An Essay: Reporting Reproducible Results of Scientific Polygraph Tests

By Raymond Nelson

©FotoliaLLC/Rawpixel.com

Results of scientific tests should be reproducible. Reproducibility of an analytic result is a differentiating characteristic between scientific tests for which test validity is based on the structured use of quantitative methods, and unquantified clinical evaluation methods for which the validity of a result will depend primarily on the subjective expertise and persona of the evaluator. Both reproducibility and replicability are related to the concept of reliability. Replicability is a scientific concern, and refers to whether the results of a scientific study can be reasonably expected to be observed again if a study were to be repeated with different sampling data. Reliability, in scientific testing, refers to repeatability – whether a particular test result can be achieved again. Discussions about reliability include both re-test reliability and inter-scorer reliability. Re-test reliability refers to the degree to which a test result will be observed again if a test is repeated. Inter-scorer reliability is an important concern whenever a test result is based on subjective feature extraction and whenever the analytic procedure will be completed manually. Reproducibility refers to whether an analysis of a dataset can be repeated to achieve the same analytic result.

Reproducibility of scientific polygraph test results will depend, in part, on the availability of information about the data aggregation and numerical transformation procedures that were used, in addition to the availability of information about the statistical reference models and decision rules. Quite obviously, re-analysis of polygraph test data with a dif-

ferent analysis protocol (i.e., differences in feature extraction, data reduction and data aggregation, statistical reference models, or decision rules) will be more likely to result in a difference in the analytic conclusion. Differences in analytic conclusions will lead inevitably to confusion, frustration and distrust, especially among the most concrete thinking persons who may be un-prepared to think probabilistically about scientific test results. In the most extreme cases, persons who have not developed an adequate understanding of the probabilistic nature of all scientific test results may engage in indiscriminate pretense or expectation for deterministic perfection, and this will lead to eventual confusion and frustration when confronted with the need to understand testing errors.

Conceptual content of a report of a scientific polygraph test result

In order to educate and inform those who receive and read the results from scientific polygraph tests, and to promote the reproducibility of test results, reported information should be standardized. A polygraph report should include sufficient information about the method of analysis that an interested reader could access more complete information about the analysis and testing model. It should also include some information about the numerical test scores and cut-scores that were used to determine the test result. Reported information should also include some information about the probabilistic test result and probability decision thresholds, along with information about the use of any statistical correction to account for multiplicity effects when completing multiple statistical comparisons, such as those that can occur when using subtotal scores. A complete written report will also provide information about the meaning of the probabilistic test result with regard to other potential conclusions that have been rejected as a result of the testing process. Finally, a written report should contain clear information about the categorical test result whenever probabilistic test results are intended to be reduced to nominal outcome categories.

Analytic method

Reported results of scientific polygraph test should include a description of the analysis method, for which there should be some published description of the analytic procedures. Absent an adequate published description of the analytic method, professionals are at risk for ignoring or inventing procedures as a matter of convenience, and quality assurance reviewers will be unable to reproduce and verify the correctness of the analysis and result. When a result is un-reproducible the strength of the conclusion will be based on the persona of the expert who will undoubtedly claim to possess some deep magic or esoteric knowledge that is not available for scrutiny. A more satisfactory form of scientific conclusion will be achieved when the result is based on a published analysis method that is evidence-based, norm-referenced and standardized. Minimally, a test report should include enough information about the method of analysis so that an interested person could access and digest the published literature in order to ascertain whether the analysis was completed in a correct and competent manner.

Numerical test scores and cut-scores

Reported results of scientific polygraph results should minimally include the grand total or sub-total scores that were used to determine the test result. The report should also include information about numerical transformations, data aggregation, and data reduction. This can often be accomplished by simply identifying the scoring protocol, for which detailed information can be obtained in available publications.

Test results that are achieved with the use of numerical decision cut-scores should include information about those decision cut-scores. Cut-scores are a matter of both science and administrative policy, and are intended to achieve identified testing objectives such as achieving high levels of test precision or accuracy, or maintaining test sensitivity or test specificity at desired levels, or constraining false-positive or false-negative errors to below specified tolerances. Quite obviously, attempts to reproduce an analysis and testing result may be ineffective if decision cut-scores are unknown or different from those used in the initial analysis.

Probabilistic test results and probability cut-scores

Probabilistic test results should be provided in the written test report. Information in a written test report should also include information about the probability reference model or computational reference model so that an interested reader can obtain more information to answer questions about model and procedural validity. This can sometimes be achieved easily by identifying the analysis protocol. Test results based on frequentist inference will require the explicit a priori declaration of a tolerance for error, and testing reports should include information describing the probability cut-scores or alpha boundaries for statistical significance. Test results based on Bayesian inference will require an explicit declaration of assumed prior probability values for the possible test outcomes. Any use of statistical error correction should be clearly described if they are used to account for multiplicity effects that are anticipated when completing multiple statistical comparisons, such as when using subtotal scores to classify polygraph test outcomes. Absent any description, a reader will assume that no statistical correction was used to optimize the desired testing or mission objective.

Statistical and probabilistic results can be provided in a number of different ways, including through the use of p-values that describe the probability of error or level of statistical significance, or through any of number statistical descriptions of the estimated effect size or strength of the testing evidence such as odds, ratios, odds ratios, risk ratios, confidence intervals, proportions, probabilities or confidence levels. There may be advantages and disadvantages to the use of different statistics. Regardless of the type of statistical information that is presented, expert professionals should be expected to achieve some familiarity with the meaning of various statistical terms.

Explanation of the realistic meaning of the test result

A report of a scientific polygraph test should include a statement describing the practical

interpretation that can be made from the probabilistic test result. This statement of explanation should answer the following question: what can be reasonably said in verbal language regarding the probabilistic test result? Interpretation of any scientific test will always be premised on a recognition that all knowledge is yet incomplete and all scientific conclusions are relative to some clearly identified alternative. For this reason, a correct interpretation or explanation of the test result will factually describe the meaning of the probability result with regard to both the observed test data and the alternative possible conclusions that the examiner set out to choose from.

In the case of a scientific polygraph test that is scored and interpreted using empirical reference distributions – calculated from empirical observation of the sampling scores of deceptive and truthful cases – the different possible conclusions may involve deception and truth-telling. A conclusion that the test data are indicative of truth-telling will be considered against the possibility that the data and conclusion are erroneous and were produced by a deceptive person. Similarly, a conclusion that the test data are indicative of deception will be considered against the potential that the data are erroneous and were produced instead by a truthful or innocent persons. In scientific decision making, a conclusion is acceptable when we have made reasonable efforts to show that we have quantified the possibility of a different possible conclusion as sufficiently low.

For those probabilistic results that are derived from a theoretical probability model – calculated from only those assumptions that can be subject to mathematical or logical proof – possible conclusions may involve the theory that underlies the scientific test. For example: the analytic theory underlying the polygraph test is that recordable reactions to different types of test stimuli are loaded as a function of deception or truth-telling in response to the investigation target issue. In this case an explanation of the meaning of a statistically significant polygraph test result might include a description of the probability that the observed data would occur if the undying theory were false – that the observed responses to different types of test stimuli are not loaded systematically but have occurred due to random chance alone.

Polygraph test results that are calculated using Bayesian methods, involving a declarative prior assumption of the probability of deception or truth-telling, can be used to mathematically update or modify the prior probability. The result will be a posterior probability of deception or truth-telling. In addition to their well-established usefulness, results from Bayesian method can also offer an advantage in that the conceptual explanations of these results are sometimes be more easily or intuitively understood by persons not trained in frequentist inference.

Categorical test result based on the probabilistic test result

All scientific tests are fundamentally probabilistic. However, professionals who make referrals for testing, and others who receive the results from scientific tests, may find it more practical and convenient to work with categorical test results. In overly sim-

plistic terms, categorical test results can be thought of as either pass or fail. Categorical results in the polygraph testing context might simplistically conclude that an examinee has told a lie or the truth, but this type of over-simplification can tend to encourage a number of overly concrete and problematic expectations for determinism or deterministic perfection. In response to both intuition and objective evidence that polygraph results are in fact non-deterministic (i.e., they are probabilistic), polygraph professionals have adopted a practice of making categorical interpretations of diagnostic polygraph test results using the abstracted terms deception indicated or no deception indicated. The terms significant reactions or no significant reactions have also been used, and these abstracted terms are equivalent to the former for practical purposes. All of these terms are contextual analogs of the more abstracted scientific terms positive or negative that are used to avoid the imposition of personal values onto conclusions that expected to refrain from subjective judgement. Most importantly, the use of abstracted categorical descriptors for scientific polygraph test results will remind readers of test reports that test results are a probabilistic conclusion and not a form of deterministic observation or direct physical measurement.

Regardless of whether described using concrete, contextualized, or abstracted categorical labels, the practical meaning of different categorical description schemes remains the same. Categorical results of scientific polygraph tests are based on probabilistic results that are based on a structured and reproducible analysis of the test data. The polygraph report should contain a clearly stated conclusion about whether a categorical conclusion is supported by the probabilistic analysis of the test data. In this way, overly-simplistic impulses to suggest that the polygraph can measure or detect lies directly – or to criticize the polygraph test because it cannot – can be either avoided or more easily recognized as a fallacious and unrealistic attempt to impose concrete expectations onto a probabilistic test result.

Summary of information for a reproducible polygraph examination result

Reproducible results of scientific polygraph tests are those that are based on data analytic methods that are published and available, and for which sufficient detail is included in the test report so that others who repeat the analysis of the test data are likely to reach the same analytic conclusion. Adequate detail in a scientific test report will also enable an intelligent and educated reader to have reasonable answers to the common types of questions that are asked about scientific test results. A complete and adequate polygraph examination report will serve to document both the analytic conclusion and the analysis parameters (i.e., reference, model, assumptions, probability cut-scores, use of statistical corrections) that were used to achieve that conclusion. A written report of a scientific polygraph examination result should include the following pieces of information:

- Name and description of the validated method used to score and interpret the test result
- Numerical test scores and cut-scores if these are used in the classification of the

test result
- Probabilistic test results and probability cut-scores including statistical error corrections
- A statement of interpretation to describe the practical meaning of the probability results
- A statement describing the categorical test result that is supported by the analysis

In addition to the test result, a complete polygraph examination report will also include information about the examinee, referral question or reason for testing, information from the polygraph interview, test questions and answers, testing and recording instrumentation. Inclusion of probabilistic information in polygraph examination reports will foster realistic expectations and reduce confusion and aggravation when people are reminded that the results of polygraph examinations are imperfect. If probability results are not clearly explained in written polygraph examination reports, it will serve only to encourage perceptions that the polygraph result is not an objective analytic process and is instead subjective – which introduces a vulnerability to subjective human bias. Subjective test results will be inherently less reliable, and therefore less desirable, than objective test results.

Much of this suggested information is already provided in many written reports of polygraph examination results. The inclusion of statistical information served to anchor and account for the probabilistic test result, and to improve the reproducibility of the analytic conclusion. Some field examiners, and perhaps some referring professionals, may be uncomfortable reporting and discussion scientific and statistical test results. Or perhaps there is discomfort in the acknowledgment of reality when discussing the potential for testing errors. Regardless, polygraph examiners who wish to claim professional expertise beyond that of an operator or technician will be obligated to learn to communicate the principles of science and scientific testing, including probability and decision theory.

A complete and satisfactory polygraph report should also include information about any constraints or limitations that might be regarded as capable of limiting the strength of the analytic, probabilistic and categorical conclusions. Such limitations might involve the medical, developmental, or psychological status of the examinee, or might involve the interpretable quality of the test data. Additionally, polygraph professionals should be careful not to misrepresent or abuse statistical concepts. For example: p-values are not an effect size. That is to say, p-values do not describe the magnitude of the observed result or level of confidence in the result. P-values only describe the expected proportion of times we expect to observe a similar test result if another data distribution is a more correct representation of the population from which the examinee is from. P-values are sometimes thought of as a probability of error when we select a particular explanation or conclusion, and reject another possible conclusion or explanation, but are formally defined as the proportion of times we would expect to see a similar experimental result under numerous repetitions of an experiment if a (usually null) hypothesis is a true or reasonable description of reality.

Discussion

Scientific test results, like the results of scientific studies, should be reported in sufficient detail such that another professional can fully understand the test result, and can reproduce the analysis and result if desired. Although not a complete verification of a test result, reproducibility of polygraph test results is fundamental to other forms of verification and validity, including quality assurance activities. Analytic results of a scientific polygraph test cannot be assumed to be a valid indicator of reality if they are not at first reproducible. In other words, if different professionals analyze the data to different conclusions using the same analysis methods, then the results cannot reasonably be argued to be a reliable indicator of reality. If unreproducible test results happen to also be correct then it is convenient, but it is nonetheless spurious.

When examiners neglect to include scientific and statistical explanations in polygraph test reports the effect will be an impulse to establish the validity of a test via social mechanisms such as bravado, simplistic hyperbole, expertizing, and credentializing. These are neither synonyms nor substitutes for scientific validity, and the result can be overconfidence in a uncertain conclusion. Inclusion of statistical information in the reports of scientific polygraph exams will remind all professionals that scientific test results will always include some potential for error. Moreover, inclusion of statistical information, regarding both test scores and probability cut-scores, will provide continuous reinforcement of professional and social knowledge that the polygraph test is a scientific test – that test results are not derived through an over-simplified social heuristic that may or may not be supported by scientific evidence but which would undoubtedly be more subjective than quantitative analysis of recorded test data.

Reinforcement of probabilistic awareness and the concepts of science and testing

The most important reasons to provide scientific and probabilistic information in polygraph test reports is to reinforce a continuous awareness of the fact that scientific tests are fundamentally probabilistic and are not expected to be perfect. Without this type of continuous awareness the polygraph profession is likely to experience continued unwarranted frustration over the fact that results of scientific polygraph tests, like the results of all other scientific tests, are imperfect conclusions regarding some amorphous phenomena that cannot be evaluated through simple and perfect deterministic observation nor through direct physical/linear measurement. The polygraph profession is also likely to continue to observe a continuous stream of individuals who attempt to assert the fallacious argument that the polygraph is invalid or unscientific simply because there is no single physiological, behavioral, or psychological phenomena that can be exploited to achieve perfect deterministic observation or direct physical measurement of deception or truth-telling, with the implication, either explicit or covert, that the polygraph is merely a bogus pipeline prop. These false and unrealistic arguments neglect the fact that virtually all scientific tests work as a function of some combination of proxy signals that are correlated with an amorphous phenomena of interest but are not them-

selves the phenomena of interest.

Scientific tests are expected only to quantify the margin of uncertainty surrounding a result or conclusion and to help constrain decision errors to within identified tolerances. Tests are expected to do this in a structured and objective manner for which the results can be demonstrated to be reliable. Reliability is established, in part, when repeating a testing procedure or analysis can be expected to reproduce a probabilistic result within some desired tolerance for uncertainty.

Reliability of test results

Test reliability, related to the reproducibility of analytic results, has been a discussion topic for many years among those who develop and use scientific tests, and has included discussions about re-test reliability and inter-scorer reliability. Questions about re-test reliability will address whether the acquisition and analysis of new testing data from the same individual will lead to the same analytic conclusion. In contrast, questions about inter-scorer reliability will address the degree to which or likelihood that different professionals will achieve the same conclusion using data from a single test administration. Similar to the way that the reliability of test will constrain the validity of a test (i.e., a test cannot be valid if it is not reliable), re-test reliability will be constrained by inter-scorer reliability (i.e., we cannot expect to consistently achieve the same result upon re-testing if we cannot first achieve the same result when different professionals analyze the same data). Increased availability of automated computer algorithms for analyzing polygraph test data could potentially reduce or even eliminate concerns about inter-scorer reliability. Results of automated computer algorithms are inherently reproducible as long as the input data and the analysis parameters are known. A satisfactory polygraph examination report will serve to document the analysis parameters regardless of whether the analysis was completed through a manual scoring rubric or automated computer algorithm.

Computerized algorithms can be useful not only to analyze test data, but also to produce satisfactory and complete reports of the test data analysis and test result. Use of computerized analytic and reporting tools can serve to increase the consistency and correctness of the content of scientific polygraph examination reports, and can also serve to disseminate correct information to polygraph professionals and others who are interested in understanding the meaning of scientific test results. The automated re-scoring reliability that computer scoring algorithms can provide could enable test development and validation efforts to be more fully devoted to the refinement of testing procedures to improve the consistency of test administration and corresponding re-test reliability.

For some forms of testing manual and automated analysis protocols may produce the same statistical or probabilistic results. For example: a psychometric test involving forced choice or multiple choice test items can produce the same numerical and probability scores regardless of whether machine scored or manually scored. Because manual analysis of polygraph time-series data has been traditionally accomplished through

subjective visual extraction of physiological reactions, polygraph results based on protocols for manual test data analysis have remained inherently less reliable than the results from automated computational machines. This places inevitable limits on the level of validity that can be achieved, simply because a test can be no more valid than it is reliable. A testing procedure and analysis method must first produce reliable outcomes; then we can begin to make realistic conclusions about the relationship between those outcomes and reality.

One fact bears reminding at this point: continued use of manual analysis methods for polygraph test data may become increasingly problematic in the future. This is because we are likely to continue to observe the development and commercialization of more objective automated scientific technologies for lie detection and credibility. The polygraph will be at risk for becoming an anachronism if the it continues to rely on subjective visual feature extraction and simplified heuristics while neglecting the development and use of more objective and more powerful statistical and computational methods. Algorithms based on statistical and machine learning principles should be more fully developed and integrated into the analysis of polygraph results. Until algorithm or machine scoring methods are more completely integrated into field practices, the continued use of manual test data analysis protocols will mean that analytic results of polygraph test data will remain subjective and more highly variable than necessary. Increased use of automation and computerized analysis methods can increase the reliability of polygraph outcomes, and can ultimately lead to increases in test accuracy.

Arguments against the inclusion of analytic details in polygraph test reports

Inclusion of probability information into polygraph reports is an explicit acknowledgement regarding the unavoidable reality that there is some potential that a test result is incorrect. Some may be uncomfortable with this acknowledgment. Others may suggest that readers of polygraph examination reports are not intelligent enough or have no interest in scientific or probabilistic test results. Still others may attempt to assert that readers of polygraph examinations report are not entitled to probabilistic and scientific information regarding the test results. These views are both shortsighted and parochial. Neglecting to include scientific and probabilistic information in polygraph reports appears to underly some of the most common criticisms against the polygraph: that the test may be overly subjective and unscientific, or may be based on pretense that is not consistent with reality, or may have some fundamental scientific flaw to conceal. Another common criticism – based on a misguided expectation for perfection – is simply that the polygraph should not be used because it is not infallible. This criticism neglects the fact that scientific tests are not expected to be perfect or infallible, and neglects the very purpose of scientific tests: to quantify in some replicable manner a phenomena that cannot be subjected to perfect deterministic observation or direct physical/linear measurement. All scientific test results should be reported in a manner that reflects the fact that the purpose of any scientific test is to quantify the margin of uncertainty surrounding a

conclusion about some amorphous phenomena.

All test results are fundamentally probability statements. In the case of a scientific polygraph test, the test results might be understood as meaning that the examinee is probably truthful or probably deceptive. Because they are probabilistic and non-deterministic, all polygraph results are at once probably correct and also probably incorrect. Ideally, the probability is high that the results are correct, with a corresponding low probability that the results are incorrect. Polygraph examiners are permitted to render a professional conclusion about deception or truth-telling when they acquire data and can provide a reproducible analysis showing that these probabilities satisfy our stated needs or are within established probability tolerances. For convenience and simplicity, field examiners and other professionals will, to some degree, omit the "probably" part when writing and communicating polygraph test results. When there is an absence of probabilistic discussion and awareness surrounding polygraph examination results it will tend to encourage misguided expectations for perfection, and this will lead to unnecessary and avoidable misunderstanding and frustration when a testing error inevitably occurs.

When people persist in maintaining naive and unrealistic misconceptions that polygraph test results should be infallible, they may attempt to incorrectly attribute all test errors to one of two causes: deficiencies in competency on the part of the examiner competency, and faking or subterfuge on the part of the examinee. The subscript of these attributions is a persistence of misguided and naive expectations that the polygraph test itself can be regarded as perfect if given a competent examiner and cooperative examinee – despite the fact that the requirements for competence and cooperation merely illustrate the potential for test imperfection. To the degree that procedural error is a cause of testing error, automation of testing procedures can be expected to reduce testing errors. However, it would be unwise to expect that a completely automated test administration would mean that all errors could be attributed to faking. A small proportion of errors can be expected even in the absence of procedural error and faking. Scientific tests are expected to provide an estimate of this potential.

The pervasiveness of expectations for perfection has been such that polygraph professionals have sometimes routinely endorsed these expectations, either deliberately or inadvertently, with statements such as "I don't like to ever get beat." Embedded in this type of statement is a message that the polygraph may be more interpersonal contest (mano-a-mano) than scientific test. This sentiment is premised on the false assumption that a probabilistic test will somehow never produce an erroneous result if it is simply administered correctly, and also implies that a testing errors can only be the result of the superiority and inferiority of the participation of those involved. These perceptions neglect the reality that the results of valid scientific test results are usually correct because of probability theory, and the corresponding reality that the results of any valid scientific test might be incorrect on some smaller proportion of occasions due to random chance alone. In other words, a quantifiable proportion of testing errors

can be expected to occur regardless of any procedural error and regardless of sophisticated subterfuge. Perhaps equally problematic, this type of statement encourages polygraph professionals to personalize (i.e., to view it as a personal failing) the existence of occasionally unavoidable testing errors. Once again, scientific tests are not expected to be perfect, and are intend only to quantify the probability of decision error so that we can reduce its occurrence to within stated tolerances. Research on test validity and test accuracy are expected to answer this: given that all test are imperfect, and given that some persons attempt to alter the test outcome, and given that there is some variability in functional characteristics, what proportion of test results can be expected to be correct or incorrect?

If the polygraph is a form of interpersonal contest and not a scientific test then polygraph professionals would be wise to divest themselves of the burden and constraints of scientific expectations and the need to understand abstractions such as alternatives and probabilities. Instead, if the polygraph is an unscientific interpersonal contest then polygraph professionals should merely train and develop themselves to win the contest, and the polygraph profession should be designed to filter and consume professional talent in the same manner as professional baseball. Major League Baseball teams are surrounded by layers of infrastructure intended to identify and develop human talent beginning at young ages and continuing through the various ranks of amature, semi-professional, minor-league and major league environments. But even the world of professional baseball has resorted to quantification and probabilistic modeling as a way to optimize desired performance outcomes. Today it would seem today that the principles of science and probability theory are used virtually everywhere. If the polygraph is a scientific test and not merely an interpersonal contest, or a bogus-pipeline prop, and then it should be an obligation for polygraph professionals to learn to think probabilistically and to communicate probabilistic test results as well as categorical test results.

Regardless of whether the polygraph is a personal contest or scientific test, the bottom line is that expectations for perfection cannot coexist with reality and cannot coexist with science. This is because science is about trying to understand reality. Because humans probably cannot ever achieve perfection and probably cannot ever know everything about reality and the universe, in the end, due to these fundamental human limitations, all forms of human knowledge and all conclusions are an approximation of reality. The goal of science and scientific testing – including scientific polygraph testing – is not to pretend perfection or infallibility where these are not possible, and not to pretend that we know the universe or reality with certainty. The goal of science, including scientific polygraph testing, is merely to solidify our probabilistic approximation of reality in a manner that is less likely to be spurious and more likely to be reproducible or replicable.

Conclusion

Professional opinions and professional conclusions are those that are based on data and evidence. Professional opinions that are not based on structured and replicable

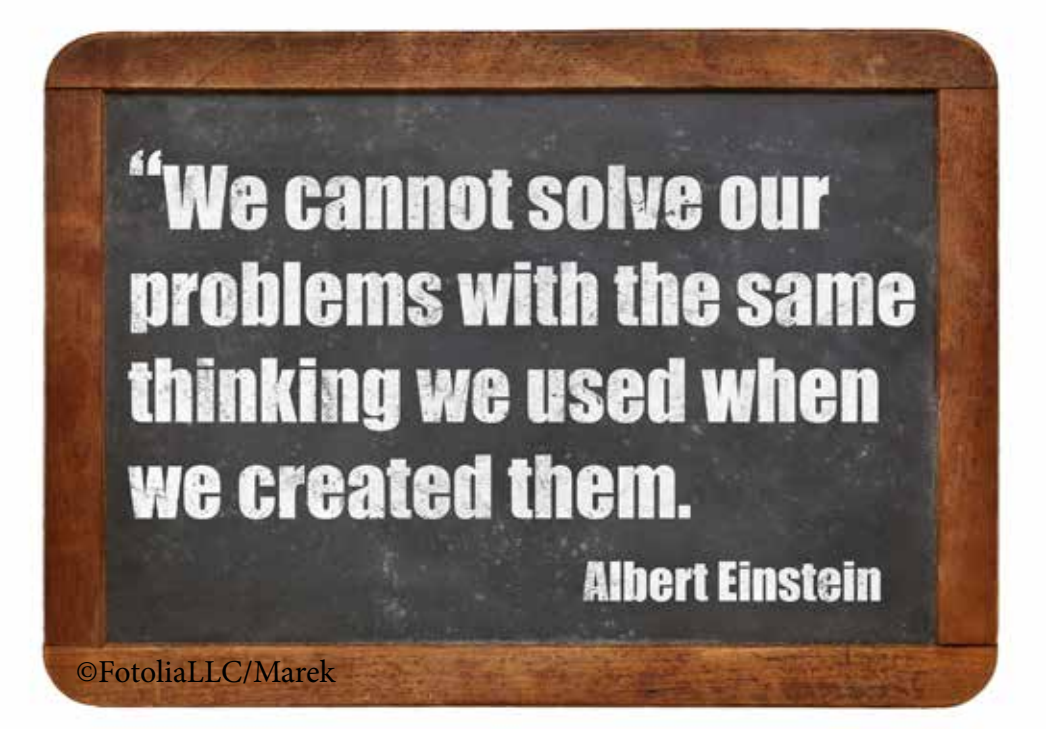
evidence are clinical methods for which the validity depends heavily on the subjective persona of the professional. Subjectivity will mean that the conclusions are wide open to alternative subjective interpretations by other professionals. When different professionals do not agree, and when there is a need for a conclusion, the ultimate determination will depend on a form of contest oriented around the social popularity or political weight of the different professionals. This form of decision making will be at risk for becoming disconnected from any form of reality that can later be supported by evidence. Criticisms that the polygraph is unscientific and invalid may begin to carry deserved weight if polygraph results are not reproducible and reliable.

Effectively written polygraph examination reports will served to correctly describe the scientific basis for the polygraph test. Effective communication of test results will educate others about how to understand both the practical and probabilistic meaning of scientific test results, and will inoculate readers against naive and unrealistic expectations that polygraph accuracy rates should somehow approach deterministic perfection. Whereas some fields of science and technology will involve margins of error and uncertainty that are so small that some individual professionals may never witness a failure, the ratios of diagnostic and error variance in the areas of human psychology and human physiology are such that many professionals can expect to occasionally confront the possibility of observing or experiencing a testing error. When errors can be reasonably expected to occur and when the professional culture creates an illusion that errors are intolerable, the result will be that field practitioners and others may experience a massive impulse to hide testing errors in order to ensure their own professional survival.

The manner in which polygraph professionals think about test results will both influence and be influenced by how they report test results. It will pervade and influence everything we assume, express and communicate. It will also influence everything we do in the acquisition and recording of the test data itself. Similarly, the manner in which polygraph professionals report test results will both influence how other professionals think about polygraph test results. If we want others to more clearly and more correctly understand polygraph results and the scientific based for the polygraph, then it will first be important for polygraph examiners to communicate the test results more effectively. In order to communicate the results more effectively it will be necessary for polygraph professionals to learn to think more clearly about the scientific and probabilistic basis of the test results. At times, the ability to clarify or improve our thinking is contingent upon our willingness to engage in self-reflection and critical analysis of the language and logic that define our present knowledge, assumptions, and field practices.

Polygraph field examiners will help to ensure their own useful future by reporting test results in a reproducible manner, including information about the categorical and probabilistic test results, and the method and parameters for analysis. Reporting information in this way will reduce and dispel criticism that the polygraph test is unscientific and will reduce misguided and naive

expectations for deterministic perfection or physical measurement where these are not possible. The polygraph profession will be wise to plan its future on the provision of scientific test results that are based on data and evidence for which validated and structured probability models can be used to provide test results that include reproducible estimates of the potential for testing error. Of course, if the polygraph were infallible then there would be no need to account for any potential for testing error.



©FotoliaLLC/Marek