

## **A Summary of the Testimony Before the Texas Legislature Regarding the Reliability and Validity of the Computer Voice Stress Analyzer**

**A Summary by Victor L. Cestaro, Ph.D. March 7, 2001**

During my tenure as a researcher at the Department of Defense Polygraph Institute at Fort McClellan, Alabama, between 1993 and 1999, I performed research using the National Institutes for Truth Verification (NITV) Computer Voice Stress Analyzer (CVSA). Prior to beginning any research using that instrument, I attended a one week CVSA examiner training course conducted by NITV's chief instructor at that time, Captain David Hughes, a retired police officer. After completing the course, I was given certification as a CVSA examiner. The training session was given at the Fort Lauderdale (Florida) Sheriff's Department and was attended by law enforcement personnel from various state and local agencies, including the Fort Lauderdale Sheriff's Department.

A major portion of the training was devoted to assessing the amount of "blocking" – squareness of – and the degree of "diagonality" --- ramping of – the voice patterns collected on the CVSA voice charts. The percentage of "blocking" -- 80% or higher – was the determining factor for deception (or the stress associated with that deception). After a significant number of hours, most students could agree on a percentage to assign to a chart tracing – within 5 to 10 percent. This scoring ability established that potential examiners could independently assign a number to a tracing and essentially be in agreement. In other words, this process established some degree of reliability in scoring *chart tracings*, at least during the training session. However, it did not establish any validity for the measure. In other words, it did not confirm that the chart tracings were actually a measure of the respondents' levels of stress. There were no laboratory sessions conducted during this training in which students could operate the instrument using subjects within a mock crime scenario. The manufacturer states that there is insufficient "jeopardy" on the part of subjects in a mock crime situation, and that the lack of "jeopardy" can compromise the validity of the instrument. It is evident that the manufacturer does not feel compelled to resort to the scientific method to validate the instrument.

During my research, I attempted to validate the instrument's ability to measure respondents' stress levels. Over a period of approximately three years of conducting research using the CVSA, I was unable to establish that the instrument could detect differential levels of stress, or provide any indication that the respondents were being truthful or deceptive. Additional independent studies I performed, using laboratory-grade equipment and computer sound spectrum analysis software, did not provide any evidence that voice analysis is efficacious for differentiating levels of stress. All of the aforementioned results were published in several government technical papers, and in the journal *Polygraph*, which is published by the American Polygraph Association. As a result of my studies, the Institute issued a policy statement regarding the lack of effectiveness of voice stress analysis for detecting deception.

Additionally, collaborative research was conducted at the Walter Reed Army Institute of Research (WRAIR) in Washington, DC. In this highly controlled research, and in several previous studies, the research staff at WRAIR had established that specific biochemical measures of blood and saliva, and physiological measures – such as blood pressure and heart rate – were directly associated with subject stress. In a robust study in which blood and saliva samples, blood pressure and heart rate, and voice samples were collected before, during, and after a highly stressful psycho-social event, the CVSA chart analyses did not correlate with the biochemical and physiological measures. All of the medical measures indicated that stress levels were highest during the stressing event. The CVSA charts tracings, collected collaterally with the biochemical and physiological measures, were randomly scrambled and scored by several examiners chosen from a list provided by the staff of NITV. The examiners scored the CVSA charts independently and were not given the identities of the other scoring examiners. Although the medical measures were highly correlated with the situations, the CVSA results were not.

Additionally, there was very little agreement among the examiners, indicating very poor scoring reliability – making the validity issue extremely questionable. The conclusion drawn in that study was that no basis was found for recommending the use of CVSA technology for medical assessment nor for the detection of deception. However, the utility of other voice stress analytic technologies was not ruled out. I am not aware of any published controlled studies using other voice stress equipment or technologies.

To my knowledge, no other scientific research has been conducted using the CVSA, and there has been no scientific evidence presented that would support the contention that the CVSA is capable of detecting differential levels of stress, or differentiating between truth and deception at any level greater than chance. Unless and until there is compelling scientific evidence to the contrary, it is my opinion that the CVSA is not capable of distinguishing truth from deception in human speech.