

Rejoinder to Matte

Barry Cushman

I am writing in response to James Matte's letter to the editor in regard to my article, "Is Matte's Inside Track the Answer to False Positives, False Negatives and Countermeasures? There is Reason to Fear that Hope is Gone" (Cushman, 2013), which is published in the current edition of this journal. I will briefly address *some* of Matte's assertions. While I tried to limit the former article to an analysis of the data and avoid a personal debate with any of the authors of the MQTZCT studies, Matte has written a response describing his personal thoughts describing how he conceived the technique and why he reasons it is superior to others. Any response of mine, therefore, may be viewed by some as an attack against the person of James Matte rather than what it is: a rejection of some of his faulty ideas, methods and conclusions. Again, I appreciate his zeal and tenacity. He is certainly a friend to the cause of elevating the status of polygraph. However, if we are to be gain credibility alongside our sister forensic sciences, it's necessary to separate fact from fiction and be responsible enough to clearly report what we can reasonably conclude to know versus that which we have only conceived in our minds. Again, I respect James Matte as a colleague, but I would not want my silence to imply agreement, and therefore I provide a response.

First, there is little, if any, new information in Matte's arguments despite the fact that his letter is significantly longer than my article. In fact, much of what appears in his response appears verbatim in some of his other publications, most all of which are freely available on his website. Matte's weak and specious arguments, despite having been re-stated multiple times, are still weak and specious arguments. Restating the same weak argument does not make it any stronger or any more valid. Moreover, it is inappropriate to reuse one's prior work without properly citing those works (i.e., engage in self-plagiarism), especially when bypassing the peer-review process by submitting a letter to the editor.

Second, Cushman (2013) analyzed the data in the three existing field studies to assess whether the inside track does what its proponents claim. As reported, accuracy is the same whether one uses or ignores the inside track. Thus, it simply cannot do what its proponents claim. Nothing Matte has stated in his letter changes that fact. In his response, Matte chose, however, to ignore the data and instead resorted to restating his unproven assumptions. Thus, it is more of an opinion editorial than a serious response.

From this point I will simply summarize or quote several of Matte's claims and follow with my responses.

Claim: Matte (2011) "...fully describes the psychological aspects and functions of each component of the MQTZCT and addresses... the effectiveness of the Inside Track."

Response: Matte needs to be clear as this is misleading. Matte (2011), as does his letter to the editor, simply discusses his reasoning for why he *thinks* the technique *should work* so remarkably. He (along with the authors of the subsequent studies) offers nothing more than demonstrably unjustified conclusions based on a wholly unscientific analysis of the available data. Other than the analysis provided in Cushman (2013), no attempt has been made to adequately assess the effectiveness of the inside track. Whereas there are now three replications of a study that consistently show the same thing (i.e., ignoring the inside track offers no benefit over scoring it), Matte's theory is clearly wrong and should be abandoned.

Matte never defines what he means by "psychological aspect" or "functions." He simply articulates several vague and unproven hypotheses. Psychology has to do with the scientific study of the mind

and human behavior. Since the mind is not something we can see, discerning what is occurring in the mind of another is difficult, to say the least. The general approach, therefore, is to present a stimulus and observe any response. Once a predictable relationship is found, then it may be appropriate to try to figure out what is happening in the mind, if possible, and what moderates or mediates the effect of interest. A simple observational study using a biased convenience sample is not sufficient to isolate all the variables necessary in order to learn if each of the alleged “aspects” of the MQTZCT does anything. While it might sound sophisticated to describe the technique in terms of its “psychological aspects,” the fact of the matter is that Matte is only continuing to offer, in large part, unfounded, illogical or disproven assumptions.

Claim: The MQTZCT both “safeguards” against and “prevent[s],” the errors identified in Chapter 9 of Matte, 1996, including the “fear of error” by an innocent examinee.

Response: Matte claims he has identified certain variables that cause errors during polygraph testing. Again, he offers no empirical evidence that any of those variables cause errors (though they very well may), much less any evidence to show he has found any solutions. He accepts as factual Ekman’s yet unproven hypothesis that a person fearful of being misbelieved might react more strongly to the RQs than the CQs. Ekman might be correct, but he might also be wrong. While it’s not unrealistic to think fear of being misbelieved might result in behavioral cues that could be misjudged as deception, no one has demonstrated that CQs are less salient in examinees who might have such concerns. (The NRC (NAS, 2003) only said the argument was “plausible” and “might” be a reason for false positives.) Interestingly, Bradley and Janisse (1981) found scores for the innocent were higher for examinees whether they were told they were detected 66% or 100% of the time on their demonstration tests – and the mean raw scores were higher at 66% than at 100%, suggesting their concerns of error, if any effect, increase differential reactivity in the opposite direction Ekman’s¹ and Matte’s hypothesis predicts.

Matte quotes Ekman² in regard to the so-called “Othello error³,” and he asserts the inside track remedies Ekman’s concerns and therefore, those raised by the NRC in its report (NAS, 2003). He fails to mention Ekman served on the committee that published the NRC’s report. He was therefore well aware of the MQTZCT since the committee included it in its report and commented on it specifically on page 351: “Another, with [an accuracy estimate of] A essentially equal to 1.0..., was the doctoral dissertation of the president of a polygraph company, based on data from polygraph tests that used a specific variant of the control question test, with an associated scoring mechanism, that the author had developed some years earlier. Of 122 polygraph tests from criminal investigations that were examined in this particular study, there were seven inconclusive tests but no false positive or false negative errors.” Rather than finding the MQTZCT to be the answer to the problems the committee raised, the committee – including Ekman – said the following (on page 108), which, when considered together, tells us they found Matte’s study too

¹ Ekman (2001) refers to, but does not cite, two studies for which he claims there are flaws that leave the question unanswered. He likely refers to Bradley and Janisse (1981), whom he cites elsewhere in the book, but I cannot be sure.

² Matte consistently cites “Ekman, P. (1985). *Telling Lies - A How-To-Guide for All those Who want to Detect Lies*. New York, NY: Berkley Books.” Interestingly, Shurany, Stein and Brand (2009) do the same. However, there is no record in the Library of Congress of such a title by that publisher; no such book was listed when I did a search of major resellers; nor does Ekman list it among his publications. They likely mean the 1985 version of the book I cite here: Ekman, 2001.

³ Matte (along with the authors of the other two studies) makes a fatal error in reasoning when he mischaracterizes Ekman’s “Othello error” by suggesting it describes an error that occurs *only* when the fear of being misbelieved is interpreted as fear of detection. Ekman (2001), however, used a more broad description that included guilt, hostility or emotions about the event itself. Thus, the inside track, if it otherwise worked as its proponents claim, does not address the Othello error as used by Ekman since it occurs any time a truthful person is misjudged.

flawed to be of any real value: “Because the studies vary greatly in quality and include several with extreme outcomes due to small size, sampling variability, bias, or nongeneralizable features of their study designs, we did not give much weight to the studies with outcomes at the extremes of the group.” Whereas Mangan, Armitage and Adams (2008) and Shurany, Stein and Brand (2009) are replications of Matte and Reusse (1989), they suffer from the same problems, and they too should not be given much weight.

Claim: Only a deceptive examinee will react to the hope of error question, and reactions to inside track questions confirm the legitimacy of reactions to CQs or RQs in the first two tracks.

Response: Matte offers no evidence to support his assertions. Moreover, the concepts themselves depend on several unproven assumptions. Matte posits that only the truthful fear errors and only the deceptive hope for errors (cf. Matte, 2011). Then, he assumes both “fear” and “hope” – or degrees thereof – can be measured by means of polygraph. How he draws such a conclusion is a mystery as none of the former studies assessed this feature. Why is it that Matte suggests there must be agreement in *two or more* charts in order for the inside track questions to be considered as validating the reactions to the CQs or RQs? If only the truthful or deceptive should react to the fear of error or hope of error questions, respectively, then what Matte is really saying is that 100% of the variance is accounted for by fear and hope.⁴ However, Matte knows that is not the case since 1) his “reaction combination guide” requires a “remedy” when a person does not respond as predicted on the inside track questions, and 2) confirmed truthful and deceptive examinees sometimes achieve negative and positive scores, respectively, in his dataset, which runs counter to his argument. Additionally, the requirement for two or more charts with inside track scores on the same side of zero also reveals that Matte recognizes that some other variable or variables (besides “fear” and “hope”) are, at times, responsible for reactions to the inside track questions. How much of the variance is not accounted for as theorized by Matte? What accounts for that unexplained (and yet unquantified) variance? None of the studies supporting the MQTZCT give us so much as a hint.

Matte’s reasoning is arguably circular but clearly invalid. For example, as stated above, he assumes that only the deceptive would hope for an error, and he further assumes asking whether a person is hoping for an error will be revealed by a negative score on the inside track’s fear question. Then, when a deceptive examinee happens to react to the question, Matte concludes that reaction verifies his assertion. That’s like making the following (fallacious) argument: If it’s raining outside, then the ground will be wet. The ground is wet. Therefore, it is raining outside. There are many reasons the ground may be wet. Maybe it rained a while ago. Maybe somebody watered the lawn. Maybe the septic system failed. (Using the language of logic, Matte’s error is a fallacious argument known as affirming the consequent.) From a review of Matte’s scoring data, it is clear there is something other than what should either be consistent “hope” or “fear” that results in a person reacting to the inside track questions. Thus, there is simply no scientific support for the view that the inside track questions actually measure hope or fear – even if we pretend for the sake of argument that the psychological community has agreed on the nature of hope and fear.

Claim: The inside track is a benefit with sex-related targets because it avoids using “stigmatic language.”

⁴ Matte also said the inside track questions “...are discussed in depth with the examinee to assure his understanding and desired comprehension of those test questions.” Thus, he argues that the examiner not only plays a part in controlling the variance in responses, but can somehow manipulate the examinee in such a way as to control 100% of the variance when the inside track questions are presented, which is inconsistent with a need for a so-called remedy.

Response: This makes little sense on its face since the primary tracks would, by necessity, directly probe the relevant issue. Matte seems to be arguing that the real benefit is with the truthful since the deceptive are expected to react to all RQs including the inside track RQ. There are a couple problems with his assertion. First, the Cushman (2013) analysis demonstrated the inside track offers no benefit since results were the same when the inside track scores were ignored. Second, *none* of the prior studies assessed the alleged effect of the inside track on the truthful who were accused of sex offenses. Matte could not have done so if he wanted since *none* of the confirmed truthful cases in his study were sex offenses! To speculate as Matte does is irresponsible and, at a minimum, misleading.

Claim: The high scores in the three MQTZCT studies "...provide a significant buffer against errors," and using lower thresholds for cut-scores could endanger examinees who arguably "hallucinated" a low score.

Response: This is a red-herring. The cut-scores used in Cushman (2013), as stated in the article, were based on Matte's (1989) own data, and they cross-validated on the other two data sets. Moreover, the argument requires Matte to hold two conflicting views. On one hand he suggests he has established a normative predictive table based upon his data, and on the other he wants to argue against using a data-driven approach out of fear that it could be difficult to counter an unscientific objection to empirically derived 'low' cut-scores. (As an aside, the mean scores for the Shurany, Stein and Brand (2009) data are not consistent with the raw data provided on Mr. Matte's website. Which is erroneous is unknown.)

If Matte is concerned that an attorney might suggest an examinee "hallucinated" a grand total score of -3 over three charts, then perhaps he could add another question to the MQTZCT that would theoretically work like he claims the symptomatic questions work. (For the sake of argument, assume for a moment that the research actually supports the efficacy of symptomatic questions.) Perhaps the question could be phrased as follows: "Did you hallucinate during the presentation of any questions on this test?" One could reason that only an uncooperative deceptive examinee would have an incentive to hallucinate, which would allow for its simple introduction during the pretest. This is silly, of course, but it is representative of the type of reasoning on which much of the so-called psychological aspects of the test are based.

While Matte makes no mention of it, Mangan, Armitage and Adams (2008) state that a blind scorer made one false positive ("Othello") error. To achieve such an error would require a score of at least -10. A correct score must be at least +6 - a 16-point difference (at least). Using Matte's reasoning, one could argue the absolute values of cut-scores are not set high enough, and yet Matte hasn't adjusted them to avoid such errors. Why? It's simple: changing cut-scores only shifts the proportion of errors and inconclusive results, and at some point a policy decision must be made as to what is acceptable. Whether one uses the higher scores on the Matte data with the inside track or lower scores without it, nothing changes. Both cross-validate on the other two data sets, revealing that the inside track does not do what is claimed and that high scores, despite the suggestion of the possibility of hallucinations, are unnecessary to classify the samples correctly.

Claim: Matte's (1989) "predictive table" provides estimates of error rates, which support the validity of the MQTZCT.

Response: Matte would be well served to review the NRC report (NAS, 2003) as it discusses the reasons why his "predictive" table cannot, due to flaws in the methodology employed in his study, predict anything. Moreover, the table is based on the scores of three people - one of which only scored three of the study's tests, so it cannot be considered normative. The table only tells us the distribution of the chart averages for the three examiners' scores. (Interestingly, the bell curves depicting the truthful and deceptive distributions [corresponding to the "predictive" table] reveal a bias against the truthful, so the comments in Cushman (2013) about bias should not be surprising to Matte.) Additionally, the table to which Matte refers is based on z-scores derived from the

average score per chart, which he suggests can be adapted for any number of charts. If that were sound reasoning, then z-scores based on two charts (and their corresponding error estimates) should equal those based on the averages in his table, but they are not (and mathematically, we would not expect them to be equal). Interestingly, the only two errors (false negatives) for which data is provided in any of the three studies have, according to the table, a 0% chance of error!

If Shurany, Stein and Brand (2009) provides us with a more representative estimate of the error rate for deceptive examinees, then it can be used to calculate the probability of finding no false negative errors in the other two data sets. Shurany, Stein and Brand (2009) reported two false negative errors out of 28 definitive decisions. Thus, they report an error rate of approximately 7%. (The 95% confidence interval is approximately 2% to 23%; however, only the point estimate will be used in the computations.) Given a false negative rate of approximately 7%, the probability of randomly selecting the Matte and Reuss (1989) confirmed deceptive cases without selecting any erroneous (deceptive) cases is 0.01. That is small enough to conclude it wasn't by chance, and therefore we can conclude there was very likely something wrong with the sampling methods. The probability of randomly selecting confirmed cases with no false negative errors in the Mangan, Armitage and Adams (2008) data set is 0.001, which is also enough to conclude the selection was not by chance. The probability of the two independent datasets failing to contain any false negative errors is therefore 0.000014, or about 1 chance in over 72,000!

As stated in Cushman (2013), there appears to be some evidence of the validity of the MQTZCT simply because it generally adheres to validated principles and the inside track seems to work much like any other CQ/RQ pair does. Thus, we'd expect accuracy consistent with any other validated ZCT. The problem, however, is that we still do not know how well (or poorly) the MQTZCT actually works relative to other ZCTs if there is a difference from one ZCT to another (and the APA meta-analysis suggests there is not). The APA meta-analysis (2011) listed the MQTZCT among the ZCTs meeting "evidentiary" standards, but it did so with a strong warning that it could have instead chosen to use as a compelling reason for excluding the MQTZCT from the ZCTs that meet the standard. The MQTZCT is clearly an outlier. Is it an outlier simply because it is a superior technique as Matte claims? No, a more reasonable explanation is it's an outlier because three seriously flawed studies conclude it has perfect or near-perfect accuracy, which is quite extraordinary. Therefore, extraordinary evidence should be required to support that claim, and it simply does not exist – and, as stated already, what does exist is inconsistent with Matte's claims (cf. Matte, 2012).

In reality, the MQTZCT, as a distinct variant of a ZCT, is one for which accuracy is unknown – especially when used by examiners not represented in the studies (which may be most). Given the high probability of sampling flaws, it certainly is not a technique that should be offered in an evidentiary setting – at least not as a superior technique that solves the problems Matte claims it does. Again, there is reason to argue its accuracy is likely consistent with any other validated, event-specific ZCT, but this is, arguably, more of a theoretical argument than an empirically based argument, and the current APA standards define a "technique" very specifically and narrowly. In other words, if one is satisfied that the research supports concluding that any properly conducted event-specific ZCT (based on validated principles) is valid, then the MQTZCT is valid, and its accuracy is likely consistent with what is reported in the 2011 APA meta-analysis (i.e., approximately 89% with an estimated 95% CI of 83% to 95%). If, however, one requires (as do the current APA standards) a specific format and scoring protocol to be independently validated with at least two studies, then there are more questions than answers for the MQTZCT, and its accuracy in average field situations is unknown.

Claim: Cushman's (2013) criticism of a lack of an experimental design is unwarranted because a control group is not feasible in a field study of polygraph. (In other words, an experimental design is not possible in field settings.)

Response: Again, Matte may want to review the NRC report (NAS, 2003) as it discusses the weaknesses of his methodology and makes some suggestions for utilizing an experimental design in field settings. Ironically, the MQTZCT's inside track probably makes it one of the easiest CQTs for which a field-based experimental design could be employed since it is essentially only necessary to manipulate the test by randomly inserting or removing the inside track questions. Even if it were true that an experimental (or even quasi-experimental) design were impossible, that would not change the fact that an observational study, for example, was more limited in what type of information it could provide. Just because Matte thinks he can't conduct an adequate field study doesn't mean his inferior methodology is an equal substitute of what would be adequate, nor does it give him license to ignore the limitations of the inferior methodology.

Claim: Mangan, Armitage and Adams (2008) "persuasively addresses" the criticisms levied by Iacono (2008) and Verschuere, Meijer, and Merckelbach (2008).

Response: This is unequivocally nothing more than Matte's opinion – an opinion with which others disagree. That's okay, but a battle of opinions will generally not settle any matter. It is interesting to note, however, that the Mangan, Armitage and Adams (2008) paper has been cited a few times since its publication, and the reasons for doing so are not flattering. For example, the textbook *Research Methods Laboratory Manual for Psychology* (Langston, 2011) cites it to support the point that "...peer review does not guarantee that something published is 'good'..." *The Handbook of Forensic Psychology's* chapter on polygraph maintained the following: "...it represents a flawed report that one published commentary characterized as a failure of the peer review system..." (Iacono & Patrick, 2014). Their comments are hardly the praises of those who found the rebuttal persuasive, and they are likely influential texts in the scientific community not only now, but for years to come. Rather than suggesting the criticisms were overcome, these authors have designated Mangan, Armitage and Adams (2008) as the poster-child for a bad (but peer-reviewed) study.

Claim: Backster's Either-Or Rule supports Matte's Dual Equal Strong Reaction Rule.

Response: The Dual Equal Strong Reaction Rule, which is likely responsible for the bias in MQTZCT scores, is not the Either-Or Rule, and the research on it, whether appropriately interpreted or not, does not and cannot validate the theory or efficacy of the Dual Equal Strong Reaction Rule.

Claim: The "double-bind effect" discussed by Bateson, Jackson, Haley and Weakland (1956) can explain why one question type (i.e., CQ or RQ) is more salient than another.

Response: The paper cited discusses a theory of schizophrenia. It has nothing to do with polygraph. According to the authors, the "ingredients for a double bind situation" (prior to "the victim [learning] to perceive his universe in double bind patterns") include the following: 1) "Two or more persons," one of which is a "victim"; 2) "Repeated experience"; 3) one of two forms of "A primary negative injunction"; 4) "A secondary injunction conflicting with the first at a more abstract level"; and 5) "A tertiary negative injunction." Why bore readers with what is likely gibberish to many? I think it's important to more emphatically make a point. Matte's inclusion of the term appears to be an attempt to commandeer it, give it a proprietary meaning, and claim a supporting history in the scientific literature (behavioral psychology). (Matte further confuses the issue by using a dictionary based definition in a footnote.) The cite is improper, and the concept of the double-bind effect as proposed by Bateson, Jackson, Haley and Weakland (1956) should not be advanced as an explanation of the psychological construct underlying the differential reactivity observed in comparison question testing.

Claim: The MQTZCT pre-test requires the examiner to remain professional, neutral and unbiased throughout, which is a best practice designed to achieve optimal results without improperly

influencing the examinee. (Said differently, a pre-test interrogation is improper, and a proper pre-test is key to accurate results.)

Response: I agree. However, I believe the same general protocols outlined by Matte are required in all of the ZCTs discussed in the APA (2011) meta-analysis, which adds to the argument that the MQTZCT is just another ZCT.

Conclusion:

It should be abundantly clear that the “psychological aspects” on which the MQTZCT is allegedly built is little more than a house of cards. Matte’s defense of his technique is not only based on old, weak arguments, but the data consistently fails to support his claims. He offers no explanation of why it is possible to ignore the inside track scores and still achieve the same “exceptional accuracy” as with its inclusion (other than to continue to speculate that the introduction of those questions impacts the CQs and RQs). We now have three studies that all support the same conclusion (i.e., the inside track cannot work as claimed), and we presumably have the best defense Matte can offer in his response here (which is essentially somewhat of an abbreviated version of Matte, 2011). Despite Matte’s appeals to the contrary, there simply is no compelling reason to accept his claims and reject the committee’s concerns about the MQTCQT as described in the meta-analysis (APA, 2011), nor is there any compelling reason to reject the straightforward conclusions in Cushman (2013). If Mr. Matte’s grandmother is right that “the proof of the pudding is in the eating,” then I’ve tasted enough. I think the pudding has spoiled.

References

- American Polygraph Association. (2011). Meta-analytic survey of criterion accuracy of validated polygraph techniques. *Polygraph*, 40(4), 193-305.
- Bradley, M. & Janisse, P. (1981). Accuracy demonstrations, threat, and the detection of deception: cardiovascular, electrodermal, and pupillary measures. *Psychophysiology*, 18(3), 307-315.
- Ekman, P. (2001). *Telling Lies: Clues to Deceit in the Marketplace, Politics and Marriage* (3rd ed.). New York, NY: W.W. Norton & Company, Inc.
- Iacono, W., (2008). Accuracy of polygraph techniques: Problems using confessions to determine ground truth. *Physiology & Behavior*, 95(1-2), 24-26.
- Iacono, W. G., & Patrick, C. J. (2014). Polygraph ("Lie Detector") Testing: Current Status and Emerging Trends. In I. B. Weiner, A. K. Hess, I. B. Weiner, & R. K. Otto (Eds.), *The Handbook of Forensic Psychology* (4th ed., pp. 552-588). Hoboken, New Jersey: John Wiley & Sons, Inc.
- Langston, W. (2011). *Research Methods Laboratory Manual for Psychology* (3rd ed.). Belmont, California, USA: Wadsworth.
- 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., & Reuss, R. (1989). Validation Study on the Polygraph Quadri-Zone Comparison Technique. Research Dissertation Abstract, LD 1452, Vol. 1502.
- 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.
- 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.