Credible score in the test.

At the conclusion of the test, a Converus Credibility Score is calculated in less than 5 minutes. Credible scores range from 50 to 99 and deceptive scores range from 1 to 49. The closer to 99, the higher the probability of correctly classifying a subject as credible. The closer to 1, the higher the probability of correctly classifying a subject as deceptive.

Also, when test data is scored, a summary report is generated and saved in PDF or HTML format. Test results and scores are available from any web browser that has two-level encryption for access. Each test scored requires a test license.

Comparison of Protocols

The chart below shows a simple comparison of the RCT, DLC and MCT protocols. The RCT and MCT protocols are used as screening tests. Each is similar in length of time required for the test, as well as comparable accuracy. The RCT allows for one RQ and the MCT allows for 3. Both tests use a CQ. The DLC is a diagnostic testing tool and focuses on specific issues.

	Use	Time	Accuracy	Topics
RCT	screening	30 min	86%	1R, 1C
MCT	screening	28 min	89%	3R, 1C
DLC	diagnostic	15 min	90%	1R, DL

More on the Algorithm

The science behind the EyeDetect credibility score is a logistic regression equation, which is, as mentioned, a statistical method for analyzing a data set with one or more independent variables.

$$Pr(Deceptive) = 1 / (1 + \exp (b0 + b1X1 + b2X2 + \ldots + bkXk))$$

where X is an ocular-motor characteristic (variable)
where b is an optimal weight

Decision Rule
If Pr(Deceptive) < .50, then "deceptive"
If Pr(Deceptive) $\geq$ .50, then "truthful"

The resulting calculation yields a binary outcome. In the case of EyeDetect, the two possible outcomes are 1) truthful or 2) deceptive. In the case of EyeDetect, the possible outcome is either "deception" or "truthfulness." The definition of the variables and the use of a logistic regression equation were devised by the science team after years of research and lab testing.

Some of the independent variables considered by the EyeDetect algorithm include pupil dilation, response accuracy, response time, gaze fixation, and reading behavior, among others. The purpose of a logistic regression equation is to find the best fit for the subject matter such as (1) Have you used illegal drugs in the past 2 years? (2) Have you stolen money, products or confidential information from your previous employer? (3) Have you asked for or taken bribes in exchange for preferential treatment of a supplier?

The aim of the equation is to obtain an answer that is reasonable and measurable and that will accurately describe the relationship between the two characteristics of interest (deception or credibility) based on the set of those independent variables considered — pupil dilation, precision response, gaze, response time, fixation, etc.

For EyeDetect, the estimate used in the logistic regression equation chooses one of two parameters that "maximizes the probability" of the classification of deception or credibility. The result is a credibility score.
The Converus Credibility Score represents the probability of belonging to the "Credible" or "Not Credible" group of scores for the test administered. A higher score means it is *more probable* that the score came from the Innocent distribution.

Scores from 50 to 99 are considered Credible and scores from 1 to 49 are considered Not Credible. The closer the Converus Credibility score is to 1, the greater the probability of deception. On the contrary, the closer to 99, the more likely it is the person is telling the truth. But a 55 and a 95 are considered "passing" scores while a 5 and a 45 are considered "failing" scores. Essentially, a score of 51 is as good as a score of 99. This is because the decision

model (algorithm) establishes “50” at the point where errors are balanced. With the decision model set this way, the test sensitivity is .83 (TP) and specificity is .89 (TN) with a mean average accuracy of .86.

This concept is similar to taking a written test on a computer in which you have “passed” the test when a certain percentage of correct responses are given. If you reach that point, the computer stops presenting questions because you are placed in the passing group.

Personal Information

The information gathered by the eye tracker during a test includes:

1. Measurements of the X and Y coordinates of eye movement, pupil diameter and dilation. These are not photographs, are not biometrics and cannot be used to identify any person.
2. The examinee’s true/false responses to the test questions.

If an organization wishes to protect the identity of any examinee for purposes of reporting testing results, during registration prior to taking a test, the Test Proctor can provide an identifying number rather than a person’s name. A person’s name can be used, but it’s not required.

In addition, the Test Proctor can choose to take a photo of the person being tested. If no photo is taken, the organization will need to find test results for a person based on the assigned ID number–after the test is taken, saved, and scored.

EyeDetect Station and Software

The EyeDetect solution is a combination of computer hardware and software.

<u>Hardware – EyeDetect Station</u>
The primary hardware components of the EyeDetect Station include an eye tracking device (infrared camera), a Windows-based computer, wireless keyboard and mouse and chin rest. Noise-cancelling headphones (not shown) are also included.

The eye tracker is a high definition, infrared camera that operates at 60 frames per second. Therefore, the eye tracker takes up to 60 measurements per second of the examinee’s eyes. Changes as small as 1/10th of a millimeter are detected. During the test, hundreds of thousands of eye measurements are recorded, as well as the examinee’s responses to the true/false statements. The eye tracker has a tracking range of 32 x 21 cm at a distance of 60 cm. It also has a +/- 20° horizontal and + 20° / - 40° vertical range.

During a test, eye measurements and test responses are temporarily stored on an encrypted drive on the EyeDetect Station. When the EyeDetect Station is synchronized with the web server, all test and eye data are uploaded to the server for scoring.

<u>Software</u>
The following software programs are used to administer and score EyeDetect tests and to view and analyze the test results:

1) **EyeDetect Software** (computer-based) – software used for test administration; runs on a Windows-based computer (tablet) provided by Converus.

2) **EyeDetect Admin** (computer-based) - software used to configure the computer to record tests and data for specific accounts (called sub accounts).
3) **EyeDetect Manager** (computer-based) – software used by a test administrator to monitor up to four EyeDetect stations running examinations simultaneously.
4) **Dashboard** (web-based) – a web-based dashboard or portal used to see and review test results, to run reports and to manage test licenses, users, and sub accounts.

Security

In terms of system security, each EyeDetect station includes either a secure external hard drive provided by DataLocker or a secure internal disk enabled by Microsoft BitLocker. Both are 256-bit AES encrypted.

All test data are encrypted and can only be accessed by users that enter the key or password to unlock the drive. Test data are encrypted using a unique key per customer before being transferred to our secure data center. Once the data are transferred to the data center it is deleted from the DataLocker or BitLocker drive.

The Converus Dashboard web application is accessible using two-factor authentication. All access to the dashboard is done through SSL. Only authorized users of an account with applicable rights can access the dashboard.
With respect to its Data Center, to store and process user data collected during testing, Converus uses standalone (non-hosted) servers owned by Converus, not the data center. Access to these servers is controlled by a firewall and incoming web traffic is monitored for threats. All servers are housed in a private, locked rack in a SSAE 16/ISAE 3402 certified data center. Access to the data center floor is controlled by key card and biometric scanners and is monitored 24/7.

Process

The examinee is seated in front of the EyeDetect Station monitor. The examinee is asked to quickly and accurately respond to a series of true/false statements for approximately 30 minutes. Responses are given with the computer keyboard. During the test, the eye tracker takes measurements of eye behavior.

Countermeasures

Corrupt examinees may attempt to "cheat" to avoid detection. Additionally, truthful examinees may attempt to influence the test outcome to show they are truthful. Preliminary validation studies indicate that these efforts do impact the testing outcome. EyeDetect monitors subtle deception cues that are impossible to control. Even highly motivated subjects cannot simultaneously control their reading behaviors, response speed, response accuracy, or pupil dilation. Examinees must answer questions rapidly and have little time to attempt physiological countermeasures.

Some examinees may attempt to cheat using makeup (eye liner) and/or eye drops (pupil dilation drugs). These measures are easily detected during the pretest eye-tracking calibration. When this happens, the test can be postponed until the examinee has cleaned off the make-up and/or the effects of the drugs or stimulants have worn off.

Some examinees may attempt to cheat by closing their eyes or squinting. This measure is easily detected in real time by the test proctor via the EyeDetect Manager application or during the test when the calibration screen appears after each brief rest period showing that data loss is significant.

Some examinees may attempt to cheat by answering all questions randomly or by answering all questions with the same response (either all true responses or all false responses). These measures are also detected and a "Not Credible" score will result, either based on "Random Responses" or "Low Reading Comprehension."

Training

The administration and use of EyeDetect requires basic training. The skills required to evaluate test responses and eye measurement data used to derive a test score have been turned over to an algorithm developed by top scientists in the field. The two types of testing roles and training are described as follows:

1) Test Proctor Training – the test proctor will administer EyeDetect tests to examinees. The basic function or role of the test proctor is to welcome the examinee, escort them to the testing station, adjust and calibrate the examinee's eyes to the infrared camera, start the test and monitor the examinee's behavior during the test. Test proctor training takes 3 hours, includes a practical exercise and can be administered remotely using Skype. If desired, there is a Test Proctor certification exam.

   Topics covered in this training include: EyeDetect Station hardware, configuring hardware, performing initial calibrations, overview of EyeDetect Software menus and options, testing room preparation, test preparation, examinee preparation, administering the test, taking a test (such as the acquaintance test, also called a blind number test, or a demo test), submitting test data, and use of EyeDetect Manager.

2) Dashboard Administrator – the administrator manages test results, individual test reports, user access, test licenses, and sub accounts. This training takes 2 hours and can be administered remotely using Skype. If desired, there is a Dashboard Administrator certification exam.

   Topics covered in this training include: accessing the Converus Dashboard, managing and reviewing test results, user login management, account and subaccount management, test licenses management.

There are also two advanced courses, described as follows:

1) Screen Test Writer – any interested person can be trained to write screening tests. The class focuses on RCT and MCT protocols. Practical exercises are required to show competence. This class can be taken remotely or in-person. This training takes one-day when attended in person.

2) Diagnostic Test Writer – any interested person can be trained to write diagnostic tests. The class focuses on the DLC protocol. Practical exercises are required to show competence. This class can be taken remotely or in-person. This training takes one-day when attended in person.

Test Topics

In most cases, EyeDetect tests are written by a trained team using time-tested protocols and techniques based on polygraph testing best practices. Customer requirements are then gathered to create the best possible fit with the pool of examinees. Note: In the U.S., lie detection tests are generally limited to federal, state or municipal government employees. However, in other countries, tests can be administered to employees of private companies.

The following is a list of potential test topics. Topics can be combined in a multiple-issue test with three as the maximum number to be considered. Or diagnostic tests can be designed (single issue).

- Theft
- Drug use
- Divulging confidential information
- Ties to gangs or organized crime
- Bribery
- Document fraud
- Drug trafficking
- Other fraud (financial)
- Money laundering
- Sex-based crimes
- Corporate espionage
- Fuel theft
- Counterfeiting money
- Cyber crimes
- Identity theft
- Terrorism
- Document fraud
- Violent crimes
- Sexual abuse
- Athlete doping
- Unauthorized financial transactions
- Use of date rape drugs
- Parole violations
- Theft of car parts
- Sports event fixing

LEPET

As a specific test example, the EyeDetect law enforcement pre-employment test (LEPET) format is used to screen public safety job applicants or lateral transfers for a variety of disqualifying behaviors such as drug use, undisclosed criminal history, or falsification of the police dept. application.

PCSOT

In the United States, post-conviction sex offender tests (PCSOT) are widely used to test convicted sex offenders now on parole or persons on restrictions or probation. There are EyeDetect templates written for the following testing scenarios:

- Instant offense (first offense) – Did you commit the crime?
- Monitoring – Have you committed the crime again?
- Maintenance – Have you violated the conditions of treatment, probation or parole?
- Sex History – Did you previously disclose all of your past victims?
- Domestic violence – Have you assaulted your spouse or partner?

And for each PCSOT test, there are EyeDetect templates written for the following offenders:

- Adult sex offender with adult victim
- Adult sex offender with child victim
- Adult sex offender – child pornography
- Adult sex offender – voyeur or exhibitionist
- Juvenile sex offender

Specific Test Types

As a more specific example of testing options, consider the following list of test topics that could be used by a financial institution, bank, or insurance company to evaluate current employees based on job description or function. In addition, tests can be written to include the organization name and other familiar information.

1) Falsification of credit applications
2) Unauthorized financial transactions
3) Falsification of insurance claim forms
4) Falsification of company checks
5) Stealing company checks
6) Stealing cash from the bank or from customers
7) Converting company assets to cash illegally
8) Asking for or receiving bribes from contractors or suppliers
9) Providing information or receiving payments from cartels or organized crime
10) Divulging confidential information to unauthorized persons

Vertical Market Use Cases

As a credibility assessment tool, EyeDetect is intended to help organizations in all types of industries in screening job applicants, current employees, law enforcement or security personnel, or any others to reduce corruption, fraud, theft and other inappropriate or illegal behaviors. The following are a few organizational use cases:

- Military school admissions
- Law enforcement pre-employment
- National security ongoing eval
- Drug enforcement ongoing eval
- Border patrol pre-employment
- Customs pre-employment
- Postal service routine evaluation
- Casino personnel pre-employment
- Armored car pre-employment
- Security guard pre-employment
- Petroleum ongoing eval
- Mining
- Shipping/logistics
- Insurance/banking
- Retail
- Manufacturing
- Hospitality
- Medical/dental

Today's Credibility Assessment Methods

Organizations worldwide use a variety of deception detection tools when assessing the credibility of job applicants, employees, and others.

Intuition - the Unassisted Human Lie Catcher Hypothesis

By far, the most frequently used lie detector is human intuition. In other words, many organizations rely on face-to-face interviews to determine the truthfulness of a job applicant or employee. Unfortunately, behaviors such as gaze aversion, touching the body or face, or covering the eyes or mouth while speaking have not been found to be reliable indicators of deception. Despite popular belief, there are no reliable nonverbal deception cues.

Liars, concerned about being believed, often come across as helpful and truthful in an interview, and put more effort into impressing their interviewer. Innocent individuals, under stress, often demonstrate many of the stereotypical behaviors associated with deception: speech errors, fidgeting, and gaze aversion.

A recent effort to summarize over 200 peer-reviewed studies demonstrated that humans have an accuracy rate of 54% at predicting deception.[7] This estimate included accuracy rates from skilled examiners trained in state-of-the-art observational and interrogation techniques. In short, as credibility assessment tools, humans are about as accurate as the flip of a coin.

Integrity Tests

Integrity tests are also frequently used by HR departments as a method of screening. Sadly, studies show that these types of tests are not accurate at determining deception. In a study entitled, "The Criterion-Related Validity of Integrity Tests: An Updated Meta-Analysis" in the Journal of Applied Psychology (2012), researchers reviewed 104 studies of integrity tests. Overall mean observed validity estimates and validity estimates corrected for unreliability in the criterion (respectively) were as follows:

- Job performance = .12 and .15
- Training performance = .13 and .16
- Counterproductive work behavior = .26 and .32
- Employee turnover = .07 and .09

With respect to the ability of integrity tests to predict behaviors:

- Counterproductive work behavior such as substance abuse, theft, and withdrawal
  - Results were moderately accurate from self reports
  - Results were poor from employee records
- Job performance, training performance, and turnover were poor
- Job performance was poorer for non-publisher studies (.12) than publisher studies (.27)

To summarize, integrity tests have a low predictive value for employee behavior for the following reasons:

- Self reporting and attitude tests are easily faked
- Tests do not directly assess many disqualifying factors, e.g., criminal relationships, drug abuse, terrorism, criminal activity, financial problems, etc.

Computer Voice Stress Analyzer

Computer voice stress analyzer technology was developed to record psychophysiological stress responses that occur to the muscles in the voice box when under duress, such as when the consequences of the person's response may be dire. The tightening or loosing of the voice box changes the sound of the voice.

One study on computer voice stress analysis showed that true positive rates (i.e., predicting deception) ranged from 50 to 65% for all conditions and types of materials. However, the "false positive" rate was just as high, and often higher.[8]

[7] Bond & DePaulo, 2006

[8] Journal of Forensic Sciences, 53(1), 183-193, Hollien and Harnsberger (2008).

Other studies that have drawn similar conclusions include:

- Journal of Forensic Sciences, 53(1), 183-193, Hollien and Harnsberger (2008).
- The Journal of the Acoustical Society of America 124(4):2458, October 2008, Hollien and Harnsberger, "Evaluation of two voice stress analyzers," J. Acoust. Soc. Am. 124(4):2458, October 2008.
- Journal of Forensic Sciences 54(3), 2009, Harnsberger, Hollien, Martin, and K Hollien, "Stress and Deception in Speech: Evaluating Layered Voice Analysis."
- National Research Council, Robert Pool, Field Evaluation in the Intelligence and Counterintelligence Context, 2009.
- Journal of Forensic Sciences, Hollien et al., "Evaluation of the NITV CVSA," 2008.

Electroencephalogram and fMRI (brain scanning)

The electroencephalogram has been shown to identify deceptive measures in the brain such as the following:

- Subject views familiar stimuli, which creates a positive baseline (subject sees a friend).
- Subject views unfamiliar stimuli, which creates a negative baseline (subject sees a stranger).
- Person is probed on topics on which truthfulness is tested.

A number of research papers were published on the electroencephalogram that claimed to have achieved results of 87% accuracy (i.e., they could accurately predict if a person was lying 87% of the time) using electrical activity at the surface of the brain. The challenge with this lie detection method is that it is very costly and it's very invasive for the subject being tested.

The fMRI can create real-time 3-dimensional models of the brain using powerful magnets to charge hydrogen protons within cells. A radio frequency is broadcast by these protons, which absorb the frequency and reflect it back at a receiver. This information is translated into an image of the area scanned. Essentially, the fMRI measure increases in blood flow to particular regions, detecting the blood oxygen level-dependent signal. Large patterns of activity are generated in initial brain scans, and irrelevant signals are filtered out. An interior structure and function map are generated to associate external stimulus to changes in localized blood flow.

The fMRI creates a map of pattern vectors associated with specific cognitive states. A radiologist can predict a cognitive state by observing these complex patterns. The radiologist looks at changes in whole-brain patterns and predicts the cognitive state the brain is experiencing. The fMRI has been used to predict simple risky behavior at 72% accuracy.[9]

However, although it can be highly accurate, the fMRI costly to administer. Testing works best when analyzing a pool of subjects, not individuals. There is no industry standardization yet and it is not certain if evasive countermeasures can be used.

Polygraph

While the worldwide standard in lie detection for many decades, the American Polygraph Association (APA) published a study[10] in 2012 that disclosed polygraph accuracy under two scenarios:

1) Event-specific questioning - 80 to 94%
2) Pre-screening and periodic questioning – 72% to 81%

Another study by Krapohl in 2002[11] revealed similar findings. Thus, polygraph remains the most accurate method of deception detection when used for event-specific questioning. However, for pre-screening employees or conducting general periodic testing, polygraph is less accurate. This could result from the condition that exists for a polygraph test that attempts to determine behavior on multiple, potentially unrelated issues. Multi-issue screening tests are less accurate because in that exploratory polygraph interview, the examinee may be questioned about separate issues for which s/he may be guilty on some and not on others. In those cases, there

[9] National Academy of Sciences, February 2014.

[10] Meta-Analytic Survey of Criterion Accuracy of Validated Polygraph Techniques, 2012, table 2.

[11] Krapohl, D. J. (2002). The polygraph in personnel screening. In M. Kleiner (Ed.) Handbook of Polygraph Testing. 217-236. San Diego: Academic Press.

are too many potentially unrelated issues being compared.[12]

The generality of relevant questions in screening polygraph examinations is desirable from the point of view of the hiring agency because these questions cover a wide range of potentially undesirable behaviors of concern. However, the generality of relevant questions may introduce ambiguity in the mind of the examinee about their guilt ("I haven't used illegal drugs in past 90 days, but I used them 6 months ago, and I know that was wrong.") The generality of relevant questions also increases their similarity to comparison questions, which are intentionally vague and broad in scope.

In addition, in screening contexts, the same set of test questions may be used repeatedly for different examinees. Standardization of test protocols lessens concerns about question formulation. However, the validity of a polygraph screening test that includes probable lie questions is likely to be compromised not only by variance in the skills of examiners but also the extent to which the relevant questions are broad enough in scope to meet the needs of the testing organization (Meijer, Verschuere, Merckelbach, & Crombez, 2008).

Conversely, in specific incident criminal investigations, if the polygraph examiner conducts a proper and professional pretest interview with well-constructed relevant questions that are clear, not subject to interpretation, have probative value, and includes a proper reading of the physiological recordings, the accuracy of probable-lie tests in specific-incident criminal investigations is approximately 90%, excluding inconclusive outcomes.[13]

<u>EyeDetect</u>

No other deception detection tool, prior to EyeDetect, has shown similar rates of accuracy as the polygraph under any scenarios. The reactions to two sets of relevant questions are compared, and each relevant issue serves as a control for the other issue.

Also, because EyeDetect is almost completely automated, test validity does not depend on the interview skills of the examiner or the ability to properly interpret physiological recordings. Because the test is administered by a computer, the examinee cannot effectively influence the examiner and the examiner cannot intimidate or discriminate against the examinee.

EyeDetect tests can be developed for any language, allowing testing in the native language of examinees and eliminating any needs for a translator.

EyeDetect is less intrusive because no sensors are attached to the examinee.

EyeDetect can potentially be ten times more efficient when considering the time required to conduct a screening test vis-a-vis polygraph. One skilled polygraph examiner can conduct 3-4 tests per day, considering each test takes about 2 hours. One EyeDetect test administrator using four EyeDetect Stations can conduct up to 40 tests per day.

EyeDetect Limitations

Examinees must have proficient reading skills. All tests require that the examinee read the true/false statements. Ongoing research may show this can be obviated by having the examinee listen to the test questions being read to them.

EyeDetect examinees must have two functioning eyes (whether they can see from both eyes is not critical); EyeDetect tests can be administered to examinees with only one seeing eye.

[12] Barland, Honts and Barger, 1981 experiment number 2.

[13] American Polygraph Association Ad Hoc Committee on Polygraph Techniques, 2011

The following list shows eye diseases or conditions and their likely impact on an EyeDetect test.

Eye Diseases	EyeDetect is OK	Potential Problems w/ EyeDetect	Notes
• Amblyopia		Yes	
• Astigmatism	Yes		OK with glasses
• Blepharitis		Yes	
• Blepharospasm		Yes	
• Cataracts		Yes	
• Allergic conjunctivitis		Yes	
• Color blindness	Yes		
• Macular degeneration		Yes	
• Entropion and Ectorpion		Yes	
• Strabismus		Yes	
• Glaucoma		Yes	
• Hyperopia	Yes		OK with glasses
• Lagophthalmos	Yes		
• Tearing	Yes		
• Myopia	Yes		
• Dry eye	Yes		
• Presbyopia or tired eye	Yes		OK with glasses
• Eyelid ptosis		Yes	
• Keratitis		Yes	
• Keratoconus	Yes		OK with glasses
• Diabetic retinopathy		Yes	
• Hypertensive retinopathy		Yes	
• Sjogren's syndrome	Yes		

If someone has any of the following conditions, they are generally not suitable for a polygraph exam, but can be tested with EyeDetect without affecting test results:

- Anxiety Disorder or Panic Disorder
- Post-Traumatic Stress Disorder
- Attention Deficit Disorder
- Asperger’s syndrome or high functioning autism
- Heart conditions such as arrhythmia or pre-mature ventricular contraction
- High blood pressure
- Asthma or related respiratory issues
- Pregnancy

References

Peer-reviewed

1. Kircher, J. C., and Raskin, D. (2016) Laboratory and Field Research on the Ocular-motor Deception Test. European Polygraph Journal, Volume 10, Number 4 (38).

2. Cook, A. E., Hacker, D. J., Webb, A. K., Osher, D., Kristjansson, S., Woltz, D. J., & Kircher, J. C. (2012). Lyin' Eyes: Ocular-motor Measures of Reading Reveal Deception. Journal of Experimental Psychology: Applied, 18(3), 301-313.

3. Patnaik, P., Woltz, D., Hacker, D., Cooke, A., Francke-Ramm, M., Webb, A., and Kircher, J. (2016) Generalizability of an Ocular-Motor Test for Deception to a Mexican Population. International Journal of Applied Psychology, 6(1): 1-9.

4. Hacker, D. J., Kuhlman, B., & Kircher, J. C., Cook, A.E., and Woltz, D.J. (2014). Detecting deception using ocular metrics during reading. In D. C. Raskin, C. R. Honts, & J. C. Kircher (Eds.), Credibility assessment: Scientific research and applications. Elsevier, pp 159-216. (AUTHOR/PUBLICATION REQUIRE PURCHASE)

5. Kuhlman, B. B., Webb, A. K., Patnaik, P., Cook, A. E., Woltz, D. J., Hacker, D. J., & Kircher, J. C. (2011, September). Evoked Pupil Responses Habituate During an Oculomotor Test for Deception. Poster presented at the Society for Psychophysiological Research convention, Boston, MA. (abstract)

6. Patnaik, P., Woltz, D.J., Cook, A.E., Webb, A.K., Raskin, D.C., and Kircher, J.C. (2015, March). Ocular-motor Detection of Deception in Laboratory Settings. Meeting of the American Psychology and Law Society, San Diego, CA.

7. Webb, A. K., Hacker, D.J., Osher, D., Cook, A.E., Woltz, D. J., Kristjansson, S. K., and Kircher, J. C., (2009). Eye movements and pupil size reveal deception in computer administered questionnaires. In D. D. Schmorrow, I. V. Estabrooke, & M. Grootjen (Eds.), Foundations of Augmented Cognition. Neuroergonomics and Operational Neuroscience (553-562). Berlin/Heidelberg: Springer-Verlag.

8. Webb, A. K, Honts, C. R., Kircher, J. C., Bernhardt, P.C., and Cook, A. E. (2009). Effectiveness of pupil diameter in a probable-lie comparison question test for deception. Legal and Criminal Psychology, 14(2), 279-292. (AUTHOR/PUBLICATION REQUIRE PURCHASE)

9. Kircher, J. C. (2018). Ocular-Motor Deception Test. In J. Peter Rosenfield, Detecting Concealed Information and Deception (pp. 187-212). Cambridge, MA: Academic Press. doi:10.1016/B978-0-12-812729-2.01001-6. (AUTHOR/PUBLICATION REQUIRE PURCHASE)

--

Non-peer reviewed

10. Osher, D. (2006). Multimethod assessment of deception: Oculomotor movement, pupil size, and response time measures. (Doctoral dissertation), University of Utah, Department of Educational Psychology.

11. Webb, A.K. (2008). Effects of Motivation, and Item Difficulty on Oculomotor and Behavioral Measures of Deception. (Doctoral dissertation), University of Utah, Department of Educational Psychology. (ISBN: 9780549980032)

12. Patnaik, P. (2013). Master's Thesis: Ocular-motor methods for detecting deception: Direct versus indirect interrogation. University of Utah, Department of Educational Psychology.

13. Patnaik, P. (2015). Doctoral Dissertation: Oculomotor methods for detecting deception: Effects of practice feedback and blocking. University of Utah, Department of Educational Psychology.