

Is Matte's Inside Track the Answer to False Positives, False Negatives and Countermeasures? There is Reason to Fear that Hope is Gone.

Barry Cushman^{1*}

Abstract

The Matte Quadri-Track Zone Comparison Test (MQTZCT), a single-issue comparison question test, contains three "tracks" (CQ/RQ pairs), one of which makes the test unique: the "Inside Track" (IT). It purportedly measures an examinee's fear or hope of error regarding the relevant issue under investigation (clearly identified in the other tracks). Its developer reasons only the truthful would fear, and only the deceptive would hope for, an error, rendering the IT's inclusion in the test a remedy for false positives, false negatives and even countermeasures, resulting in a superior CQT with exceptionally high accuracy. The three known MQTZCT validation studies were reassessed by comparing accuracy with and without the IT after optimizing cut-scores for a two-track test – something never done in any of the prior studies. There were no differences in accuracy with and without the IT, resulting in a conclusion that the IT does not offer the benefits its developer asserts and the extreme accuracy previously reported is likely due to methodological flaws and unsound reasoning.

Introduction

There is no question the National Academy of Sciences' landmark report on polygraph, "The Polygraph and Lie Detection" (2003, hereafter, "NAS Report"), has had a profound impact on the polygraph community. Whether one accepts all, some or none of its findings, the report essentially contained a warning that could not be ignored: Polygraph, as a forensic science, stood at a crossroad. Practitioners had to choose whether to remain a rather loosely knit group of freelancing entrepreneurs (with examiners developing their own techniques of preference and declaring them valid based on perceptions of authority within the field) or a discipline of public interest in which

practitioners are bound by evidence-based best practices (through collective and peer-reviewed research within the scientific community). Choosing the former would come with a cost: the credibility of the profession, and potentially, its future demise. The APA chose the latter option. In 2007, it incorporated into its bylaws that as of January 1, 2012, its members would be required to utilize only those techniques that have been empirically validated (based on published research). However, that too would come with a cost: a lack of freedom in testing practices for the individual practitioner, i.e., some traditionally accepted practices would have to be subjected to more rigorous examination, and if found inadequate, abandoned.

¹ Barry Cushman is a 20+ year police veteran, the past 10 of which he has been a polygraph examiner. He is a full member of both the AAPP (where he serves as Chaplain) and the APA (where he serves on the Board of Directors, currently as President). He has published and presented on a variety of criminal justice and polygraph topics. The views and opinions expressed herein are his own and do not necessarily represent any organization with which he is affiliated. He can be reached at president@polygraph.org.

* I must include a special thank you to an anonymous peer-reviewer(s) for some very thoughtful and thorough suggestions and comments, many of which have been directly incorporated into this article and are deserving of more credit than I can offer. I am extremely grateful for the contributions.

As 2012 approached, and on the heels of similar warnings to other forensic sciences (NRC, 2009), the APA published its Meta-analytic Survey of Criterion Accuracy of Validated Polygraph Techniques (2011), which listed those techniques known to meet the APA's minimum validity standards. The report includes, among other things, a description of how the committee selected or rejected studies for inclusion, accuracy profiles for the techniques it reviewed, and information on techniques for which the committee could find no support.

Recently (in *European Polygraph*), Matte (2012) criticized the American Polygraph Association's (2011) Meta-analytic Survey of Criterion Accuracy of Validated Polygraph Techniques stating its authors failed to appreciate the reasons for the "exceptional accuracy" of the technique he developed, the Matte Quadri-Track Zone Comparison Test (MQTZCT). He asserted his technique was an outlier not because of any methodological flaws in the research upon which he relies, but rather the superiority of the technique itself. He maintained the unique elements of the test that render the technique superior, namely the "Inside Track" questions (described below), are responsible for the extraordinary accuracy reported.

In his 2012 article, Matte refers readers to a prior article (Matte, 2011) in which he explains the alleged benefits of the Inside Track. In that paper, he essentially argues the Inside Track is a remedy to false positives, false negatives and counter-measures. If true, we owe him a great debt of gratitude; however, Matte's analyses are seriously flawed. In this paper it is argued that his data (Matte & Reuss, 1989) and that of the two other validity studies he cites², Mangan, Armitage & Adams (2008) and Shurany, Stein & Brand (2009), when properly analyzed, clearly demonstrate the Inside Track does not function as he claims.

For those unfamiliar with the technique, the "Inside Track" consists of two questions that are the same in every test. The

first is a comparison question: "*Are you afraid an error will be made on this test regarding the target issue?*" The second is a relevant question: "*Are you hoping an error will be made on this test regarding the target issue?*" The "target issue" is the relevant issue addressed in the two primary relevant questions appearing earlier in the test (Mangan, Armitage and Adams, 2008; Matte, 2011). Matte theorized that truthful examinees fear an error and liars are hoping for an error, and thus the reason for the hope / fear question pair (Matte and Reuss, 1989a). (Thus, he contends both hope and fear - and degrees thereof - can be measured physiologically with a polygraph.) The format of the MQTZCT is as follows: 1) Neutral, 2) Sacrifice Relevant, 3) Symptomatic one, 4) Comparison one, 5) Relevant one, 6) Comparison two, 7) Relevant two, 8) Comparison three (*fear of error*), 9) Relevant three (*hope of error*), 10) Symptomatic two. Each relevant question is scored against the preceding comparison question (Mangan, Armitage and Adams, 2008; Matte, 2011).

Before directly addressing the main issues, discussion of two assertions Matte (2011) makes concerning possible MQTZCT validation pitfalls are in order. First, Matte claims the Inside Track scores can confirm the legitimacy of reactions to the other two relevant questions. Second, he claims the pre-test review of Inside Track questions "... can have a positive influence and effect on the salience of the [comparison]/relevant questions in the Primary and Secondary tracks." Matte continued, stating the following:

It is therefore incorrect to reach a conclusion based solely on the scores of the Primary and Secondary tracks without considering the psychological influence and effect that the Inside-Track questions had on the examinee and ensuing pairs of [comparison]/relevant questions contained in the Primary and Secondary tracks whose scores could have been weaker without that Inside-Track influence.

² The raw scores for these studies is available for download on Matte's website: <http://www.mattepolygraph.com/matte/valid-polygraph-techniques.html>

Thus, Matte claims – but offers no evidence – the mere discussion of the Inside Track questions can influence the salience of the remaining comparison and relevant questions. By arguing that it is improper to evaluate the scores of the first two relevant questions without considering the Inside Track scores, Matte seems to argue there is no way to validate the efficacy of the Inside Track, and yet also argues he and others have done that very thing.

Using the approach to which Matte appears to object, all three prior studies compare accuracies using the total score for tracks 1 and 2 to the score when using all three tracks, but they do not optimize scoring cut-offs for a two-track test. In other words, they use the same cut-scores optimized for a three-track test – which has about 33% more data – on a two-track test and then argue how the third track significantly reduces inconclusive decisions and errors. For example, Matte and Reuss (1989b) report correctly classifying only 43% of the truthful (with 52% inconclusive and 5% errors) when excluding the Inside Track compared to correctly identifying 91% (with 9% inconclusive) when including it. Mangan, Armitage and Adams report correctly classifying only 67% of the truthful (with 33% inconclusive) when excluding the Inside Track compared to correctly identifying 100% (with 0% inconclusive) when including it.

After comparing results with and without the Inside Track, the authors conclude the Inside Track offers something any standard CQ/RQ question pair does not. Mangan, Armitage and Adams (2008) state the

results, when comparing two track decisions to three track decisions, “[testify] of the significant fear that innocent examinees may have regarding the accuracy of the test and the threatening aspects of the relevant questions compared to the structurally less intense (not as threatening) control questions.” Arguing that inconclusive results may mask countermeasure use, they extrapolate further, “Clearly, the Inside Track prevents the successful use of countermeasures.”³ Despite their attempt to evaluate the Inside Track in a logical manner, the assumptions and research methodology employed, as others have pointed out, lack scientific rigor (Verschuere, Meijer, & Merckelbach, 2008). Therefore, their conclusions cannot be considered persuasive.

Because neither Matte nor the other investigators used a true experimental research design with control groups⁴, the only viable option is to do as they did and compare the efficacy of the first two tracks to that of all three – with one distinction: compare results with optimized cut-scores for a two-track test with those of a three-track test using traditional cut-scores. If accuracy using the first two tracks is no different than using all three, then there is no evidence to support Matte's theory – unless one accepts Matte's proposition that the discussion of the Inside Track questions enhances reactions to the other question pairs.⁵ Only if a real difference were found between the two and three track models would it be appropriate to begin to theorize the reasons for any difference, which would then require additional research to test the validity of the new theory.

³The application of the “Dual-Equal Strong Reaction Rule,” discussed below, is also credited.

⁴In other words, nobody was randomly assigned to a test condition in which the format was stripped of the Inside Track questions and associated discussion. Thus, there is no way to compare a non-manipulated group (i.e., the control without the IT) to a sample in which the IT variable was introduced (i.e., the experimental group).

⁵Of course, you could have such a discussion during the pre-test of *any* polygraph examination (without presenting the Inside Track questions in the test). However, without empirical tests using a good research design, it is difficult to know if introducing the Inside Track questions does what Matte suggests – or even the exact opposite of what he suggests – in regard to impacting the salience of other questions.

Method

Data from the three published studies were used to calculate correct decisions, errors and inconclusive results using the first two tracks (ignoring the Inside Track), which were then compared to the results using all three tracks (i.e., the original results reported in each study). First, optimized cut-scores were created for a two-track test using the data from the first published study.

Matte (1989a) reported scores for each individual case in his study, which included total scores for the first two relevant questions

(tracks) and scores for the Inside Track comparison / relevant pairs. Therefore, it is possible to compare results with the Inside Track (i.e., the total test score with all three tracks) versus results using only the first two relevant questions (i.e., ignoring the Inside Track scores). Scores were entered into a spreadsheet and sorted from lowest score to highest score using only the first two relevant questions. (In other words, only the grand total scores of R1 and R2 (combined) were used.) As can be seen in Figure 1, the score at which the truthful appear to be best separated⁶ from the deceptive is a score of -2.

Figure 1. Approximate Center of Distribution of Track 1 and 2 Combined Scores

TK12	-13	-12	-12	-12	-12	-11	-11	-10	-9	-2	-2	-2	-2	-2	0	0	0	0	1	1
GT	D	D	D	D	D	D	T	T	D	T	T	T	T	T	T	T	T	T	T	T

“TK12” = Total Exam Score (i.e., all charts) for Tracks 1&2

“GT” = Ground Truth (“D” = Deceptive; “T” = Truthful)

After visually estimating optimal cut-scores for a two-track test, accuracy was assessed using cut-scores of -2 / -3. That is, if the score was -2 or greater, the decision was no deception indicated (NDI); if it was -3 or less, the decision was deception indicated (DI). No inconclusive zone was used. This scoring method was then cross-validated on the Mangan, Armitage & Adams (2008) and Shurany, Stein & Brand (2009) data available from Matte’s website (see footnote 2). For simplicity, from this point on each dataset will be referred to by the name of the respective study’s primary author: Matte, Mangan, & Shurany.

One might conclude the data support optimal cut-scores that include an inconclusive zone where the distributions overlap. For unknown reasons, there are no scores between -2 and -9 in the Matte sample, which is not the case in the other samples, and which makes the selection of an

inconclusive zone little more than a guess. Using a visual inspection of the data, the distributions overlap from -9 to -11. Scores for the deceptive and truthful then begin at -12 and -2, respectively. Thus, the data support, albeit weakly due to the gap in expected scores, the possibility of an inconclusive zone from -3 to -11.

An inconclusive zone of -3 to -11 does not cross-validate on the other samples. Using the same process, the Mangan data supports an inconclusive zone from -2 to -1; the Shurany data, -2 to 1. Therefore, the wide inconclusive zone was not used for the analysis. For informational purposes only, results were assessed using the wide inconclusive zone: The proportions correct and erroneous were not significantly different but the proportions of inconclusive classifications were simply because of the large number of deceptive subjects that were classified as inconclusive. There were no,

⁶ For the sake of simplicity, I estimated based on a visual inspection of the data. Statistical norming is also an option, but it is neither necessary nor helpful since I cross-validated the scoring rules on the two remaining data sets.

differences in the proportions of truthful subjects classified as inconclusive.

Intuitively, some might suggest simply adjusting a two-track test cut-score to two-thirds of the cut-score for a three-track test. (For example, use a score of +2 per chart without the Inside Track – instead of the +3 used when the Inside Track is considered – as the cut-score for a truthful classification.) That assumes the inside-track and the primary tracks result in comparable scores, with each CQ/RQ pair contributing approximately one-third of the total score. However, a t-test comparing the Inside Track scores to those of the averages of the two primary tracks in the Matte data reveals a significant difference: $t(126) = -6.93, p < .0001$ (deceptive) and $t(114) = -5.72, p < .0001$ (truthful). (The other data sets suffer from similar problems.) Consequently, reducing cut-scores by one-third will not achieve a true comparison of the value (or lack thereof) of the Inside Track.

Returning to Figure 1, the very low scores (some negative) generated by the truthful may surprise some. Those scores are likely due, in part, to Matte's "Dual Equal Strong Reaction Rule" which "...demands that when the *red* [relevant] and *green* [comparison] zones being inter-compared both contain timely, specific and significant reactions of maximum and equal strength, a minus one (-1) score is assigned to that spot." Matte reasons the rule is necessary to combat defective comparison questions and countermeasure attempts, thereby reducing

inconclusive classifications (Matte, 2011). Additionally, over half of the exams in the (Matte) dataset contained only two charts, and fewer charts should, at least in theory, result in lower (but not necessarily negative) scores. In any event, decisions were simply based on static score totals (regardless of the number of charts) as described above.

Results

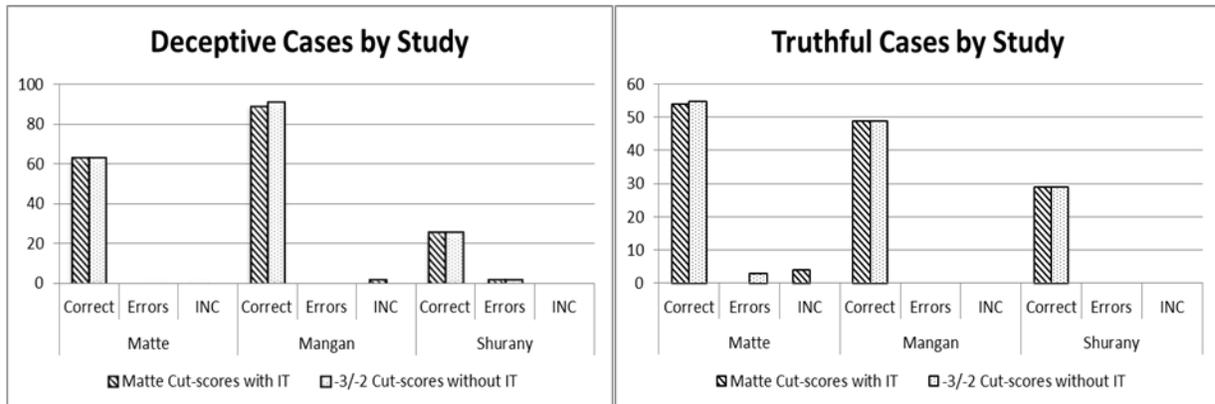
The Matte data resulted in 117 correct, 0 errors and 5 inconclusive when all three tracks were considered, and 118 correct, 4 errors and 0 inconclusive when only the first two tracks were considered. The Mangan data resulted in 138 correct, 0 errors and 2 inconclusive when all three tracks were used and 140 correct, 0 errors and 0 inconclusive when only the first two tracks were considered. The Shurany data resulted in 55 correct, 2 errors and 0 inconclusive with and without the Inside Track. (That is, the results were identical when the Inside Track was ignored.) The truthful and deceptive results are broken down by study in Table 1 and Figure 2. (All cases are combined for comparison purposes and labeled "Total" in Table 1.) Fisher exact statistics were calculated separately for truthful and deceptive. Additionally, the same statistics were calculated with and without inconclusive results. When inconclusive results were considered, they were considered as errors. Thus, calculations were based on 2X2 tables of results. There were no significant differences in any of the results with or without the use of the Inside Track ($p > .05$).

Table 1. Original and Cross-validation Results of Scoring Only the IT with Adjusted Cut-scores

DECEPTIVE CASES	Matte			Mangan			Shurany			Total		
	Correct	Errors	INC									
IT	63	0	1	89	0	2	26	2	0	178	2	3
No IT	63	1	0	91	0	0	26	2	0	180	3	0
TRUTHFUL CASES	Matte			Mangan			Shurany			Total		
	Correct	Errors	INC									
IT	54	0	4	49	0	0	29	0	0	132	0	4
No IT	55	3	0	49	0	0	29	0	0	133	3	0

Note: Original study results (using Matte scoring rules with the IT) are included for comparison.
 IT = Matte scoring rules applied to all three tracks (original study results)
 No IT = Optimized scoring rules (-3/-2) applied to first two tracks only, ignoring IT scores

Figure 2. Original and Cross-validation Results Comparisons for Deceptive and Truthful Subjects



Note: "Matte Cut-scores with IT" are the original results (using Matte scoring rules with the IT) reported in their respective studies.

Discussion

Three studies have been produced by proponents of the MQTZCT, and those three studies consistently support the same conclusion: scoring the Inside Track provides no benefit over ignoring it and scoring only the first two tracks. Whereas scoring the Inside Track offers no benefit, then it cannot reasonably be concluded that it reduces errors or inconclusive classifications (which allegedly may mask countermeasures). In other words, the theory behind the Inside Track's efficacy cannot be correct, and the "extreme" accuracy reported must be due to some other factor, several possibilities of which were raised in the APA's (2011) meta-analytic survey.

Whether the Inside Track is scored or ignored, the prior MQTZCT studies offer some support for the validity of the technique overall. However, the precise accuracy estimates by its proponents are questionable given the suspected methodological flaws that have been identified by Patrick and Iacono (1991), Iacono (2008) and Verschure, Meijer & Merckelbach (2008), along with the flaws

discussed in the APA survey (2011). However, something more than rebuttals, objections and quibbles are necessary to better assess the extent to which some of the possible flaws actually affect the accuracy of the MQTZCT (or any other technique for that matter). Moreover, all MQTZCT studies use convenience – not probability – samples, making generalization risky. Thus, it appears the Inside Track questions, for all practical purposes, function much the same as traditional CQ/RQ pairs. (However, see the below discussion regarding the Dual Equal Strong Reaction Rule.) Even if the Inside Track questions had otherwise served as Matte suggests, it seems implausible that the mere discussion of those questions could influence the salience of the other CQ/RQ pairs given that the fear of error or hope of error, if either exists in any given examinee, would be present without a discussion and review.⁷ Nonetheless, it is erroneous to conclude their introduction or use has any special impact given the inadequate research design employed in all three studies coupled with the results presented here.

⁷ Perhaps the "fear" and "hope" are introduced by the examiner, but there is no way of knowing if that is the case, so any conclusions are speculative. To test the idea, one would need to assess each examinee's level of "hope" and "fear" before the exam, in so doing, introducing them if they did not already exist. To quantify them as in the MQTZCT (with a score, index, etc.), psychological testing would be necessary before polygraph testing. The bottom line is that the burden of proof is upon the promoters of the technique, and to date, they have not made a persuasive argument.

The Dual Equal Strong Reaction Rule deserves consideration⁸ given its purported benefit as a countermeasure defater (Mangan, Armitage & Adams, 2008; Shurany, Stein & Brand, 2009). However, the rule results in a shift of the scores of the MQTZCT's first two tracks in the negative direction, but it does so asymmetrically. That is, it pushes the truthful scores in the negative direction, but is essentially neutral for the deceptive. Other evidentiary techniques that do not use the Dual Equal Strong Reaction Rule or the Inside Track report inconclusive rates of approximately 10% for both the truthful and deceptive (cf. APA, 2011). The proponents of the MQTZCT consistently report unusually high rates of inconclusive scores for their truthful samples (without the Inside Track), as mentioned in the introduction. For the truthful samples, the MQTZCT studies report an average of 38% inconclusive when the Inside Track scores are ignored. Their deceptive samples result in approximately 15% inconclusive scores when the Inside Track is ignored.⁹ The Dual Equal Strong Reaction Rule therefore appears to hamper the rendering of a conclusive and accurate classification for the truthful, but that is not true for the deceptive.

Matte's main remedy to the biasing effect of the Dual Equal Strong Reaction Rule is the Inside Track, a question pair that is structurally different from typical CQ/RQ pairs in that its "Fear-Hope" questions target the examinee's *emotional rapport* with the testing situation (*not involvement in, or knowledge of the issue* specifically targeted by the primary relevant questions). The "fear" question either addresses or induces a concern in the truthful about a possible diagnostic error, i.e., a false positive; conversely, the "hope" question either

addresses or induces a concern in the deceptive about being found deceptive, i.e., about not making an error and exposing the deception (a true positive).

By targeting the examinee's *emotional rapport* with the testing situation, the Inside Track questions operate on a level that is psychologically different from that of a typical CQ/RQ pair. However, the Inside Track cannot operate directly by selectively altering the salience of the comparison or relevant questions (during the test itself) as Matte speculated (Matte, 2011). If it did, there would be no need to add the Inside Track's score to that of the first two tracks in order to bring the inconclusive scores into a conclusive range. The result, practically speaking, is a CQ/RQ question pair that acts, functionally, as a traditional CQ/RQ pair. However, the benefit of the Inside Track appears limited to those techniques that use the Dual Equal Strong Reaction rule, which is bound to fade away as one adopts sound, standardized practices for ZCTs as established through research, starting with standard numerical scoring procedures and empirically optimized cut-scores.

A cautionary note is perhaps in order. Some may be tempted to argue that since the technique seems to be valid overall, then each of the technique's component parts must therefore be valid. Such a conclusion is based on a logical fallacy known as the "fallacy of division." Just because the overall technique has a certain property – in this case, the property of being "valid" – that does not mean that each of its component parts shares that property. For example, water – H₂O – has the property *wet*, but if we were to separate the parts, that is, the hydrogen from the oxygen, each part would not have that same property.

⁸ Again, I must thank and credit an anonymous reviewer(s) for pointing out the rule's biasing impact and Matte's remedy, the Inside Track. This portion of the discussion would not have produced here had it not been so thoroughly discussed by the reviewer(s).

⁹ Shurany, Stein & Brand (2009) report a dramatically high rate of inconclusives for their deceptive sample (71%), which is an exception to this interpretation. However, the status of this outlier result is highly questionable as it runs contrary to any outcome the Dual Equal Strong Reaction Rule would predict. The reason for the outlier is unknown, but the more positive scores reported by Shurany et al. (track 1 and 2 combined mean score of -3.4 as opposed to mean scores of -7.3 and -6.2 for the Matte and Mangan data, respectively) may be explained by undetected countermeasures or more conservative scoring, for example. Until this outlier result can be replicated and its causes better understood, caution is in order.

Additionally, just because the Inside Track questions appear to be “valid” (in that they ultimately appear to function like any other CQ/RQ pair), that does not mean the underlying theory prompting their use is valid, which is why the results here should not be surprising.

Increasing polygraph accuracy (and reducing errors) is a noble task. Regretfully, despite what the authors of these studies claim, the MQTZCT has not been proven to do so. It seems clear the Inside Track questions serve as little more than a more complex but unnecessary CQ/RQ pair. When all is said and done, the MQTZCT is just another ZCT, and as such, we should not expect accuracy any better (or worse) than any other validated single-issue ZCT. Thus, the evidence that the MQTZCT is an outlier as concluded in the APA meta-analysis (2011) is now even stronger. The most parsimonious explanation for its “exceptional accuracy” is likely nothing more than a methodological flaw (or flaws), rendering the accuracy results reported in the prior studies non-generalizable to field settings.

Matte’s (2012) criticisms are appreciated, but his premises are based on conclusions from, as demonstrated here, previously flawed findings and reasoning. The theory upon which the MQTZCT is based has now been adequately proven implausible, likely rendering the MQTZCT’s use in the field impractical. Matte has spent a good portion of

his life promoting and defending the underlying theory of the MQTZCT. From its inception, Matte’s goals have been noble: to reduce errors and inconclusive classifications. He observed a problem (errors), theorized the reasons for them, and created a possible solution to remedy them (Matte, 1978). However, theory building is different from theory testing, and those practicing and researching in the field have a duty to advance the science by self-policing and establishing an historical record of theories, hypotheses, successes, failures and criticisms, etc. Scientific discoveries begin with personal observations, but science requires objectivity – insofar as that is possible for a human experimenter – basing conclusions on good evidence and solid reasoning. The NAS Report (2003) warnings cannot be ignored. If polygraph as a science is to flourish, the polygraph community must embrace and promote independent research, and it must begin to question its untested, or inadequately tested, theoretical assertions. It must also be prepared to accept that some of those theories will not survive careful scrutiny, as is the case here. That is the nature of scientific investigation. It is not personal, but it is necessary in order to be accountable to consumers of polygraph. James Matte’s commitment and many contributions to the field of polygraph are to be applauded, but the evidence simply does not support that the MQTZCT offers the benefits he states it does, and his criticism of the APA meta-analysis in that regard is therefore without merit.

References

- American Polygraph Association. (2011). Meta-analytic survey of criterion accuracy of validated polygraph techniques. *Polygraph*, 40(4), 193-305.
- Iacono, W. (2008). Accuracy of polygraph techniques: Problems using confessions to determine ground truth. *Physiology & Behavior*, 95(1-2), 24-26.
- Mangan D., Armitage T., & Adams G. (2008) A field study on the validity of the Quadri-Track Zone Comparison Technique. *Physiology and Behavior*, 95, 17-23.
- Matte, J. (1978). Quadri-Zone comparison technique. *Polygraph*, 7(4), 266-280.
- Matte, J., & Reuss, R. (1989a). *Validation Study on the Polygraph Quadri-Zone Comparison Technique*. Research Dissertation Abstract, LD 1452, Vol. 1502.
- Matte, J., & Reuss, R. (1989b). A Field Validation of the Quadri-zone Comparison Technique. *Polygraph*, 18(4), 187-202.
- Matte, J. (2011). Psychological Aspects of the Quadri-Track Zone Comparison Technique and Attendant Benefits of its Inside Track. *European Polygraph*, 5(2), 41-59.
- Matte, J. (2012). Critique of Meta-Analytic Survey of Criterion Accuracy of Validated Polygraph Techniques. *European Polygraph*, 6(1), 19-34.
- National Research Council (2003). *The Polygraph and Lie Detection*. Washington, D.C.: National Academy of Sciences.
- National Research Council's Committee on Identifying the Needs of the Forensic Sciences Community (2009). *Strengthening forensic science in the United States: a path forward*. Washington, D.C.: National Academy of Sciences.
- Patrick C., & Iacono, W. (1991). Validity of the control question polygraph test: the problem of sampling bias. *Journal of Applied Psychology*, 76, 229-38.
- Shurany, T., Stein, E., Brand, E. (2009), A Field Study of on the Validity of the Quadri-Track Zone Comparison Technique. *European Polygraph*, 1(7), 5-23.
- Verschuere, B., Meijer, E., & Merckelbach, H. (2008). The Quadri-Track Zone Comparison Technique: it's just not science. A critique to Mangan, Armitage, and Adams (2008). *Physiology and Behavior*, 95(1-2), 27-28.