Pre-test Breathing Instructions Increase Perceptions of Respiratory Countermeasures

Walt Goodson,[1] Charles R. Honts, Mark Handler,[2] Raymond Nelson,[3] Matt Hicks, and Dennis Westerman[1]

Abstract

Thirteen experienced Texas Department of Public Safety examiners participated in a study of the potential effects of pre-test breathing instructions on examiners' impressions of respiratory countermeasures and artifacts in the polygraph recordings. Data were obtained from 60 field examinations, where half included a pre-test breathing instruction. Judgments of both respiratory artifacts and countermeasures were found to have low reliability. Nevertheless, a significant and undesirable impact of instructions on the perceptions of both artifacts and countermeasures was observed. Breathing instructions resulted in a greater proportion of blind evaluator reports of artifacts and countermeasures. Field practices, assumptions, and policies pertaining to pre-test instructions about breathing during testing should be considered in light of these data. These results underscore the importance of evidence-based field practice policies, and the potential hazards associated with well-intended field practice policies derived in the absence of evidence.

Keywords: Polygraph, Respiration, Countermeasures, Artifacts

Pre-test Breathing Instructions Increase Perceptions of Respiratory Countermeasures

An important assessment field examiners make is to visually inspect all test data during their evaluation. The goal of this visual inspection is to ensure that response data subjected to quantitative numerical or computer-based analyses are of adequate interpretable quality. Data that are distorted by instrumental or subject artifacts should be edited to remove the artifacts or the distortions should be excluded from subsequent analyses. A valid concern exists for the deliberate manipulation of data through the use of countermeasures during polygraph testing. However, numerous studies show that the spontaneous use of countermeasures is ineffective for guilty subjects (Honts, 1987; Honts, Amato, & Gordon 2001; Honts & Crawford, 2010; Honts & Hodes, 1983; Honts, Hodes, & Raskin 1985; Honts, Raskin, & Kircher, 1987; 1994; Honts, Raskin, Kircher, & Hodes, 1988). Moreover, those same studies clearly demonstrate that human polygraph experts are not effective at differentiating countermeasures from other artifacts when reviewing the physiological data from polygraph examinations (Honts, Amato & Gordon, 2001; Honts & Hodes, 1983; Honts,

[1] Texas Department of Public Safety
[2] Honts, Handler, and Hartwig, LLC
[3] Lafayette Instrument Company

Author Note
Please email correspondence to Walt Goodson at Walt.Goodson@dps.texas.gov.

Acknowledgements
The authors are grateful to Kyle Patrick for his thoughtful review and comments to an earlier draft of this paper. The views expressed in this article are solely those of the authors, and do not necessarily represent those of the Texas Department of Public Safety or the Lafayette Instrument Company.

Hodes & Raskin, 1985; Honts, Raskin & Kircher, 1994; Ogilvie & Dutton, 2008). It is notable that unassisted attempts by examiners to detect countermeasures consistently result in a substantial number of false positive errors (see the review by Honts, 2014).

There is some evidence that statistical and instrumental approaches can increase the accuracy of detection of physical countermeasure attempts (Honts et al., 1994; Raskin & Kircher, 1990; Elaad & Ben-Shakhar, 2009; Honts, et al, 1987; Ogilvie & Dutton, 2008; Stephenson & Barry, 1988). The ability to detect mental countermeasures remains an open question (Honts, 2014). An additional concern is raised by the fact that Honts and Crawford (2010) found that it was common for truthful test subjects to spontaneously produce respiratory patterns that appear identical to the countermeasure respiratory patterns taught in countermeasure strategy training materials (Maschke & Scalabrini, 2005; Williams, 2001). It is important to note that the patterns of respiratory countermeasures attributable to the countermeasure training materials have never been substantiated as valid indicators of countermeasure use in a scientific study. In fact, many of the respiratory patterns taught in these countermeasure training materials are derived from the training materials used in some polygraph schools (Department of Defense, 2004; Maschke & Scalabrini, 2005; Williams, 2001).

It was counterproductive for truthful subjects to engage in activities intended to enhance their polygraph test reactions with the goal of increasing their chances of passing (Rovner 1979 & 1986; Rovner, Raskin & Kircher, 1979). Those subjects actually increased their likelihood of being classified as deceptive. Countermeasure attempts by truthful subjects resulted in the production of more deceptive test scores (Honts, Amato, & Gordon, 2001). The use of countermeasure attempts to enhance the test results of truthful test subjects was not advisable (National Research Council, 2003).

Nevertheless, potential countermeasures, manipulation, and data integrity remain important concerns to field examiners and program managers. The most general conclusion at this time is that the spontaneous use of countermeasures is ineffective for guilty subjects and may reduce the accuracy of polygraph test results for the innocent (Honts, 2014). Trained countermeasures may reduce, and in some laboratory studies have substantially reduced, the accuracy of the polygraph with guilty subjects (Honts, 2014). The trained use of countermeasures is a potentially serious concern to polygraph programs involved in government and operational security screening, law-enforcement and public-safety screening, and convicted offender compliance screening. These concerns extend to the potential effects that countermeasures can have on diagnostic test results or the confidence consumers may place on those results.

Although statistical and computational methods have not been widely exploited to detect possible countermeasure attempts, the OSS-3 algorithm includes a procedural requirement to review data for interpretable data quality and mark any segments of data that contain apparent artifacts produced by movement or other problem activity (Nelson, Krapohl, & Handler, 2008). The OSS-3 algorithm will calculate the number and location of possible artifacts and calculate the statistical probability that observed artifacts have occurred due to random causes based upon the background variability. When the likelihood falls below an established boundary for statistical significance, the algorithm will alert the examiner to the possibility that a test subject may have attempted to systematically or intentionally alter the recorded physiological data. However, the effectiveness of the OSS-3 approach to countermeasures has never been demonstrated in a scientific study.

The American Polygraph Association now requires the use of activity sensors as a standard operating procedure (American Polygraph Association, 2012). This requirement may reduce the occurrence of discrimination errors attributed to data artifacts that should not be scored due to countermeasures or other causes. However, there are no validated criteria, scoring rules, or decision criteria for using the currently available movement sensors. Moreover, there are simply no studies of the reliability and validity of their use with known innocent, guilty, and countermeasure subjects. Until such data are obtained, caution must be exercised in making decisions about countermeasure use.

Among the potential countermeasure strategies recommended in countermeasure publications is the control or manipulation of one's breathing during testing (Maschke & Scalabrini, 2005; Williams, 2001). A primer on respiration responses and polygraph testing explores the effects that breathing has on polygraph testing (Handler, Reicherter, Nelson & Fausett, 2009). Because respiration data may be more easily subjected to voluntary control, field examiners have claimed that respiration data can provide clinically rich information regarding a test subject's behavior during testing. Subjects can attempt to suppress, augment, or distort their normal or natural breathing-related reactions to polygraph test stimuli. With the exception of simplistic efforts to distort the recorded test data beyond interpretation, skillful and effective use of breathing manipulation or countermeasure strategies to suppress or augment reactions may prove challenging. Success will require that a test subject simultaneously interact with the examiner and respond to the test stimuli in a convincing manner that will not arouse suspicion. The test subject must also mimic normal autonomic respiration patterns for the duration of the recorded examination and feign normal breathing movement responses to test target stimuli. It is important to note that there are no studies demonstrating that deliberate manipulations of respiration can successfully produce false negative outcomes.

Field examiners have become increasingly alert to the presence of “messy” or apparent artifacts in respiration data as potentially capable of influencing other recorded physiological responses (Department of Defense, 2004). As a counter-countermeasure technique, some field examiners have incorporated discussion, during the pre-test interview, to instruct or admonish test subjects against voluntary manipulation of breathing during testing. The pragmatic objective underlying the use of a pre-test breathing instruction may be two-fold. First, there may be an attempt to convince innocent test subjects to refrain from activities that may hinder their ability to produce non-deceptive test results. Additionally, examiners may believe they are discouraging deceptive subjects from engaging in countermeasures, alerting the subject that the examiner will be watching for countermeasure activities. No previous studies could be found regarding the use of pre-test breathing instructions but the use of pre-test countermeasure instructions were shown to have no effects (Honts, et al., 1994). This study was designed to investigate the hypothesis that the use of a pre-test breathing instruction would reduce the perception of both artifacts in respiration data and suspected countermeasures in the recorded physiological data.

Method

Participants

Two groups of examiners participated in this study. The polygraph data examined in the study were 60 actual polygraph examinations conducted by three experienced field examiners employed at the Texas Department of Public Safety. The second group of examiners was comprised of 13 experienced polygraph examiners, also employed at the Texas Department of Public Safety. The second group of examiners completed blind evaluations of the physiological data and made assessments of respiratory artifacts and countermeasures in general. The examiners were all male and had an average of seven years’ experience. All were licensed in the state of Texas, and all were members of at least one professional association. Examiners reported an average of 120 hours of continuing education during the preceding two years. Examination data were collected from 60 police applicants, who underwent polygraph testing prior to the pre-employment background investigation. Examinees were advised, in writing, that their test data would be reviewed by others for quality control and “other purposes.” Test subjects ranged in age from 22 to 42 years (M = 27 and Mdn = 26), with an average of three years of higher education. Forty-nine of the test subjects were male, and 11 were female. One test subject had previous law enforcement experience, and 18 had completed previous polygraph examinations.

Procedures

All polygraph examinations were completed using standard commercially available Lafayette instruments (Lafayette Instrument Company, Lafayette, IN). All examinations were conducted with the use of a movement sensor located in the seat of the test subject's chair. Examinations were conducted using the Directed Lie Screening Test, a format that is somewhat more easily standardized than other polygraph screening formats, and which has been shown to provide a reasonably good balance between test sensitivity and test specificity, of which both are significantly greater than chance (American Polygraph Association, 2011; Handler, Nelson, & Blalock, 2008; Handler, Honts & Nelson, 2013). All examinations consisted of a pre-test interview, test data recording phase, test data analysis phase, post-test interview, and quality control review.

Three examiners (data collectors) were identified who used pre-test respiration discussions during the physiology and instrument explanation portions of their pre-test interviews. The data collectors also provided specific breathing instructions to their test subjects prior to data collection. During those instructions, the data collectors specifically told their test subjects to *breathe normally* during the data collection and to avoid deep breaths. These same three data collectors were then instructed to refrain from providing test subjects with any discussion or instruction regarding respiration or breathing for all examinations for a period of time during the study. During the no instruction period the data collectors were instructed to refrain from any use of the words “breathing” or “respiration” to the extent possible. Discussion and explanation of the respiration sensors was kept to a minimum and test subjects were told the respiration sensors recorded movement in the upper and lower body areas. Data collectors were instructed that, if a test subject asked a specific question about breathing or respiration, they should answer in a manner that was factual, neutral, and designed to redirect the test subject's attention to the respiration sensors as movement sensors attached to the upper and lower body areas.

Case sampling from the archive was performed by an administrative assistant who had no knowledge of the purpose of the selection or study design. Thirty examinations were randomly selected from the archive of examinations conducted from a three-month window before the examiners were told to cease all breathing instructions and discussions. Thirty examinations were also randomly selected from a three-month period following the instructions to cease breathing instructions and discussions. Thus, a total of 60 examinations were randomly sampled from the examinations conducted by the data collectors during the six-month period of the study.

Recorded data from the 60 examinations were then subjected to analysis by the second group of 13 examiners (data evaluators). Data evaluators were asked to review and score all chart data and to then answer two additional “Yes/No” questions: 1) “Do the respiration data appear unusually messy or unstable?” and 2) “Are the data suspicious for attempts to fake or alter the test result through breathing manipulation?” It was not necessary, for the purpose of this study, to ascertain or confirm the guilt vs. innocence status of the sample examinations.

Results

Reliability

The “Yes/No” responses of the 13 data evaluators to the two test questions about countermeasures were entered into a spreadsheet. They were coded so that “Yes” responses were given the value of 1 and “No” responses were given the value of 0. Reliability across the 13 data evaluators was assessed with the *Intraclass Correlation Coefficient* (ICC: McGraw & Wong, 1996; Shrout & Fleiss, 1979). The ICC for the respiration artifact data was 0.379, with a 95% confidence interval of 0.288 to 0.49. This ICC was significantly larger than chance, $F(59, 808) = 10.064$, $p < .001$. The ICC for the countermeasure data was 0.343, with a 95% confidence interval of 0.253 to 0.455. This ICC was significantly larger than chance, $F(59, 649) = 8.433$, $p < .001$. Although these reliability values were greater than what would have been expected by chance, they were markedly less than the reliability normally considered necessary for measurements in applied setting. Normally an ICC of at least 0.8 is considered minimal when important decisions are being made (Graham, Milanowski, & Miller, 2012). The unreliability of decisions about respiration artifacts and countermeasures in this study is consistent with the literature on countermeasure detection by polygraph examiners where countermeasure decisions have consistently been found to be unreliable and invalid (Honts, 2014).

Study Questions

To consider the impact of breathing instructions on the examiners' perceptions of breathing artifacts and possible counter-measure use, we created new data vectors where each examiners' judgment was considered as independent. Table 1 shows the distribution of judgments of respiratory artifacts crossed with whether or not test subjects received a breathing instruction. The distribution of decisions about respiratory artifacts was different than what was expected by chance, *Chi-Square* (1) = 10.86, p = .001. An examination of Table 1 indicated that the presence of breathing instructions increased the frequency of perceived respiratory artifacts. Although the distribution was significantly different than chance, the relationship between breathing instructions and the perception of respiratory artifacts was modest, r (778) = 0.118, p = .001. Table 2 shows the distribution of judgments of countermeasures crossed with whether or not subjects received a breathing instruction. The distribution of decisions about counter-measures was different than what was expected by chance, *Chi-Square* (1) = 10.56, p = .001. An examination of Table 2 indicated that the presence of breathing instructions increased the frequency of perceived countermeasures. Although the distribution was significantly different than chance, the relationship between breathing instructions and the perception of countermeasure artifacts was modest, r (778) = 0.116, p = .001.

Table 1. Examiner judgments about the presence or absence of respiratory artifacts

	Respiratory Artifacts		
Breathing Instruction	Absent	Present	Totals
No	214 (56.0%)	176 (44.2%)	390
Yes	168 (44.0%)	222 (55.8%)	390
	382	398	780

Numbers in parentheses are column percentages.

Table 2. Examiner judgments about the presence or absence of countermeasures

	Countermeasures		
Breathing Instruction	Absent	Present	Totals
No	299 (53.7%)	91 (40.8%)	390
Yes	258 (46.3%)	132 (59.2%)	390
	382	398	780

Numbers in parentheses are column percentages.

Discussion

This study evaluated the relationship between the use of pre-test breathing instructions and subjective assessments of respiratory data instability (artifacts) and possible countermeasures during a blind review of the test data. The review data were collected from 13 experienced examiners who conducted evaluations on examination data collected by three different examiners who tested 60 police applicants. Both the assessments of respiratory artifacts and the assessments of countermeasures were found to have weak reliability. Our analyses showed a significant relationship between the perception of data problems, including respiratory artifacts and suspected countermeasure use, and the use of a pre-test breathing instruction with the instruction increasing the perceptions of data problems. The nature of this relationship was unexpected when one considers the purported objective(s) of the pre-test breathing instruction was to have test subjects to breathe normally and refrain from attempts to manipulate the test data. It is important to note that although the undesirable effects of breathing instructions were statistically reliable, the observed effect sizes for those effects were modest. Taken together with weak reliability results, the robustness of these results should not be assumed without replication.

Field experiments such as this are limited in their ability to provide unambiguous causal attributional relationships because optimal control could not be obtained for all testing conditions. Moreover, no attempt was made to study the exact wording of pre-test breathing discussion on the observed results. It is possible that some forms of discussion are more problematic than others, and such potential differences remain unstudied at this time. The generalizability of the results of our field experiment may also be limited by the setting, the examiners, and the relatively small number of polygraph examinations that were evaluated. Test subjects in this study were police applicants who can be assumed to possess a variety of different and strong characteristics motivating them to produce optimal test results. Some differences in motivation may, or may not, be expected when comparing this cohort of test subjects to other groups, such as criminal suspects, convicted offenders, or current government employees. It will be important to consider these possible variables in future attempts at replication before any strong conclusions can be made about the broad impact of breathing instructions.

A final limitation of this project is that blind judgments regarding data artifacts/ instability and possible countermeasures were made with an emphasis on unstructured clinical assessment procedures. It is thus perhaps not surprising that the assessments were unreliable. No use was made of statistical and/or automated methods to evaluate data quality or possible countermeasure use. Although statistical and computational methods have not been widely exploited to detect possible countermeasure attempts, the OSS-3 algorithm (Nelson et al., 2008) includes a procedural requirement to review data for interpretable data quality and mark any segments of data that may contain artifacts caused by movement or other problem activity. Additional research is needed to validate the OSS-3 approach. The use and development of statistical models should be included in future studies with an eye toward the development of validated statistical countermeasure algorithms.

Use of a pre-test breathing instruction was not supported by this study as contributing to improved data quality or a reduction of indications of possible countermeasures during testing. In fact, our findings suggest that breathing instructions have *negative impacts on data quality*. To our knowledge, this is the first study of this type, and future studies may be able to provide more information upon which to formulate more affirmative suggestions regarding field practices and the use of pre-test breathing instructions. Nevertheless, the results of this study provide insight into a fairly important issue of concern to field examiners and program managers. Examiners, supervisors, and program managers should be careful to ensure that field practices are based on evidence. A corollary to the need for evidence-based field practices is the need for caution around simplistic assumptions that adding discussions of breathing to the pre-test interview will reduce attention to breathing during testing.

References

American Polygraph Association. (2011). Meta-analytic survey of criterion accuracy of validated polygraph techniques. *Polygraph*, 40, 194-305.

American Polygraph Association. (2012). Model policy for law enforcement/public- service pre-employment polygraph screening examinations. Retrieved from American Polygraph Association. http://www.polygraph.org/section/resources/ download-policies-and-acts.

Department of Defense (2004). Test Data Analysis: DoDPI Numerical Evaluation Scoring System, August 2004. Downloaded, July 22, 2014 from https://antipolygraph.org/documents/dodpi-numerical-scoring-08-2004.pdf

Elaad, E., & Ben-Shakhar, G. (2009). Countering countermeasures in the concealed information test using covert respiration measures. *Applied Psychophysiological Biofeedback*, 34, 197-208. DOI 10.1007/s10484-009-9090-5

Graham, M., Milanoski, A., & Miller, J. (2012). Measuring and promoting inter-rater aggrement of teacher and principal performance ratings. Center for Educator Compensation Reform. Downloaded, July 18, 2014 from: http://files.eric.ed.gov/fulltext/ED532068.pdf

Handler, M., Honts C., & Nelson, R. (2013). Information gain of the directed lie screening test. *Polygraph*. 42(4), 192-202.

Handler, M., Nelson, R., & Blalock, B. (2008). A focused polygraph technique for PCSOT and law enforcement screening programs. *Polygraph*, 37(2), 100-111.

Handler, M., Reicherter, J., Nelson, R., & Fausett, C. (2009). A respiration primer for polygraph examiners. *Polygraph*, 38(2) 130-144.

Honts, C. R. (1987). Interpreting research on countermeasures and the physiological detection of deception. *Journal of Police Science and Administration*, 15, 204-209.

Honts, C. R. (2014). Countermeasures and credibility assessment. In, Raskin, D. C., Honts, C. R., & Kircher, J. C. Credibility assessment: Scientific research and applications: First Edition (pp. 131-158). Oxford, UK: Academic Press. http://dx.doi.org/10.1016/B978-0-12-394433-7.00003-8

Honts, C. R., Amato, S. & Gordon, A. K. (2001). Effects of spontaneous countermeasures used against the comparison question test. *Polygraph*, 30(1), 1-9.

Honts, C. R., & Crawford, M. (2010). Polygraph countermeasures cannot be detected from respiratory signatures: Government policy puts the innocent at risk. Paper presented at the 2010 meeting of the American Psychology Law Society, Vancouver, British Columbia, Canada.

Honts, C. R. & Hodes, R. L. (1983). The detection of physical countermeasures. *Polygraph*, 12(1), 7-17.

Honts, C. R., Hodes, R. L., & Raskin, D. C. (1985). Effects of physical countermeasures on the physiological detection of deception. *Journal of Applied Psychology*, 70, 177-187.

Honts, C. R., Raskin, D. C., & Kircher, J. C. (1987). Effects of physical countermeasures and their electromyographic detection during polygraph tests for deception. *Journal of Psychophysiology*, 1, 241-247.

Honts, C. R., Raskin, D. C., & Kircher, J. C. (1994). Mental and physical countermeasures reduce the accuracy of polygraph tests. *Journal of Applied Psychology*, 79, 252-259.

Honts, C. R., Raskin, D. C., Kircher, J. C., & Hodes, R. L. (1988). Effects of spontaneous countermeasures on the physiological detection of deception. *Journal of Police Science and Administration*, 16, 91-94.

Maschke, G. W., & Scalabrini, G. J. (2005). The Lie Behind The Lie Detector. Retrieved July 24, 2014, from https://antipolygraph.org/lie-behind-the-lie-detector.pdf -detector.pdf

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, 1, 30-46.

National Research Council (2003). The Polygraph and Lie Detection. Washington, DC: The National Academies Press.

Nelson, R., Krapohl, D. J. & Handler, M. (2008). Brute force comparison: A monte carlo study of the Objective Scoring System Version 3 (OSS-3) and human polygraph scorers. *Polygraph* , 37(3), 185-215.

Ogilvie, J., & Dutton, D. W. (2008). Improving the detection of physical countermeasures with chair sensors. *Polygraph*, 37(2), 136-148.

Raskin, D.C. & Kircher, J.C. (1990). Development of a computerized polygraph system and physiological measures for the detection of deception and countermeasures: A pilot study. (Preliminary Report Contract 8-L655300-000). Salt Lake City, UT: University of Utah.

Rovner, L.I. (1979). The effects of information and practice on the accuracy of physiological detection of deception. *Dissertation Abstracts International.* (University Microfilm # AAD80-05308).

Rovner, L.I. (1986). The accuracy of physiological detection of deception for subjects with prior knowledge. *Polygraph*, 15(1), 1-39.

Rovner, L. I., Raskin, D. C., & Kircher, J. C. (1979). Effects of information and practice on detection of deception. *Psychophysiology*, 16(2), 983.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86, 420-428.

Stevenson, M., & Barry, G. (1988), Use of a motion chair in the detection of physical countermeasures. *Polygraph*, 17(1), 21-27.

Williams, D. (2001). *How to Sting the Polygraph*. Norman, OK, U.S.A: Sting Publications.