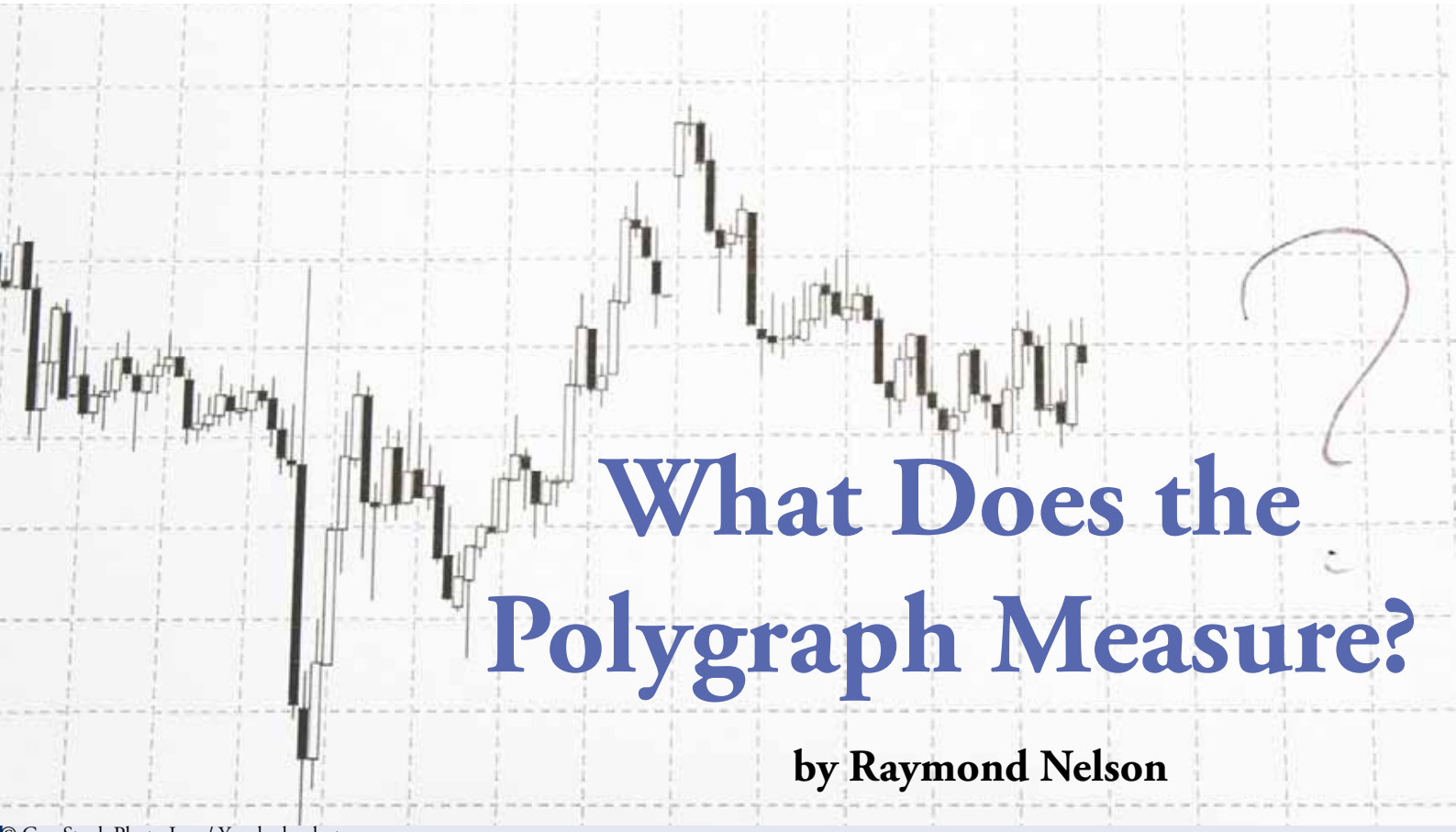

What Does the Polygraph Measure?

by Raymond Nelson

Does the polygraph measure lies? Why is it called a lie detector test? Does it measure fear? Is there some physiological response or activity that is uniquely associated with lying? Can lies be measured in a manner similar to the ways that physical things or events can be measured? Is there such a thing as a lie detector test, or a test that actually detects lies? Are there any other causes or activities besides lying that could cause reactions to polygraph questions? Are the results infallible? Have the test result ever been wrong? Can someone learn to fake the test results? What normative data exist to quantify our present knowledge regarding reactions from guilty or innocent persons? What is the confidence level or probability of error and associated confidence interval for test accuracy?

These and other questions have been heard repeatedly by examiners and scientists regarding the polygraph test. They are not merely rhetorical questions, and they are not asked for the simple purpose of being critical. Rather, these questions are often be asked in the spirit of administrative, scientific and legal, scrutiny. Neglecting to study and develop realistic evidence-based answers to these questions is simply unwise. Underlying these questions are

more fundamental questions about the underlying theory or construct on which the polygraph test functions. While the pursuit of answers to these questions may be the domain of researchers and scientists more than field practitioners, working professionals at all levels face the potential for damaging the outcome of individual cases or damaging the credibility and stature of the profession if they are completely unprepared to respond confidently to these questions.

Pioneers in polygraph testing faced the challenge of making order out of chaos for the first time, and did not enjoy the luxury of an existing knowledge-base in the way we do today. In that context test administration was largely a clinical process in which the intuition and experience of the examiner was centrally featured as the basis of test effectiveness and test accuracy – in the absence of procedural methods and normative data to support a structured quantitative solution. This can be seen historically in the methods devised by Keeler, in which quantitative analysis was not a large concern; by Reid, in which numerical analysis began to become important - though Reid, in the absence of an existing knowledge-base to guide decisions about cut-scores, wisely refrained from imposing the use of fixed cutscores and instead emphasized a more clinical approach to the decision concerning the final test result; by Backster who sought to overcome the myriad of emerging subjective clinical approaches through the development of a more structured procedural methods, using numerical cut-scores that seem to have been devised either intuitively or through some unpublished study; by the U.S. Government polygraph programs that sought to standardize the test by clarifying the structure of decision rules and fixed cutscores; by the group of researchers at the University of Utah, who published their empirical and statistical analysis regarding numerical and statistical scoring methods; and by the more recent Emperical Scoring System, based on decades of prior knowledge and evidence, in attempt to structure and simplify the application of norm-reference statistical procedures to polygraph scoring tasks. Clinical methods are excellent solutions in the learning context, and remain excellent solutions whenever validated structured objective methods do not yet exist.

One obvious upside of a clinical approach is the sense of professional satisfaction that ensues when

exercising a high level of expertise to an effective solution and resolution of an important matter. Another advantage of a clinical approach is the rich volume of information that can be developed for later scrutiny. As with most things there are advantages and disadvantages to every solution. The downside of a clinical approach is that the methods are inherently subjective, and also inherently difficult to teach and learn. As with other forms of testing, the emergence of methods for numerical scoring and statistical analysis have led to increases in both test accuracy and test reliability. That is where we stand today: it is difficult to find an ethical position from which to advocate for a less accurate and less reliable clinical approach when we have quantitative models that have been shown to provide better overall validity. The optimal solution will be to leverage the advantages of the two approaches: a clinical approach to develop rich sources of information, and a quantitative approach to ensure validity and reliability.

Returning to the question about 'what does the polygraph measure,' answering this questions is a process itself which can benefit from some explanation and definition. Scientific questions can be said to be enduring questions, in that we will never know everything: there is always more to learn. In that context it will be important to refrain from the temptation or impulse to express our present knowledge as if it is absolute knowledge that it not subject to change or modification when new information becomes available. Indeed resistance to change in the context of new knowledge – including observed phenomena that cannot be adequately accounted for with existing explanations – is a hallmark of dogma, pseudoscience, and the pursuit of goals that pertain more to individual egos, businesses or other organizations than to knowledge or the profession as a whole. The pursuit of knowledge is a never ending process, simply because it is humanly impossible to know everything. There is always more to learn. Our task is to continue learning, and to continue making good use of new emerging knowledge, including information and technology. Neglecting to make use of new knowledge is to risk becoming an anachronism.

What does the polygraph measure? The simplest and most general explanation is that the polygraph, like other tests, measures response to a stimulus. Stimulus-and-response, in some form,

is a construct that is central to virtually all forms of testing: present a stimulus and observe/measure the response. Although interesting, this explanation is a bit too generic to be very useful. At a more practical level the polygraph test can be said to measure physiological responses, involving the respiratory, cardiovascular and integumentary systems, in addition to somatic activity in skeletal muscle. Although more descriptive, this still does not satisfy the initial inquiry about what the polygraph actually measures. Indeed, it would be avoidant of the question to stop at this point. Advancing deeper into the detail, we note that those physiological reactions are associated with arousal in the autonomic nervous system, occurring in response to the test stimulus questions. This begs the next question: what is the relationship between autonomic arousal and lying? And the next question - what is the potential to observe the same autonomic arousal when someone is not lying? In the realm of science, there is always another layer of detail to investigate.

Questions about the correlation between autonomic arousal and lying are empirical questions, for which the answer exists in the form of a statistical correlation to describe the strength of association between observed autonomic responses and lying about a past behavior. These questions also involve the potential that observed autonomic arousal is caused by something other than lying. More precisely, what is the potential to observe autonomic response when a person is lying or not lying in response to a stimulus question that describes a past behavioral fact? In the polygraph testing context this question is further refined - what is the potential to observe certain differences in autonomic responses to target and control stimuli when a person is lying about a behavior described by the target stimuli? These questions underscore the importance of the study and development of normative data to quantify and describe our knowledge about responses that are expected to be typically observed among guilty or innocent persons who undergo polygraph testing.

Related empirical questions will involve the psychological basis of physiological responses. It was suggested many years ago that the psychological basis of reaction was emotion and fear, and this was a satisfactory explanation until the theoretical premise was evaluated with respect to physiological responses that are common to both

comparison question techniques and concealed information tests, and with regard to the similar effectiveness of comparison question tests formulated with probable-lie questions and directed-lie questions. As with other forms of science our working theories and hypotheses must be evaluated with consideration for both historical data and also for new information and new observed phenomena. When the data and evidence do not agree with the theory or hypothesis, then one of them must begin to change. Good scientific investigation, like good criminal investigation, does not allow for changing the data or evidence; it is the theory or hypothesis that must begin to change. In the polygraph testing context the emotion-fear hypothesis may not have been completely wrong, but is now viewed as largely incomplete. A more satisfactory working theory would recognize that research in psychophysiology tells us clearly that physiological differences between various emotions cannot be observed or measured by field polygraph instruments. A more complete working theory would therefore include a range of potential emotions as potentially underlying responses to test stimuli, and would also recognize that cognition or mental activity and behavioral conditioning appear to play a greater role than had been discussed in decades past.

While it is useful to develop a more satisfactory working theory about the psychological basis of response, there is always more to learn. Continued scientific inquiry may eventually help us to more effectively discriminate between various emotions, and may help us to better describe the mechanisms and relationship between behavioral experience and responses to test stimuli, and the relationship between the mental activities of deception and physiological responses to test stimuli. Continued scientific investigation may eventually help us to understand more precise details about the physiological mechanisms and actions that can be observed and recorded – and the correlation between these mechanisms and deception or other human phenomena.

Regardless of the depth of our knowledge, there will always remain more to learn – simply because it is humanly impossible to know everything. And regardless of the depth of our knowledge about the minute details of the psychological and physiological mechanisms underlying responses to test stimuli, the meaning

and interpretation of recorded and observed responses will remain an empirical and probabilistic question. This is because, as far as our present knowledge suggests, all physiological and all psychological activities have multiple causes and multiple purposes. Answers to empirical questions are unavoidably probabilistic concerns for which we must be careful that there is no false expectation of a deterministic or perfect solution for which randomness and error play absolutely no role. The alternative to learning to account for polygraph testing in a probabilistic manner would be for us to engage in pretense and dishonesty around the potential for absolute certainty.

Because there is no such thing as a perfect test that works every time without error for every person, our task – as a scientific community – is to continue to develop our knowledge-base and normative data to quantify and account for the probability that our test result is correct or incorrect. Our task – as a community of field practice professionals – is to learn to understand and make use of the best scientific knowledge available. The choice to neglect the best knowledge and information, or to emphasize procedures, practices and assumptions based on a less accurate or less complete knowledge-base, would betray our scientific obligation and betray the trust placed in us by our communities, agencies and countries to provide the most accurate possible assessment of an individual credibility.

What does the polygraph measure? The polygraph test, like other tests, measures response to the test stimulus. Responses to polygraph test stimuli are observed and measured in the form of differences in the strength of responses as the examinee is presented with comparison or control stimuli in sequence together with investigation target stimuli that describe the examinee's possible involvement in or relationship to a factual or behavior issue of concern. While the basis of responses is psychological, the observed and recorded responses are physiological, allowing the polygraph instrument and polygraph examiner to observe and record the relative salience or differential salience of the different types of test stimulus questions. Salience can be thought of as a function of the basic principles and mechanisms of psychology, including emotion, cognition, and conditioning. Because the test stimuli refer to behavior, physiological responses are interpreted with the assumption that they are correlated with behavior – as long as

the observed responses are timely with the test stimuli and no interference or distracting stimuli is present during testing. Published evidence has repeatedly confirmed the operational construct that responses will tend to load onto relevant or comparison stimulus questions as a function of deception or truth-telling. Recorded data from multiple physiological sensors are aggregated together in the form of a numerical structural model that is compared with our knowledge, in the form of published normative data, about responses commonly observed among deceptive and truthful persons. If our knowledge is anchored in the form of quantitative information that is representative of the population and the examinee, recorded responses to test stimuli can be used to determine the empirical or statistical likelihood of a correct or incorrect result when the examiner has concluded that the responses from the examinee conform more closely to those of persons known to be deceptive or truthful.

One answer to the question concerning the question 'why it is called a lie detector test,' is this: the term lie detector test is a term of convenience, not a term of science. It is a term that conveys the goal and purpose of polygraph testing in a concise way that is easily understood by untrained persons and non-professionals. Despite the value of simplicity, it would be unwise for working professionals to actually limit or restrict understanding and intelligence to the simplistic hyperbole of 'lie detector test.' Instead, experts, professionals, and even technicians are all obligated to understand the details

Scientific questions can be said to be enduring questions, in that we will never know everything; there is always more to learn. In that context it will be important to refrain from the temptation or impulse to express our present knowledge as if it is absolute knowledge that it not subject to change or modification when new information becomes available.

© ShutterStock / Sergey Nivens

and nuances of the test, so that they can better use the test capabilities and better account for their work when the need arises. In scientific terms the purpose of the test is not so much to detect lies but to discriminate between deception and truth-telling.

Tests can be said to be valid, in a scientific sense, if they increase the effectiveness of our discrimination, beyond what we might achieve by random chance alone, at statistically significant levels. This leads to another inevitable discussion about the need for statistical confidence intervals – and the potential hazards of neglecting these – but that is a topic for another time. Published studies at the present time indicate that some validated polygraph techniques are capable of average accuracy over 90% with a 95% confidence range from 86% to 96%. Scientific reviews tend to converge at an average accuracy rate that averages 89% for diagnostic exams with a confidence range of 83% to 95%.

Data at this time suggests that polygraph tests interpreted with the assumption that criterion state of the target questions vary independently may provide lower accuracy, with an average near 85% and a confidence range from 77% to 93%. The aggregated average accuracy of the range of validated techniques has been reported as 87% with a confidence range from 80% to 94%. While the reported averages tell us the trend or central tendency, the lower limit of the confidence range gives a more cautious or conservative view of the worst-case-scenario. Claims of accuracy in excess of the upper limits of these ranges are not consistent with the volume of available published evidence, and may simply be too good to be true.

Discussions about polygraph accuracy will inevitably prompt questions or discussions about the potential that someone can learn to fake the test result or defeat the polygraph test. Potential vulnerability to faking is a concern that is pertinent to all types of testing. This is again a complex topic for another time. Minimally, we must remember that there is no such thing as a perfect test. Polygraph accuracy depends, in part, on effective interviewing, effective test administration, and effective test data analysis. Test accuracy also depends on good instrumentation, good questions, and a suitable examinee who is represented by a knowledge-base and normative data. Effective faking strategies would have to increase testing errors beyond the confidence interval

surrounding known error rates. In the event that a testing error does occur it would be difficult to know whether the cause of error is related to intentional effort, or is simply within the known range of test accuracy and inaccuracy. The National Research Council (2003) wrote, "Claims that it is easy to train examinees to "beat" both the polygraph and trained examiners require scientific supporting evidence to be credible" p147.

Scientific proof of the effectiveness of faking strategies would have to show, with no potential for corrective action or remedy, a reproducible result in which the rate of error could be increased above the upper limit of the confidence range of the normally expected error rate. Of course, many tests, including the polygraph test, may include methods designed to identify, detect, and deter faking attempts.

Why is it not a truth-detector? Why not try to use the polygraph in a more optimistic paradigm? Questions about truth, and the compelling discussions they provoke, are also for another time. For now, it will have to suffice to say that truth is a philosophical question in the realm of epistemology and knowledge – with its own complex quandaries involving what kinds of things can be true, and what it means to say that something is true. The short answer is that opinions, beliefs, emotions and future events – or anything that cannot be factually verified – are not epistemological truths, unless we endorse a postmodern or deconstructionist view of truths as a form of subjective perspective or subjective experience.

A somewhat rationalist perspective would hold that truths are statements about facts that can potentially, though not always easily, be verified through empirical observation. From a statistical measurement perspective, positive test results inform us that the test result does not conform to our knowledge and evidence about how truthful persons respond. Polygraph test results that are statistically significant for deception signify that there is a high probability that external evidence will eventually be discovered to show that the person was involved in the behavioral issue of concern. In the practical work of polygraph and lie detection, we engage a form of pragmatic truth in which we accept persons as truthful when we are reasonably confident – based on a stated alpha level that describes our required confidence level or tolerance for error - that they are not lying.