Scientific (Analytic) Theory of Polygraph Testing

Raymond Nelson

The analytic theory of polygraph testing is that greater changes in physiological activity are loaded at different types of test stimuli as a function of deception and truth-telling in response to relevant target stimuli.

Stimulus and *response:* this is the basic idea underneath a scientific test. For example: a stimulus in a lie detection or credibility assessment test such as the polygraph test is a question that describes something for which the subject can be either truthful or deceptive. Although sometimes procedurally complex, using a scientific test is a conceptually simple matter: present the stimulus, then observe and record the response. Complete several stimulus and response trials. Aggregate the response data, and compare the result to a probability reference model. Probability models can be empirical models, calculated from observed data, and can also be theoretical models, calculated from information subject to mathematical and logical proof – the strongest form of scientific information.

Scientific terms: hypotheses, theories, laws of science

Scientific ideas begin as questions about how reality, the universe, works. Suggested answers or explanations to questions about reality and the universe are referred to as *hypotheses,* and these must be subject to testing and investigation before they are accepted. When an idea or explanation is inconsistent with reality (i.e., inconsistent with evidence) it is referred to as a *false-hypothesis* and must be discarded. Continued reliance upon a false hypothesis as if it is an adequate scientific explanation for reality has been referred to as *pseudoscience* [See Shermer (2011) for more information about pseudoscience.] False-hypothesis must be discarded and replaced with ideas that can be reconciled with evidence from reality.

When an idea cannot be falsified after reasonable attempts to do so – when an idea is not inconsistent with available evidence – it then is accepted as a working *theory*. Scientific theories are hypotheses that are *supported* by the available evidence. A theory will describe what we can reasonably say, based on the available evidence, about how reality and the universe works. A theory is a *model*, but a theory itself is neither reality nor the universe. To paraphrase the physicist Neils Bohr "It is wrong to think of the task of [science] as the study of nature. [Science] is the study of what we can say about nature" (Peterson, 1963). A theory is simply an abstract representation – an attempt to understand and explain – reality and the universe, constructed in verbal language or other structured and replicable form of expression and communication.

A good theory will account for the broadest range of evidence with the simplest explanation, referred to as

parsimonious, and will be consistent with other scientific knowledge and theories. Scientific ideas are never actually proven, they are merely supported by available evidence. To the extent that there is always more to learn, scientists will continue to apply and test our theories with increasingly available evidence. When a scientist or scientific thinker discovers that theory or idea cannot be reconciled with some new evidence or other scientific ideas then some modification or replacement of the theory is necessary. To the extent that all theories and all models can only incompletely describe reality and the universe, it has become an aphorism in science that "All models are wrong but some are useful" (Box, 1976). All ideas and all assertions about reality are ultimately an approximation.

A general principle of science is that we may never know everything about the universe, and so the task of learning and increasing our knowledge will be forever ongoing. The purpose of a scientific investigation or scientific experiment is to test a hypothesis against a basis of evidence. A fundamental requirement of any scientific idea or explanation is that it is *falsifiable* (Popper, 1959). That is, there exists some means to test a scientific idea so that we can reject it if it is incorrect.

Rationale for scientific testing

The purpose of a scientific test is to quantify some interesting phenomena that cannot be subject to perfect deterministic observation – for which the outcome is not influenced by human behavior, and not affected by random variation and is therefore always exactly the same – and also cannot be subject to physical measurement – subject only to measurement error – which would require both a physical substance and a well defined unit of measurement. While much of science can be thought of as attempting to explain or understand the outcome of some process, scientific tests, including the polygraph test, are often concerned with *classification*, and *prediction*.

Classification refers to the determination of a class or category to which a case can be assigned. Prediction can refer to the expected likelihood that a conclusion about a single case will concur with other information from reality and the universe, or to the likelihood that concurrent information is available in reality and the universe. Prediction can also refer to an expected proportion of possible cases for which we can expect to achieve a correct classification.

Scientific tests, because they are in-

tended to quantify phenomena for which there is no physical substance. Tests accomplish the task of quantification through the measurement and combination of proxy data, for which there is a known statistical relationship between the data and the criterion of interest. Scientific tests are based on scientific theories. Without a theory it is difficult or impossible to understand and evaluate the strengths and limitation of a test and also difficulty to quantify the level of effectiveness (i.e., range or confidence interval) that can be reasonably expected. And it is similarly difficult or impossible to quantify the range of error that can be reasonably anticipated when a test is applied to reality.

Scientific tests, because they are intended to quantify amorphous phenomena, are not expected to be infallible (e.g., tests of personality adjustment or intellectual functioning). Tests are only expected to quantify the margin of uncertainty or level of confidence associated with a conclusion. A good test will produce probabilistic estimates for which both experience and evidence from reality and the universe will tend to agree with the estimated or predicted proportions of correct conclusions and classification error. A poor test will give probabilistic estimates that are inconsistent with reality.

Although direct control over outcomes cannot often be achieved, the ability to make better outcome predictions will enable us to weigh the occurrence of correct decisions with the economic costs associated with correct decision errors, and ultimately increase the effectiveness at decision-making and the achievement of practical and operational goals and objectives. Professionals who use scientific tests and those who use scientific test results are expected to learn to communicate and think probabilistically and to make use of probabilistic information.

An analytic theory for polygraph testing

The analytic theory of polygraph testing is that greater changes in physiological activity are loaded at different types of test stimuli as a function of deception and truth-telling in response to relevant target stimuli. An advantage of this analytic theory is that it does not depend on mind-reading or guessing about the un-falsifiable and un-verifiable subjective emotional experience of the examinee. Instead, it describes what we expect to observe in the recorded data. This analytic theory does not depend on metaphoric language such as "strong reaction," for which there is no physical strength involved. The phrase "greater changes in phys-

iological activity" is a factual and descriptive statement that begins with an assumption that physiological activity is an ongoing process for which changes will occur in response to test stimuli.

According to this analytic theory, interpretation of changes in physiological activity is a matter partitioning the observable, measurable, and quantifiable variation in the recorded data for different types of test stimuli. Most importantly, this analytic theory for polygraph testing does not depend on the false hypothesis that responses are driven by fear, or any other single emotion, or any single psychological process. Nor does it depend on the false premise that different emotions will manifest in physiological response differences that can be observed or recorded by field polygraph recording instrumentation.

Implicit in this analytic theory for polygraph testing is the idea that human physiology is active, and that presentation of the test stimuli can be expected to induce observable changes in activity. The issue of interest to the polygraph is whether the changes in physiological activity are systematically (i.e., non-randomly) loaded for the different types of stimuli. Interpretation of systematic loading is a matter of whether the numerically quantified data do or do not achieve a statistically significant level. In this way the test data, test scores, and the reproducible analytic test results serve as a basis of evidence to support a scientific conclusion about deception or truth-telling (Nelson, 2015).

An analytic theory for polygraph testing does not depend on unquantified subjective or impressionistic judgments about observed *patterns* or activity signatures in the plotted or displayed waveform for the recorded time-series data. The pattern of interest during the analysis of recorded polygraph data is not the graphical shape of the plotted wave-forms after signal processing. Instead the pattern of interest is the loading of responses for different types of test stimuli. This pattern can be observed in the recorded physiological data only when a sufficient volume of data is recorded. The analytic theory of the polygraph test is a falsifiable empirical theory: analysis of field and laboratory sampling data will show that greater changes in physiological activity either are or are not loaded at different types of test stimuli as a function of deception or truth-telling in response to the test target stimuli.

Finally, this analytic theory for polygraph testing does not attempt to fully account for or describe the under-

lying psychological or physiological processes that explain or account for the recorded signals. Without doubt there are interesting questions about the cognitive and emotional and behavioral basis for observable and recordable physiological responses to test stimuli, just as there are important questions about the exact details of the physiological mechanisms that are captured by the recording sensors themselves, and difficult questions about the correlation of recorded signals with the criterion of interest and covariance of sensor data with data from all of the other sensors. These complex and difficult questions are best addressed incrementally, else progress towards a complete systematic theory will be potentially handicapped by wasted time and attention on hypotheses that are not consistent with reality.

Contrast with earlier hypotheses

In years past, if we were to ask polygraph field practitioners to explain the theory of the polygraph and we would likely have heard discussion about fear, threat and consequences as the basis of responses to relevant and comparison stimuli. It was also hypothesized, though incorrectly, that different emotions might manifest differently in recordable physiological activity.

The fear hypothesis was first suggested at a time when the use of statistical models and analytic methodologies was beyond the skill set or imagination of most polygraph field practitioners. At that time, polygraph data was recorded by tracing ink onto a moving paper – where the ink on the paper was the actual data. This is in contrast to the polygraph instrument of today, for which the displayed data is recorded in a time-series (i.e., a series of successive recorded samples) of digital and numerical values that can be subject to signal processing, feature extraction, statistical analysis, and graphical display. The idea of fear as a theoretical explanation for polygraph responses was first suggested at a time when numerical scoring was regarded as simply a teaching tool, a crutch, for practitioners who lacked sufficient experience and expert judgment to render decisions by merely looking at the recorded data.

Few people could have correctly anticipated the importance and implications of what had been demonstrated by Meehl (1954) about clinical and statistical conclusions in the mid-20th century. Instead, it seems to have been expected that polygraph field practitioners with sufficient experience and expertise would not relay on numerical scoring and would instead simply look at the data to achieve conclu-

sion based on visual analysis alone. In the absence of attention to statistical analysis and probabilistic conclusions, the emphasis in the polygraph profession was solely on attempting to explain the psychological processes or mechanism that underlie observed differences in physiological reaction to different types of polygraph test stimuli. In contrast, and analytic theory for the polygraph attempts to explain the data and what the data can tell us about whether practical conclusions of deception or truth-telling are likely to concur with reality.

The *fear and threat hypothesis* has been referred to as the idea of *psychological set* within the polygraph profession, though it has been pointed out by Krapohl (2001), Honts (2000) along with Handler and Nelson (2007) and Senter, Weatherman, Krapohl and Horvath (2010) that the term is not used in the field of psychology in the same way as it is in the polygraph profession and is without scientific support as an explanation for polygraph responses. The notion that fear is a basis of observable and recordable physiological activity is inconsistent with published evidence showing that polygraph techniques that make use of directed-lie-comparison (DLC) questions (Department of Defense, 1995a, 1995b; Honts & Raskin, 1988; Horowitz, Kircher, Honts & Raskin, 1997; Prado, Grajales & Nelson, 2015a, 2015b) can provide criterion accuracy rates that may equal or exceed that of probable-lie comparison (PLC; Reid, 1947; Summers, 1939) question formats (American Polygraph Association, 2011; Horowitz, Kircher, Honts, & Raskin, 1997).

A corollary to the fear hypothesis would be that the polygraph test might not be effective with psychopathic persons, for whom some evidence has shown have low levels of fear conditioning (Birbaumer *et al.*, 2005; Veit *et al.*, 2013). Fear conditioning may be related to the ability to learn from one's consequences and subsequently modify future behavioral choices. Evidence again does not support the fear hypothesis, as the polygraph test has shown to be effective with psychopathic persons at rates similar to non-psychopathic persons (Balloun & Holmes, 1979; Barland & Raskin, 1975; Patrick & Iacono, 1989; Raskin & Hare, 1978), despite their differences in the ways they subjectively experience emotions such as fear.

Taken together, similar levels of polygraph effectiveness with psychopathic and non-psychopathic persons, the effectiveness of DLC polygraph techniques, and the fact that polygraph instrumentation is known to be incapable of discriminating between basic

emotions such as fear, anger, disgust, sadness and happiness (see Kahn, Nelson, & Handler, 2009 for a discussion) and also incapable of discriminating the reason for an emotion (e.g. fear of the examiner or fear of consequences for a behavior) all indicate that the fear hypothesis, though perhaps at one time interesting and useful, is in need of wholesale replacement as an explanation for polygraph responses.

The fear hypothesis becomes even more problematic when considering that the consequences for an innocent/truthful person whose polygraph test results appear deceptive (i.e., a false-positive error) are generally identical to the consequences for a guilty/deceptive person who produces a deceptive polygraph result. Instead of attempting to guess about the subjective experience of the examinee, a satisfactory analytic theory for the 21st century polygraph test will describe what we can observe to observe quantitatively and probabilistically in the recorded polygraph data when a person is deceptive or truthful.

The analytic theory described herein can be applied and tested, and has been, for both PLC and DLC techniques, both types of which have been shown to produce similar effect sizes. It is also consistent with the more general concept of salience, as suggested by Handler and Nelson (2007) and *differential salience* as applied to the polygraph test by Senter, Weatherman, Krapohl and Horvath (2010). It can also be generalized to other types of polygraph tests such as the *concealed information test (*also described as a *guilty knowledge test* in some literature) and even the *relevant-irrelevant test*. All that is necessary is to develop suitable statistical reference model to quantify the probabilistic values associated with different possible conclusions about different types of test stimuli. For example, in the concealed information test the types of stimuli can be thought of as the investigation target stimulus and all other stimuli. The analytic theory states that greater changes in physiological activity will be loaded at different types of stimuli as a function of concealed information in response to the investigation target stimuli. Because responses to concealed information test stimuli are encoded as 0, 1, or 2, for several trials (referred to as keys), the reference model for the concealed information test is a multinomial distribution [See Handler, Nelson & Kuczek, 2015 for a discussion.]

Summary and conclusion

The future of the polygraph test and the polygraph profession depends in part on the identification of a hypoth-

esis or theory that is consistent with the requirements of science and available scientific evidence. The analytic theory proposed herein meets those requirements, and does so without introducing new ideas and without introducing necessary changes to testing methodologies. Scientific studies have for decades supported the validity of this theory (American Polygraph Association, 2011; Honts & Peterson, 1997; National Research Council, 2003; Office of Technology Assessment, 1983; Senter *et al.*, 2010). [See Nelson and Handler, (2013) for a brief history of scientific reviews of polygraph test accuracy].

The analytic theory of polygraph testing describes the data that field examiners can work with numerically, statistically and analytically. It is one which field examiners and scientists have been using for decades time whenever they numerically score and quantify polygraph test results. The analytic theory for polygraph testing is falsifiable. Perhaps most important, it is consistent with decades of scientific research on the effectiveness of the polygraph test at discriminating deception and truth-telling by evaluating differences in the loading of greater changes in physiological activity in response to different types of test stimuli.

Without doubt there are deep and important questions that remain to be explored concerning the underlying physiological responses to polygraph stimuli, along with perhaps even deeper and more difficult questions about the subjective cognitive and emotional experiences of the examine, the degree to which these are conscious or unconscious, the degree to which these experiences are correlated with past behaviors and experiences, and the degree to which polygraph questions may function as a form of conditioned stimulus will remain important but are beyond the scope of work of most field polygraph examiners. For field practitioners, our present knowledge of the psychological basis for observed responses to polygraph stimuli can be assumed to involve multiple psychological processes, including emotion, cognition, attention, motivation, memory and conditioned learning. In a larger sense, it will be wise to continue to expand our polygraph theories as generally within the scope of all available knowledge from psychology, physiology, recording instrumentation, measurement and analytic theories. For the present time, the simplest and most effective approach towards a working theory for practical or applied polygraph testing will be to limit the discussion to information that we can expect to observe in the test data.

If the polygraph is merely a bogus-pipeline tool (Jones & Sigall, 1971) or interrogation prop to be used for obtaining confessions – if the test results themselves are never to be regarded with any value of their own – then the definition of a suitable working theory is neither important nor useful, nor necessary. But if the polygraph is merely an interrogation prop, then it will only be a matter of time before some other technology begins to replace the polygraph test in circumstances for which a scientific test result is desired.

If it is correct that there exist some physiological activities for which there are identifiable differences in their correlation with deception and truth-telling, then it is only a matter of time before scientists and technologists begin to exploit that those physiological activities in a commercialized or productized test format and algorithmic decision model. It will be a mistake for the polygraph profession to attempt to coexist with new scientific credibility assessment tests while relying on a false explanatory hypothesis that centers on un-testable and un-falsifiable subjective emotional experiences that cannot be discriminated by available polygraph recording instrumentation. A polygraph test that premised on false hypotheses, developed at a time when interpretation of polygraph data was limited to subjective expert/clinical judgment that did employ numerical scoring, statistical decision theory or data analytic methods that are possible today, will be vulnerable to becoming an anachronism.

In years past the polygraph test was the only scientific test for credibility assessment and lie detection, and there may have been little motivation for the polygraph profession to advance its foundational explanatory theory. A hypothesis that had superficial appeal was satisfactory even if inconsistent with known phenomena. Today, in the early 21st century, new technologies are emerging and will continue to emerge in the lie detection and credibility assessment space. Those new technologies will not successfully enter the marketplace without a sound working theory, without effective recording technology, and without proven methodologies based on analytic and statistical models.

A scientific theory for polygraph testing must not be inconsistent with reality or other knowledge, even if it means temporarily limiting the range and depth of phenomena for which a theory attempts to make assertions. Attempts to develop a basis of scientific knowledge, or an area of professional practice, on ideas that are inconsistent with reality will result only in ritualism, mysticism, and discon-

nection from other areas of science and technology, and will not result in intellectual, technological and engineering advances that will ultimately contribute toward the achievement of human goals and objectives.

A theory for polygraph testing that is consistent with reality and other scientific knowledge will enable the profession to advance. Professional practices based on false hypothesis will instead remain static, unable to make use of new knowledge and new technologies and new analytic methods. A satisfactory theory for polygraph testing will explain the observable data and evidence, and will be consistent with our knowledge in other areas of science, including measurement theory, test theory, physiology, and psychology.

Because tests are not intended to be a form of perfect deterministic observation, for which neither random variation nor human behavior will change the outcome, nor a form of physical measurement, which requires both a physical phenomena and a physical unit of measurement, all scientific test results are inherently probabilistic and are therefore inherently analytic. In a larger sense, all scientific conclusions, whether from field study, laboratory study, meta-analysis, Monte-Carlo or other, give only statistical approximation of reality.

If it were possible to achieve a precise measurement of reality, or if it were possible to satisfy our important questions with simple and perfect deterministic observation, then we would not need a test. It is a paradox of reality and the human condition that some of the most interesting and important things that we may want to quantify may turn out to be the most difficult things to quantify. The purpose of a scientific test is to obtain and analyze data that can serve as a statistical proxy to improve our conclusions and decision-making by probabilistically quantifying some amorphous phenomena. Our task is to understand the basis of scientific testing and probabilistic measurement so that we can continue to improve our conclusions and decision-making.

The analytic theory for polygraph testing – that greater changes in physiology are loaded at different types of test stimuli as a function of deception or truth-telling in response to relevant stimuli – describes what we expect to observe when we obtain and analyze polygraph data. Experience with reality will continue to tell us whether this theoretical model, as a description of what we can expect to observe in polygraph data, is satisfactory to assist us with the task quantifying the lev-

el of confidence or margin of uncertainty associated with polygraph test results. Ultimately, the validity of our theoretical and probability reference models will be observed as a function of whether the predicted proportion of correct and incorrect polygraph results corresponds to our calculations and predictions about the correct and incorrect classification of deceptive and truthful polygraph results. A clearly defined analytic theory will help the polygraph profession to advance.

References

American Polygraph Association (2011). Meta-analytic survey of criterion accuracy of validated polygraph techniques. Polygraph, 40(4), 196-305. [Electronic version] Retrieved August 20, 2012, from http://www.polygraph.org/section/research-standards-apa-publications

Balloun, K. D. & Holmes, D. S. (1979). Effects of repeated examinations on the ability to detect guilt with a polygraphic examination: A laboratory experiment with a real crime. Journal of Applied Psychology, 64, 316-322.

Barland, G. H. & Raskin, D.C. (1975a). Psychopathy and detection of deception in criminal suspects. Psychophysiology, 12, 224.

Birbaumer, N., Veit, R., Lotze, M., Erb, M., Hermann, C., Grodd, W., & Flor, H. (2005). Deficient fear conditioning in psychopathy: a functional magnetic resonance imaging study. Archives of General Psychiatry 62(7), 799-805.

Box, G. E. P. (1976). Science and statistics. Journal of the American Statistical Association, 71, 791–799.

Handler, M. & Nelson, R. (2007). Polygraph terms for the 21st Century. Polygraph, 36, 157-164.

Handler, M., Nelson, R., & Kuczek, T. (2015). A review of the Lykken scoring and evaluation system for the concealed information test. APA Magazine, 48(2), 49-57.

Honts, C. R. (2000). A brief note on the misleading and the inaccurate: A rejoinder to Matte (2000) with critical comments on Matte and Reuss (1999). Polygraph, 29, 321-325.

Honts, C. R. & Peterson, C.F. (1997). Brief of the Committee of Concerned Social Scientists as Amicus Curiae United States v Scheffer. Available from the author.

Honts, C. & Raskin, D. (1988). A Field Study of the Validity of the Directed Lie Control Question. Journal of Police Science and Administration, 16(1), 56-61.

Horowitz, S. W., Kircher, J. C., Honts, C. R. & Raskin, D.C. (1997). The role of comparison questions in physiological detection of deception. Psychophysiology, 34, 108-115.

Jones, E., Sigall, H. (1971). The bogus pipeline: a new paradigm for measuring affect and attitude. Psychological Bulletin, 76(5), 349–364.

Kahn, J., Nelson, R. & Handler, M. (2009). A Exploration of Emotion and Cognition during Polygraph Testing. Polygraph, 38, 184-197.

Krapohl, D. (2001). A brief rejoinder to Matte & Grove regarding 'psychological set'. Polygraph, 30, 203-205.

Meehl, P. E. (1954). Clinical versus statistical prediction. University of Minnesota Press.

National Research Council (2003). The Polygraph and Lie Detection. National Academy of Sciences.

Nelson, R. (2015). Polygraph testing as a single-subject scientific experiment. APA Magazine, 48(5), 101-106.

Nelson, R. & Handler, M. (2013). A brief history of scientific reviews of polygraph accuracy research. APA Magazine, 46(6), 22-28.

Office of Technology Assessment (1983). The validity of polygraph testing: A research review and evaluation. [Re-printed in Polygraph, 12, 198-319.].

Patrick, C. J. & Iacono, W.G. (1989). Psychopathy, threat and polygraph test accuracy. Journal of Applied Psychology, 74, 347-355.

Petersen, A. (1963). The philosophy of Neils Bohr. Bulletin of the Atomic Scientists 7, 8-14.

Popper, Karl (1959). The Logic of Scientific Discovery. New York, NY: Basic Books

Prado, R., Grajales, C., & Nelson, R. (2015). Laboratory study of directed lie polygraphs with Spanish speaking examinees. Polygraph, 44(1), 79-90.

Prado, R., Grajales, C., & Nelson, R. (2015). Laboratory study of a diagnostic polygraph technique in a single sequence: a replication. Polygraph, 44(2), 1-12.

Raskin, D. C. & Hare, R.D. (1978). Psychopathy and detection of deception in a prison population. Psychophysiology, 15, 126-136.

Reid, J. E. (1947). A revised questioning technique in lie detection tests. Journal

of Criminal Law and Criminology, 37, 542-547. Reprinted in Polygraph 11, 17-21.

Research Division Staff (1995a). A comparison of psychophysiological detection of deception accuracy rates obtained using the counterintelligence scope Polygraph and the test for espionage and sabotage question formats. Report number DoDPI94-R-0008. DTIC AD Number A319333. Department of Defense Polygraph Institute. Fort Jackson, SC. Reprinted in Polygraph, 26(2), 79-106.

Research Division Staff (1995b). Psychophysiological detection of deception accuracy rates obtained using the test for espionage and sabotage. DoDPI94-R-0009. DTIC AD Number A330774. Department of Defense Polygraph Institute. Fort Jackson, SC. Reprinted in Polygraph, 27, (3), 171-180.

Senter, S., Weatherman, D., Krapohl, D., & Horvath, F. (2010). Psychological set or differential salience: a proposal for reconciling theory and terminology in polygraph testing. Polygraph 39(2) 109-117.

Shermer, M. (2011). What Is Pseudoscience?: Distinguishing between science and pseudoscience is problematic. Scientific American: 92.

Summers, W. G. (1939). Science can get the confession. Fordham Law Review, 8, 334-354.

Veit, R., Konicar, L., Klinzing, J. G., Barth, B., Yilmaz, Ö., Birbaumer, N. (2013). Deficient fear conditioning in psychopathy as a function of interpersonal and affective disturbances. Frontiers of Human Neuroscience 7, 706.