

**Voice Stress Devices and the Detection of Lies**  
**Donald J. Krapohl, Andrew H. Ryan, and Kendall W. Shull**

Published 2002 in:

*Policy Review*, the official publication of the International Association of Chiefs of Police (IACP) National Law Enforcement Policy Center.

### **Introduction**

Throughout history, one of the most common and difficult problems for law enforcement officers has been determining whether a potential suspect is lying or telling the truth. Many an investigation has stalled, been diverted, or failed, because the statements of a suspect couldn't be verified, regardless of how much effort the investigators expended. Law enforcement has often turned to science with this problem, and science has tried to provide a solution. In the last 100 years several techniques and devices have been proffered to police to help them separate fact and fiction in suspects' stories, from interrogation drugs to word association tests, reaction time tests to interpretations of body postures, eye contact and gestures, - even a machine that registers hand trembles. The polygraph notwithstanding, no lie-detection method has enjoyed much longevity. In the last 25 years, however, newer methods have been introduced that are purported to offer more convenience, accuracy, and utility than even the polygraph. These are the voice stress devices, and by now most police agencies have seen or heard about them. They come with modest up front costs, making them very attractive to cash-strapped departments. Marketers of voice stress instrumentation portray their product as an important tool to hundreds of local police agencies, that they solve crimes quickly, and help in the selection of qualified police candidates.

The prospect of detecting lies in a speaker's voice has intuitive appeal. A common experience for most people is to have spotted a fib simply by noting a change in the tone of a speaker's voice. It would seem logical then, that with the help of computers and advanced technology, the lies of suspects could similarly be detected. If a device could find that special something that happens in the voice when a person is trying to deceive us, it could be a boon for the criminal investigation process, as well as for a variety of other uses such as business negotiation, confirmation of treaty compliance, airport security, and insurance claim verification, to name a few. The purpose of this article is to review what is known about voice stress, and to assess to what degree this technology can provide a reliable means for detecting deception.

### **The Benefits**

Voice stress devices offer several potential advantages over the standard polygraph, the reigning lie-detection technology. The training time to operate a voice stress device is less than that for polygraph training, and there are no academic prerequisites to receive that training. Very low training and education requirements can save taxpayer money, and put the devices in the hands of more officers. The voice stress examinations themselves take little time, averaging about 45 minutes per session, or about a half or third of the time needed for a typical polygraph examination. There are no sensors placed on the body, only a small microphone clipped to the examinee's clothing. Because only the voice is used, the examinee need not even be present during the examination. A recording from a remote location or time can be processed with the equipment. Not only might this be more convenient than transporting the suspect to the voice

stress technician, but it also opens the door to surreptitious processing of previously recorded voices.

Though thousands of the devices have been sold over the years, a far smaller number remain in service after a few years. Despite convenience and low cost, there are problems with voice stress devices that the product manufacturers have not yet overcome. The most pressing shortcoming appears to be the level of accuracy these machines deliver. As will be taken up later in this article, the track record of voice stress analysis in careful empirical studies has been lackluster. This has forced promoters to rely heavily on personal testimonials as evidence of accuracy. Accuracy may not be all that important in some applications, of course, such as when the device is used only as an adjunct to an interrogation. In an interrogation setting, a sophisticated–appearing machine in the room that is represented as a lie-detector may offer the interrogator a psychological wedge to encourage more candor from the suspect. Whether the machine really detects anything is secondary, so long as the suspect believes it works. Repeated judicial decisions have supported the use of ploys and trickery by law enforcement to help obtain a confession, so long as the tactic is not so coercive as to “shock the conscience.” Given the non-intrusiveness of current voice stress products, it seems likely that they would likely withstand that test.

### **The Costs**

Low validity is not without drawbacks, some potentially severe. Poor accuracy could have profound consequences for a department if investigative decisions are based on the outcome of the voice stress examination. Precious manpower resources could be misdirected, or a criminal could escape while another citizen is wrongly pursued, affecting not only public safety, but community confidence, as well. Also, the use of the devices in a surreptitious mode raises very imposing legal questions, issues that are beyond the scope of this article, but are especially important when placed in the context of the devices’ unimpressive accuracy.

Of more immediate concern are those instances where departments use voice stress tests in the hiring process for new officers. In 1999, the American Bar Association published an article (Palmatier, 1999) which indicates that the use of a voice stress device for hiring decisions may constitute a violation of the Equal Employment Opportunity Commission rules, and that the departments could find themselves in legal peril if they use them in this manner. It is because of the twin problems of validity and potential litigation that voice stress technology has played a limited roll in state and local law enforcement, and none at the federal level.

### **Federal research**

Recognizing its possible applications, the US Government’s interest in voice stress technology can be traced back at least to the 1960s. A number of government agencies independently investigated the potential of this method of lie detection, but since the mid-1980s the task has fallen largely to the research facilities at the US Department of Defense Polygraph Institute (DoDPI). In its charter, DoDPI is responsible for providing research in new concepts and technologies with relevance to the detection of deception. Though DoDPI’s research into alternative methods, including voice stress, has been ongoing for over 10 years, those efforts have taken on new emphasis since the terrorist attacks on New York City and Washington D.C. Voice stress is currently one of the hot topics, and DoDPI has conducted or collaborated in several studies on voice stress devices that can provide answers to agencies and departments weighing the potential costs and benefits of fielding them.

The premise of all voice lie detectors is that certain pitch parameters, allegedly associated with certain nervous system activities, are not under voluntary control. Voice stress device marketers suggest that there is an inaudible component in the vocal spectrum, called the micro-tremor, which changes during stress. Micro-tremors are oscillations in the FM component of the voice, in the range of 8 to 14 cycles per second, and purportedly are markers for the stress associated with the act of deception. According to descriptions by the manufacturers, micro-tremors are normally seen in relaxed, natural speech. Their disappearance signals stress, with the inference that the speaker is uncomfortable with what he is saying. In the field, a voice stress technician asks a structured series of questions during the voice stress examination, some questions relating to the crime and others being neutral. By noting the presence or absence of micro-tremors on the crime questions, a decision is rendered regarding the examinee's truthfulness.

If such a relationship between these micro-tremors and deception were empirically sound, government security professionals and law enforcement would have a powerful new tool, not only to replace the polygraph, but for applications where the polygraph cannot be used. A review of the current research is presented here to bring the reader up to date with the findings.

The first significant commercially available product to analyze vocal signals was the Psychological Stress Evaluator (PSE), introduced in the early 1970s by Dektor Counterintelligence and Security, Inc. of Springfield, Virginia. All analyses were conducted off line, using an audio recording of the examinee's voice taken during a structured test. Quick and inexpensive, hundreds of PSEs were sold, though they never achieved the acceptance enjoyed by the polygraph, due in large part to the lack of supporting evidence that it could actually detect deception (Brenner, Branscomb, & Schwartz, 1979; Hollien, Geison, & Hicks, 1987; Horvath, 1978, 1979; Lynch & Henry, 1979; Timm, 1983; Waln & Downey, 1987).

In the late 1980s, the National Institute for Truth Verification, Inc. (NITV) of West Palm Beach, Florida produced what they termed a computer voice stress analyzer and trademarked it as CVSA. The CVSA is marketed as a convenient replacement for the polygraph. Like the PSE, the CVSA analyzes micro-tremors in the vocal signal, but unlike the PSE, the CVSA provides real-time graphical outputs or charts that examiners can score. Sales of the CVSA have been brisk in recent years, easily overshadowing other brands of voice stress devices. The US Government purchased a small number of units, and trained a few personnel, but after field trials with the devices did not meet expectations, the equipment was discarded. Widespread use of the CVSA in the law enforcement sector, combined with the Government's continuing interest in new lie detection methods, prompted DoDPI to conduct and sponsor a number of studies to answer two important questions. First, can micro-tremors in the vocal signal be used effectively to detect deception? Second, how does this compare to the current gold standard, the polygraph?

Cestaro and Dollins (1996) examined the utility of using the vocal responses of subjects during a low-stress test involving the examinee concealing a number. Parameters examined were spectral energy distribution of the voice response, fundamental frequency, response energy, response duration, and pitch variations around the fundamental frequency. No significant relationships were found in the voice data for vocal components and deceptive answers. The authors concluded that none of these parameters, in isolation, was a reliable and valid discriminator of truth and deception. However, they left open the possibility

that multiple measures extracted from pitch information might be useful as indicators of deception.

A two-part study by Cestaro (1996) was conducted because controlled lab research to test the validity or reliability of the CVSA instrument or the techniques employed in its use had not been conducted. The first part was designed to determine whether the CVSA detects micro-tremors in the fundamental frequency of presented signals as the manufacturer claimed. The second part was designed to determine whether accuracy rates obtained using the CVSA differed from those using the traditional polygraph. The results of part one indicated that the CVSA functioned electrically according to the manufacturer's theory of operation. Changes in the frequency of the input signal caused deflections of the CVSA display in proportion to the frequency of the input signal. Because the study demonstrated that the CVSA recorded what it purported to be recording, the issue of decision accuracy was then ready to be investigated.

In the second part, the CVSA was compared to the polygraph, again in a low-stress lab study. Forty-two subjects were tested with both instruments. The difference in accuracy between the polygraph and CVSA was significant: polygraph decisions were significantly greater than chance, while the CVSA decisions were not. The authors concluded that poor instrument or procedure sensitivity of the CVSA was the cause for the lack of accuracy.

It had been shown that the CVSA did not perform well in low-jeopardy scenarios, and it therefore became important to test it in settings in which the outcome was more meaningful to the examinee. In a mock crime study Janniro and Cestaro (1996) again evaluated the accuracy of the CVSA to detect deception. One hundred nine subjects were tested; half were asked to commit a realistic and engaging mock crime while half did not participate nor had knowledge of the mock theft. CVSA examiners conducted and scored the exams in accordance with NITV procedures. Charts were blind-scored by three other CVSA evaluators. The variable of interest was the number of correct decisions, with chance accuracy set at 50%. Blind CVSA evaluators made correct decisions on 49.8% of the cases, while the testing CVSA examiners achieved 48.6% accuracy. These accuracy rates were not different than chance. The authors concluded in this laboratory paradigm that, though the examiners consistently employed the NITV scoring methods, the CVSA sensitivity to detect lies was low.

The last DoDPI study with this device was a collaborative project conducted jointly by DoDPI and the US Army Walter Reed Hospital (2000). The project examined the capabilities of the CVSA in a well-understood and controlled stressful interview model (U.S. Army Soldier of the Month Board). In this study, voice responses before, during, and after the interview were transferred to CVSA charts for blind scoring by CVSA evaluators. In addition, other indices were recorded before, during, and after the interview using validated medical measures of physiological stress. These included heart rate, arterial blood pressure, and plasma ACTH. Salivary cortisol measures were made before and after the interview. The results showed that the interview paradigm elicited stress at significant levels, as indicated by the medical markers of stress. Results for the CVSA did not correlate with the medically confirmed stress at any level, neither low nor high. In addition, inter-CVSA examiner agreement proved to be low. The authors conclude that the CVSA analysis of voice features does not reflect well-validated tonic responses to acute stress. In other words, whatever the CVSA may record, it is not stress. The makers of the CVSA have since suspended cooperation with federal research with their product, and have required some new

buyers of their equipment to agree not to participate in government-sponsored validity research.

The studies outlined above cast strong doubt on the ability of micro-tremor analyses to detect deception better than chance. A new product, Vericator by Integritek Systems, Inc. of Tampa, Florida, has recently been introduced that claims to extract information from the entire vocal signal to produce decisions (2000). It has been marketed in a manner that emphasizes its flexibility and utility across a wide range of situations and circumstances. DoDPI's interest has grown in the possible use of the Vericator as a tool that could facilitate the work of inspectors at security checkpoints (e.g., US Customs), a setting where polygraph examinations are not practical. DoDPI commenced a two-site comprehensive study to assess the accuracy of the Vericator to detect deception. Detection rates at both sites, which involved very realistic and stress-inducing laboratory paradigms, were quite disappointing and did not exceed chance (Brown, Senter, & Ryan, 2002; Sommers, Brown, & Ryan, 2002.)

### **Conclusion**

Over the last 30 years other researchers outside of the government have also researched voice stress for lie detection, and published their findings in scientific journals. The general conclusion has been that the accuracy is modest to poor for a handful of experimental approaches, and uniformly poor for those relying on the micro-tremor (see [www.voicestress.org](http://www.voicestress.org) for a summary of the available research). This does not prevent some as-yet untried analytical approach from someday yielding a valid voice lie-detector, and the Government is still aggressively seeking such a capability for the important advantages it would afford. As a practical consideration, the poor validity for the current voice stress technology should provide a caveat to agencies considering adding voice stress to their investigative toolboxes.

The controversy over the use of voice stress analysis will surely continue for years to come. Additional research offers the prospect that voice could become one channel in the next generation of lie-detection instrument that might also include brain waves, eye movement, thermal imaging, remote sensing, or some technology that does not yet exist. Though none of the current voice analysis technologies are valid for detecting deception, the US Government's continuing investigation in this area might one day find one that works, with the goal of better protection of our communities and our nation.

### **References**

- Brenner, M., Branscomb, H., & Schwartz, G.E. (1979). Psychological stress evaluator: Two tests of a vocal measure. *Psychophysiology*, *16*(4), 351-357.
- Brown, T.E., Senter, S.M, & Ryan, A.H. (In press). Ability of the Vericator to detect smugglers at a mock security checkpoint. Abstract.
- Cestaro, V.L. (1996). A comparison between decision accuracy rates obtained using the polygraph instrument and the Computer Voice Stress Analyzer (CVSA) in the absence of jeopardy. *Polygraph*, *25*(2), 117-127.
- Cestaro, V.L., & Dollins, A.B. (1996). An analysis of voice responses for the detection of deception. *Polygraph*, *25*(1), 15-342.

- DoDPI Research Division Staff, Meyerhoff, J.L., Saviolakis, G.A, Koenig, M.L., Yourick, D.L. (2000). Physiological and biochemical measures of stress compared to voice stress analysis using the Computer Voice Stress Analyzer (CVSA). Report No. DoDPI98-R-0004. Fort Jackson, SC.
- Hollien, H., Geison, L., Hicks, J.W. (1987). Voice stress analysis and lie detection. *Journal of Forensic Sciences*, 32(2), 405-418.
- Horvath, F.S. (1978). An experimental comparison of the psychological stress evaluator and the galvanic skin response in detection of deception. *Journal of Applied Psychology*, 63(3), 338-344.
- Horvath, F.S. (1979). Effect of different motivational instructions on detection of deception with the psychological stress evaluator and the galvanic skin response. *Journal of Applied Psychology*, 64(3), 323-330.
- Janniro, M.J., & Cestaro, V.L. (1996). Effectiveness of detection of deception examinations using the Computer Voice Stress Analyzer. *Polygraph* 27(1), 28-34.
- Lynch, B.E., & Henry, D.R. (1979). A validity study of the psychological stress evaluator. *Canadian Journal of Behavioural Science*, 11(1), 89-94.
- Palmatier, J.J. (1999). The Computerized Voice Stress Analyzer: Modern technological innovation or 'the emperor's new clothes'? *GP Solo & Small Firm Lawyer*, 16(4), 42-45.
- Sommers, M.S., Brown, T.E., & Ryan, A.H. (In press). Evaluating the reliability and validity of Vericator as a voice-based measure of deception. Abstract.
- Timm, H.W. (1983). The efficacy of the psychological stress evaluator in detecting deception. *Journal of Police Science and Administration*, 11(1), 62-68.
- Vericator User Manual* (2000). Integritek Systems, Inc.: Tampa, Florida.
- Waln, R.F. Downey, R.G. (1987). Voice stress analysis: Use of telephone recordings. *Journal of Business and Psychology*, 1(4), 379-389.

### **About the authors**

Donald J. Krapohl is a researcher with the Department of Defense Polygraph Institute. He can be contacted at [krpohld@jackson-dpi.army.mil](mailto:krpohld@jackson-dpi.army.mil).

Dr. Andrew H. Ryan is the Chief of Research, US Department of Defense Polygraph Institute. He can be reached at [ryana@jackson-dpi.army.mil](mailto:ryana@jackson-dpi.army.mil).

Kendall W. Shull recently retired as Chief of the FBI's Polygraph Unit. He is now in private practice as Kendall Shull Investigations and Polygraph Services in Knoxville, Tennessee. He can be contacted at [KendallShull@aol.com](mailto:KendallShull@aol.com), or (865) 742-7744.

The views expressed in this article are those of the authors, and do not necessarily represent those of the Department of Defense or the US Government.