sciendo

EUROPEAN
POLYGRAPH
PUBLISHED SEMI-ANNUALLY
2022 VOLUME 16 NUMBER 2 (56)

DOI: 10.2478/EP-2022-0008

Thoughts on the Inconclusive Zone in Comparison Question Test (CQT)

Avital Ginton[1]

AV-GN, polygraph, applications & research[2], Tel-Aviv, Israel

Abstract

The Inconclusive Zone, in one form or another, is an integral part of almost any data analysis method of CQT. It represents the existence of built-in uncertainties in the test and indicates how the test, in general, and the examiners, in particular, bear this limitation in practice. Unfortunately, it is so common that most examiners tend not to spend too much time and attention wondering about its meaning. The following are some reflections on the concept of an Inconclusive Zone, and its actual use, including some resulting recommendations. That includes internal aspects of the test, such as the relation between the extent of using multiple comparison points between relevant and comparison questions and the optimal Inconclusive

[1] Avital Ginton, Ph.D., a psychologist, past head of the Israel Police Polygraph section (1977–1994), avital.ginton@gmail.com.

[2] AV-GN, polygraph, applications & research is a private independent institute. The author also teaches part time in Tel Aviv University – The Buchmann Faculty of Law (c/o Prof. S. Shoham), Israel. Email: ginton@zahav.net.il Alternative Email: avital.ginton@gmail.com

Zone size. External aspects affecting the Inconclusive Zone's size and shape are also stressed when looking at the CQT through a prism of aiding decision-making rather than a mere means to sort truth from lies.

Key words: Comparison Question Test, CQT, Inconclusive Zone, Polygraph, True score theory, Pay-off matrix in polygraph decisions

Assuming that polygraph subjects know whether they are telling the truth or lying to the relevant questions, it is clear that an inconclusive result does not reflect their state of mind regarding the truthfulness of their answers to the relevant questions. Rather, it reflects the Comparison Question Test (CQT) method (Krapohl & Shaw, 2015) or its execution failure to identify the subjects' state of mind in that respect. More precisely, it is a failure to reach a definite conclusion about that, whether true or false.

In the CQT, the primary rule in deciding whether the examinee is found to be deceptive or truthful to the relevant questions states that if the magnitude of the response to the Relevant Questions is found to be stronger than to the Comparison questions, namely, **R >C** – the subjects are deemed Deceptive (DI) and if the opposite is detected, i.e., **C >R,** they are considered truthful (NDI). Thus, in the CQT, the indicator for lying or telling the truth to the relevant questions is the difference between the magnitude of physiological responses to the two categories of questions. The rationale behind these decision rules is that the response magnitude is highly correlated with the amount of concern from the questions. And while deceptive subjects are mostly concerned about the relevant questions, the opposite is found among truthful examinees (Vrij, 2008: 302)[3].

However, due to the ongoing fluctuations in the states of mind, cognitively and emotionally, and in the body's physiological activity, any two measuring of responses even to the same objective stimuli (i.e., questions) may subjectively be responses to somewhat different inputs that induce different output intensities (e.g., the same question might be detected as more or less intimidating in different repetitions)[4].

[3] This rationale has been critically questioned by many scientists in the academic world (Vrij, 2008: 304–311) but the present paper is not the place to argue on that.

[4] Physiological response to psychological stimulus (e.g., a question) is influenced by the psychological state of the individual at the moment when the stimulus is "thrown" at them. How focused they are? How much are they self-absorbed or lost in thought at that moment? To what degree they are aroused or excited at that particular minute? And so forth. All these are matters that by nature contain some sort of lability resulting in mental baseline fluctuations which modify the way

These fluctuations are considered background noise that obstructs the detection of the pertinent signal. One characteristic of the noise is its masking quality that interferes with identifying the signal. But once we manage to detect what seems to be a signal, we still should wonder whether this detected signal is enough to conclude about the truthfulness of the examinee regarding the relevant questions.

The theoretical rationale of CQT states very clearly the expected direction of the differences between R and C questions; however, it says nothing about the size of these expected differences in the magnitude of reactions.

Why is it then that we are looking for a certain minimum in the magnitude of the difference before rendering a decision as manifested in applying an Inconclusive Zone? Somehow we assume, probably justifiably, that slight differences due to incidental noise might

frequently occur, as opposed to big differences, which are very rare to occur due to random fluctuations per-se. Hence we want to see a big enough difference before attaching any significance to it. While that might seem quite logical, one should keep in mind that to avoid relying on an observed difference that occurred just by random fluctuations or other kinds of non-systematic noise, a consistency factor is much more important than the size of the difference (Ginton, 2013).

To reduce the subjectivity of estimating differences in response magnitudes, a technique attributed to Cleve Backster (1962, 1963), who termed it Numerical Scoring, was developed. For those unfamiliar with the numerical scoring technique, the following is its skeletal description: It is based on comparing the detected physiological responses to a Relevant question each time it is presented with those of an adjunct Comparison question. This is done per each physiological measure, which usually consists of Respiratory, Cardiovascular, and Electrodermal activities. Each comparison is considered a separate comparison point, and the difference in the response magnitude in each point gets a number

---

the individual experiences the stimuli. Therefore, even if we objectively throw the same stimulus (a specific question) at the subject twice, they may experience them differently (as long as we assume fluctuations in his momentary mental state) and react with different physiological responses. To that, we should add the fact that the mere body activity is constantly in flux, and the physiological response is always built on top of the physiological state of the responder at the said moment. The bottom line is that there is a high probability that two stimuli that are objectively considered equals create different reactions. Because of this potential variation, it is necessary to repeat the stimuli several times to pick the commonality and reject the momentary differences.

indicating the difference in the response strength between the compared questions in that particular point. If there is no substantial difference, the number is Zero; if the Relevant response is stronger than the Comparison, it gets a Negative number, and a Positive number is given to a comparison point, indicating a more robust response to the Comparison question relative to the Relevant. Originally the scoring technique applied a seven points scale from −3 to +3, indicating the degree of difference found in each comparison point and its direction. However, later a three-point scale indicating only the direction of the difference was also adopted, i.e., −1, 0, +1. The numbers from all the comparison points, including repetitions of questions, add to a total score across repetitions of each relevant question and, when applicable, across all the relevant questions together. Based on prior decided positive and negative cut scores[5], the examiner could determine whether the difference between the Response's magnitude to R and C questions is big enough to make a call. A Negative final score that reaches the cutpoint means a Deceptive examinee, and a Positive score indicates a Truthful examinee. The range between the two cut scores (the negative and the positive) is an Inconclusive Zone, meaning that getting a final score in this range is not enough to make a call. Thus, the moment of birth of the Inconclusive Zone was when the decision rules in numerical scoring set two separate cut scores, one as the threshold for a DI call and the other as the threshold for an NDI ruling. This does not mean that with smaller differences, the rate of correct decisions that would have been made if these were determined by the mere direction of the difference is similar to flipping a coin. It should still be above random success; However, the estimated error rate may be too high to be used as a practical method of distinguishing lies from truths. For instance, Ginton (2012) showed that relying on a decision criterion that did not leave any undecided test, which practically means not having an Inconclusive Zone, resulted in an error rate of twenty percent. Is that still good enough to be used for decision-making, or is it too costly? That is a matter of policy.

The Inconclusive Zone has been an integral part of any numerical scoring technique and, with some needed adaptation, also part of many computerized data analysis algorithms. Hence, discussing various aspects of this practice seems essential to the polygraph profession, as presented in the following reflections on that issue.

---

[5] Based on theoretical reasons, on empirical research or just arbitrarily.

The Inconclusive Zone and the True Score Theory

As already mentioned, the CQT is based on the assumption that deceptive examinees are more concerned about the relevant questions than the comparison questions. In contrast, the opposite is true for truthful examinees, and these differential concerns are manifested in the relative magnitude of the physiological reactions. The differences found between the magnitude of the physiological responses that accompany the relevant questions and those that appear in the comparison questions indicate the veracity of the examinee about the relevant questions but does not mean that the veracity of the subject's answers to the relevant questions is changing with any change of the detected difference. The subject's veracity stays the same because it does not depend on the test but on the subjective cognition of the examinee. What is changing is the degree one can rely on the detected difference to reflect the truthfulness of the examinee's answers to the relevant questions. Looking for essential consistency in that matter means repeating the measuring of the difference in reactions between R and C questions to determine whether the detected difference keeps pointing in the same direction.

In the field of test theories, consistency is closely related to reliability. If a measurement device or measurement procedure consistently assigns the same score to individuals or objects with equal values, they are considered reliable, but if the scores assigned to the same people or objects vary in repeated measurements in the absence of any known and understandable reason, the device or procedure is considered unreliable. Reliability is the extent to which a test or any measuring procedure yields the same results on repeated measures or trials or the extent to which a repeated test yields consistent outcomes and scores. Thus, reliability can be defined by the consistency of a measurement procedure and its outcomes. Reliability is the extent to which a test or any measuring procedure yields the same results on repeated measures or trials or the proportion to which a repeated test yields consistent outcomes and scores. Thus, reliability can be defined by the consistency of a measurement procedure and its outcomes. In CQT (and other polygraph techniques), consistency is not just a desirable option to look for but a critical factor for decision-making. We want to ensure that the outcomes we get on the test using our measuring procedure represent the examinees' actual mental states regarding their relative concerns about the relevant versus the comparison questions and not mere reflections of random fluctuations or other nonstable irrelevant factors.

The desire to differentiate the true state of a person who takes the test in the variable we are interested in measuring, from non-relevant influences that do exist in any test can be found in the "True Score Theory" (Lord, Novick & Birnbaum, 1968; Trochim, 2000) that, in its simplified version, when adapted to our field, states that – the measured physiological responses to any question and the measured difference between the strength of the responses to relevant (R) vs. comparison (C) questions, are a combination of "True score" plus "Error." That is to say that, in principle, the output itself may always contain irrelevant aspects, and the measurement device or the act of measuring might add another irrelevant value, namely Measurement Error. Thus the polygraph test variance includes true variability in the valence held by the examinees plus variance that reflects the effect of irrelevant factors[6]. For example, a person's test score might deviate from the true score because they are sick, anxious, in a noisy room, inattentive, got stomach cramps, are distracted by the examiner's inappropriate behavior, etc.

A significant portion of these irrelevant factors is the existence of random fluctuations, for instance, in the momentary level of the examinee's attention capacity or focus and in the pattern and the strength of the psycho-physiological reactions. These fluctuations might mask the true values and interfere with making the right decision. The need to nullify or at least to weaken the effect of these fluctuations is crucial for any test, and a polygraph test is no exception. Repeating the questions and looking for consistency over the repetitions is the primary way to achieve that.

The logic behind this tactic is that random fluctuations tend to nullify themselves in repeated measures and enable the true score to surface. The degree to which a particular technique or test is immune to the effects of such fluctuations on the outcomes is defined by the Reliability index that takes values between 0.0 to 1.0, indicating the extent to which an assessment tool or procedure is consistent, i.e., free from random measurement error. The closer the computed index is to 1.0, the higher the test output's consistency and the lower the noise's final effect.

To nullify or reduce the weight of the random errors factor that masks the true score takes repeating the measurements, and as long as its effects are presumed to exist, the Inconclusive Zone assumes the role of being the primary buffer from its adverse effects.

---

[6] It is important to understand that the variable here is not deception vs. truth-telling rather it is the difference between the reactions to the relevant and the comparison questions that supposed to reflect the difference in concerns between the questions. This difference can take various values which are the combination of true score (i.e. reflecting the amount of concern) plus error. There are two kinds of errors, random and systematic. Striving for consistency deals with the effort to reduce random error but can not remedy the effect of possible systematic error.

Having stressed the crucial role of repetitions, one may wonder about the number of repetitions needed to gain confidence that the outcome is reliable. Most techniques would ask for three repeated tests or three charts. Is that enough? Given the fact that in numerical scoring, as for now, it is customary to give the same weight to whichever repetition and every comparison point[7], using two charts with three relevant questions equals three charts with two relevant questions (in a single issue, single facet examination). In light of this, one may wonder whether this equality represents a latent unintended and surely undeclared attitude toward question repetitions which is rooted deeply in qualitative considerations far beyond the numerical scoring technique.

The role of having multiple comparison points

Multiple comparison points help expose relevant factor values and eliminate non-relevant ones, such as random fluctuations, or possible systematic irrelevant noise effects, such as the order by which the questions are presented.

By repeating the measuring of responses to the repeated stimuli, the random fluctuations of the noise tend to cancel each other and change the order of questions between repetitions to nullify possible systematic irrelevant noise. In contrast, the pertinent signals, i.e., the values of the responses to the "pure" stimulus that usually points to the same direction, surfaced above the noise. That is the theoretical reason for repeating the questions several times. Whenever the number of measurements is small, the relative effect of the noise on the variance or, in other words, the weight of the noise effect in the measured activity relative to the pertinent signal tends to be pretty strong, so there is a need to look for robust signals to be sure that what is detected is a pertinent signal and not a random fluctuation of the noise. However, when applying more measuring points, the noise effect is reduced, and even a relatively weak signal can surface above the noise and be detected. This should affect the size of the Inconclusive Zone. The smaller the number of comparison points, the bigger the minimum signal value to which significance can be attached, which means an increased Inconclusive Zone. Whereas having numerous comparison points, one can be satisfied with a much narrower Inconclusive Zone since after nullifying most of the random noise, even a moderate signal is enough to be consequential. This seems to be at odds with what some polygraph experts think, as they indirectly claimed that when using fewer comparison points (e.g., two relevant

[7] Some scoring methods, e.g., ESS or OSS3, double the EDA weight.

questions instead of three relevant ones), a narrower Inconclusive Zone is enough to distinguish DI from NDI calls because the maximum possible score is potentially smaller than the maximum possible score when more comparison points are used. That has falsely led them to opposite conclusions concerning the required size of the Inconclusive Zone for preventing errors due to random or non-random noise[8].

Exact repetitions of stimuli are the best technique to eliminate random noise since the noise tends to nullify itself due to the random quality of its changing directions and intensities. Ideally, in the end, we are left with the non-random effects that tend to pile up each time it is measured. However, there are also non-random irrelevant stimuli and responses that grow to build up too with repetitions and can add or subtract their values to the pertinent signals depending on their direction. There are several kinds of non-random irrelevant stimuli that we would like to reduce their impact. One type encompasses some characteristics embedded in the relevant questions that prompt concern in the truthful examinees from reasons other than being deceptive, such as being ashamed or embarrassed by the way the questions have been phrased or concerned about possible public exposure of the questions. Another type is emotionally loaded characteristics that do not induce concern but do activate the physiological system pretty much the same, such as mentioning in the questions specific names or circumstances. A third type is known as an outside-issue matter which mostly means that something outside the scope of the test or the Relevant and Comparison questions spheres bothers the examinees, not allowing them to direct enough attention to the R or C questions. Then, of course, there is always the general non-specific tension induced by the testing situation and its procedure and the concern of being diagnosed mistakenly. This does not exhaust all possibilities of directly affecting the examinees, but for now, I would like to enlighten another angle of possible effects of irrelevant noises. That type of noise affects the examiners, and only through them, occasionally, might roll onto the examinees' responses. An example of that can be the contamination by prior information about the guilt or innocence of the examinee.

The best way to lower the adverse effects of such non-random noise is by building a controlled testing atmosphere and procedure that prevents them from showing up. And in case they do show up, to speed the pace of their habituation with repetitions. Unfortunately, speeding up habituations of non-random irrelevant contamination without affecting the relevant stimuli is a matter that needs some sophisti-

[8] Personal written communication with anonymous reviewers and the chief editor of the Polygraph & Forensic Credibility Assessment: A Journal of Science and Field Practice, 2022.

cation that is beyond the material learned in polygraph schools and may contradict the policy of moving towards stiff standardization across cases, circumstances, and individual differences between examinees.

Unfortunately, even applying all these precautions is never enough to ensure that the effects of the irrelevant factors have been nullified. We should assume that the final measure contains a bit of random noise and, more dangerously, some non-random interference. Thus, the reason for applying an Inconclusive Zone is twofold. First, to buffer the random fluctuation remains and second, to function as a safeguard against the adverse effect of non-random variables hoping that we have managed to weaken their impact to such a low level that they would not cross the Inconclusive Zone barrier. Of course, there is nothing sacred in any size or shape of an Inconclusive Zone, but we should follow certain logic and steps in opting for the optimal zone.

The pay-off matrix of the inconclusive rate vs. errors and the optimal Inconclusive Zone

The general role of the Inconclusive Zone is to prevent wrong decisions. However, it comes with an inherent cost tag of reducing the number of correct decisions. Almost by definition, there is a negative correlation between the size of the Inconclusive Zone and the number of definite calls made by the test. The number of both kinds of decisions, whether wrong or correct, decreased whenever the range of the Inconclusive Zone increased. So, a wise approach is to think through a pay-off matrix and look for the optimal zone breadth. However, the optimum is not a fixed value; what might be regarded as optimal in one circumstance might not suit a different situation. Sticking to a rigid size of an Inconclusive Zone is like saying that all decisions have similar importance and consequences. Ideally, it should be left to the clients – who use the polygraph as a means to aid their decision-making to decide directly or undirectly how big the Inconclusive Zone should be. The case of homeland security is quite different from catching a thief, and so is the difference between sexual harassment and cheating on a school examination. This is a matter of social and institutional preferences based on economic, moral, and ethical considerations. Still, an evidence-based approach is essential to enable reasonable assessments of the proper range of the Inconclusive Zone. Alas, due to the lack of data, it is inconceivable to expect evidence to support such a differential approach with high resolution. That leaves us with a less-than-perfect choice to base the decision about the appropriate size of the Inconclusive Zone mostly on gross estimation developed from insufficient data and conventions stemming from reason and intuition rather than factual data.

For years under the influence of a variety of American government and law enforcement polygraph institutions, supported by the Utah group led by David Raskin, the Inconclusive Zone was symmetrical around Zero though there were indications that probably an asymmetrical zone would function better as was recommended by Backster (1963), (e.g. Elaad, 1999; Krapohl & McManus, 1999). As for now, some polygraph formats recommends symmetrical Zone, whereas other have adopted asymmetrical Zone based on research and some pay-off matrix considerations (Krapohl & Shaw, 2015).

Usually, the pay-off matrix, in its beneficial side, deals with improving the accuracy in polygraph testing, i.e., increasing the percent of correct decisions (True Positive and True Negative). However, polygraph testing also includes the idea of fairness that describes the wish to give any examinee a fair and equal chance to prove their innocence (True Negative) and, in the same vein, not to let any examinees to have a better chance than others to pass the test if they were lying (False Negative error).

By definition, random fluctuations do not take sides. They do not incline to one side. Thus, using an asymmetrical Inconclusive Zone can not be a remedy to cope with its potentially adverse effect on decisions. Yet, it was demonstrated that using an asymmetrical Inconclusive Zone in some circumstances may improve accuracy by reducing the rate of errors (Elaad, 1999; Krapohl, 2005). So, if it is not a matter of coping with random fluctuations, it is probably a way to compensate for some weaknesses or deficiencies in the neat rationale underlying the CQT and/or its derived rules resulting in asymmetrical pressure to augment responses to one type of questions. For example, suppose it is found that, as a general tendency in some circumstances, the reactions to R questions are more robust than the comparison ones. This may violate the fairness principle against innocent subjects and increase the False Positive rate. In such a case, an asymmetrical Inconclusive Zone is meant to correct for inherent basic inclination towards one side and evenly split the expected error rates between false-positive and false-negative types. Thus, keeping the fairness principle by applying an asymmetrical Inconclusive Zone usually increases accuracy. However, this is not necessarily the result of applying an asymmetrical inconclusive zone, and sometimes it may even reduce the overall accuracy. Yet, one can adopt it as a policy if he cares to preserve the fairness principle and choose the balance between accuracy and fairness in using the CQT accordingly.

Another way to describe the importance of applying an asymmetrical Inconclusive Zone is to remember that the polygraph is not a means of discovering the truth for its own sake; instead, it is a means to improve decision-making whenever a decision is needed, and as such, the asymmetrical Inconclusive Zone can also serve policy that weighs the cost of one type of error higher than the other. For instance, the police would not like to miss a serial rapist to a greater degree than it cares about casting false suspicion on someone who, by further investigation, can prove his innocence.

Conclusions

The Inconclusive Zone is not part of Polygraph theory in general and the Comparison Question Test in particular. This assertion is valid no matter which theoretical framework of CQT is adopted. Be it the basic FFF notion with its distinct Fear branches – fear of lie detection and fear of possible consequences – that Reids used to embrace (1977) and are still used in the polygraph theoretical sphere for almost a hundred years or so[9]; The Backster's Psychological Set (1963); Ginton's Relevant Issue Gravity (RIG) (2009); Senter et al.'s Differential Salience (2010); Palmatier & Rovner's Preliminary Process Theory (2015); (all five are mentioned in Krapohl & Shaw, 2015: 201–204); Hont's Cognitive Load (2014), or any other theoretical framework.

The Inconclusive Zone does not explain anything about how it is possible to conclude from physiological measures about the state of mind or the awareness of an examinee regarding the truthfulness of their accounts. So, it is possible to ignore it and still be within the sphere that enables inference of truth and lies by measuring physiological activity. However, the Inconclusive Zone is an important practical concept in applying the test since, in any measuring situation, there is a built-in component of several kinds of errors, as has been suggested in the "True Score Theory" that asserts that any test score or measurement consisted of the "True Score" and "Error(s)" (Lord, Novick & Birnbaum, 1968; Trochim, 2000). The Inconclusive Zone aims to reduce the impact of potential errors, mainly derived from random noise and the errors derived from irrelevant systematic noise, that is left due to failure to nullify such factor by applying specific procedures and means. This is not the only means for reducing potential errors but by no means

[9] The author is not awared of any reliable document that attributes the first use of this notion in polygraph testing to any scientist or practioner.

is a primary one. It should be recognized that there is a negative relationship between the size of the Inconclusive Zone and the number of decisions (right or wrong) that can b taken. Thus, the right way to decide about the Zone's size and shape (symmetrical or asymmetrical) should follow thinking through a pay-off matrix. Finally, if one takes the polygraph as a decision-making aid and not as a "truth machine" for the sake of getting the truth per-se, it is wise to consider the non-decisive area with flexibility that follows the importance of the specific decision to be made.

References

Backster C. (1962), Methods of strengthening our polygraph technique. *Police*, 6, 61–68.

Backster C. (1963), Polygraph professionalization through technique standardization. *Law and Order*, 11, 63–64.

Elaad E. (1985), Decision Rules in Polygraph Examination. In: *Anti-terrorism, Forensic Science, Psychology in Police Investigations* (pp. 167–179). A Book of proceeding. First Published, 1985; Imprint Routledge. 2019. https://doi.org/10.4324/9780429036590.

Elaad E. (1999), The Control Question Technique: A search for improved decision rules. *Polygraph*, 28, 65–73.

Ginton A. (2009), Relevant Issue Gravity (RIG) strength – a new concept in PDD that reframes the notion of psychological set and the role of attention in CQT polygraph examinations. *Polygraph*, 38 (3), 204–201.

Ginton A. (2012), A non-standard method for estimating the accuracy of lie detection techniques demonstrated on a self-validating set of field polygraph examinations. *Psychology, Crime & Law*, 19, 577–594. https://doi.org/10.1080/1068316X.2013.765137

Ginton A. (2013), The Importance of the Consistency Factor in CQT and Other Polygraph Tests. *Polygraph*. 2013, 42, 146–162.

Honts C.R. (2014), Countermeasures and credibility assessment. In: D.C. Raskin, C.R. Honts & J.C. Kircher (eds), *Credibility assessment: Scientific research and applications* (pp. 131–158). Academic Press.

Krapohl D.J. (2005), Polygraph decision rules for evidentiary and paired-testing (Marin Protocol) applications. *Polygraph*, 34, 184–192.

Krapohl D.J. & McManus B. (1999), An objective method for manually scoring polygraph data. *Polygraph*, 29, 209–222.

Krapohl D.J. & Shaw P.K. (2015), *Fundamentals of Polygraph Practice*, Academic Press.

Lord F.M., Novick M.R. & Birnbaum A. (1968), *Statistical theories of mental Test scores*. Addison-Wesely, Oxford, England.

Matte J.A. (1996), *Forensic Psychophysiology using the polygraph*. JAM Publications, Williamsville N.Y.

Palmatier J.J. & Rovner L. (2015), Credibility assessment: Preliminary Process Theory, the polygraph process and construct validity. International, *Journal of Psychophysiology*, 95 (1), 3–13.

Raskin D.C. & Honts C.R. (2002), The comparison question test. In: M. Kleiner (ed.), *Handbook of Polygraph Testing* (pp. 1–48). Academic Press, New York.

Reid J.E.& Inbau F.B. (1977), *Truth and deception: The Polygraph ("Lie Detector") Technique.* Williams and Wilkins, Baltimore, MD.

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

Trochim W.M. (2000), *The Research Methods Knowledge Base*, 2nd edition. Atomic Dog Publishing, Cincinnati, OH.

Vrij A. (2008), Detecting lies and deceit: Pitfalls and opportunities, 2nd edition. John Wiley & Sons Ltd. New York.