**Effect of the Location of the Numbers Test on Examiner Decision Rates in Criminal Psychophysiological Detection of Deception Tests**

**Richard Widup and Gordon H. Barland**

**Department of Defense**
Polygraph Institute
Building 3165
Fort McClellan, AL 36205-5114

**Department of Defense**
Polygraph Institute
Building 3165
Fort McClellan, AL 36205-5114

**Abstract**

The U.S. Army Criminal Investigation Command (USACIDC) conducted 251 psychophysiological detection of deception (PDD) examinations on suspects of criminal offenses in which a numbers test (stim) was conducted between the first and second tests of the main test, as is their standard procedure. Another 231 examination were conducted in which the numbers test was conducted prior to the first test of the main test series. Study results suggest the location of the numbers test had no effect on the inconclusive rate or the number of deception indicated (DI) and no deception indicated (NDI) decisions.
Effect of the Location of the Numbers Test on Examiner Decision Rates in Criminal Psychophysiological Detection of Deception Tests

Special Agent Richard Widup
U.S. Army Criminal Investigation Command

and

Gordon H. Barland, Ph.D.

March 1994
Department of Defense Polygraph Institute
Fort McClellan, Alabama 36205-5114
Telephone: 205-848-3803
FAX: 205-848-5332
Effect of the Location of the Numbers Test on Examiner Decision Rates in Criminal Psychophysiological Detection of Deception Tests

Special Agent Richard Widup
U.S. Army Criminal Investigation Command

and

Gordon H. Barland, Ph.D.
Department of Defense Polygraph Institute

March 4, 1994
Director's Foreword

This is another study wherein a field practitioner served as the principle investigator (PI) in conducting a field research project. One of the technical issues in psychophysiological detection of deception (PDD) has to do with the use of a numbers test in the test sequence. Some examiners use it before the main test series; others use it after the first test of the main test series; and others do not use it at all.

The results of this study supports the position that it does not matter if the numbers test is used before or after the first test of the main test series as regard to an effect on the inconclusive rate or the number of deception indicated (DI) or no deception indicated (NDI) diagnostic decisions.

This study did not, however, address the issue as to the value of the numbers test as regard to whether such a test actually improves the quality of the physiological data collected and if that in turn produces higher diagnostic accuracy. In other words, does a numbers test really do what PDD examiners claim it will do? An analog study, where ground truth is known, will be required to answer this important question.

William J. Yankee, Ph.D.
Director
Acknowledgments

The authors wish to thank Special Agent Joe Gallimore for conceiving the study, and the forensic psychophysicists of the United States Army Criminal Investigation Command (USACIDC) for conducting the examinations. This study was funded by the Department of Defense Polygraph Institute as project number DoDPI90-P-0006. The views expressed in this article are those of the authors and do not reflect the official policy or position of the Department of Defense or the U.S. Government.
Abstract

WIDUP, R. and BARLAND, G. H. Effect of the location of the numbers test on examiner decision rates in criminal psychophysiological detection of deception tests. March 1994, Report No. DoDPI94-R-0015. United States Army Criminal Investigation Command and Department of Defense Polygraph Institute, Fort McClellan, AL 36205. The U.S. Army Criminal Investigation Command (USACIDC) conducted 251 psychophysiological detection of deception (PDD) examinations on suspects of criminal offenses in which a numbers (stim) test was conducted between the first and second tests of the main test series, as is their standard procedure. Another 231 examinations were conducted in which the numbers test was conducted prior to the first test of the main test series. Study results suggest the location of the numbers test had no effect on the inconclusive rate or the number of deception indicated (DI) and no deception indicated (NDI) decisions.

Key-words: stim test, numbers test, psychophysiological detection of deception (PDD), polygraph, forensic psychophysiology
Executive Summary


Many examiners include a "numbers test" when they conduct a psychophysiological detection of deception (PDD) examination on a suspect of a criminal offense. In that test, the subject chooses one of several numbers and is instructed to lie when asked about the chosen number. The examiner can then show the subject how his physiological response, when he lied, was recorded by the polygraph. The purpose of the numbers test is to demonstrate how reactive the body is to even an inconsequential lie. It is generally believed that such a demonstration minimizes the possibility of errors by reassuring the innocent-but-nervous subject, thereby reducing the magnitude of the physiological response to the relevant questions. On the other hand, the demonstration is believed to have the opposite effect on the guilty person. By reinforcing the guilty subject's fear of detection, response to the relevant questions should be enhanced.

The numbers test traditionally has been conducted between the first and second iterations of the crime questions, because the examiner wished to see if the subject's level of physiological responding decreased or increased as a result of the numbers test. Recently, examiners have argued that if the numbers test actually reduces errors, it should be employed at the very outset of the examination, prior to the first iteration of the crime questions. This study used real-life PDD examinations to examine the effect of the numbers test location on the inconclusive rate and the rates for decisions of deception indicated (DI) and no deception indicated (NDI).

Forensic psychophysiologists of the U.S. Army Criminal Investigation Command (CID), conducted 251 PDD examinations on suspects in criminal investigations, in which the numbers test was conducted between the first two iterations of the crime questions, as is their standard procedure. The examinations of another 231 suspects of criminal offenses were conducted with the numbers test being used prior to administering the first test of the main test series. Because ground truth was not known in most cases, no conclusions can be drawn about the effect of the location of the numbers test on the accuracy of decisions.
As tested, the location of the numbers test had no effect on the inconclusive rate or the number of decisions of DI and NDI. Additional research is needed to determine what, if any, effect the presence or absence of the numbers test has on the quality of the data obtained during a PDD examination and the accuracy of decisions based on that data.
# Table of Contents

Title Page ................................................................. i
Director’s Foreword ......................................................... ii
Acknowledgments .............................................................. iii
Abstract ................................................................. iv
Executive Summary ......................................................... v
List of Tables ............................................................... vii
Introduction ................................................................ 1
Procedure ................................................................. 3
  Subjects ................................................................ 3
  Examiners ................................................................ 3
  Equipment ................................................................ 3
  Procedure ................................................................ 4
Results ........................................................................... 5
References ..................................................................... 10
List of Tables

1. Frequency of examiners’ trichotomous decisions. . . . . . . 6
2. Frequency of examiners’ decisions versus inconclusive results. . . . . . . . . . . 6
3. Frequency of examiners’ decisions on zone comparison test (ZCT) examinations. . . . . . . . . . . . . . . . 7
4. Numerical scores at end of three tests on ZCT examinations. . . . . . . . . . . . . . . . . . . . . . 7
5. Numerical scores at end of first crime series test. . . . . 8
Some examiners advocate the use of a numbers test because it can be helpful in setting control questions and in post-test interrogations (e.g. Barland, 1978; Raskin, 1988, p. 257). Other examiners believe that a numbers test, designed to impress an examinee with the psychophysiological detection of deception's (PDD) effectiveness of responses, increases the accuracy of the PDD tests (e.g. Abrams, 1978). They reason that the innocent-but-nervous suspect is reassured no error will occur, thereby reducing concern for the relevant questions while increasing concern regarding the control questions. The guilty suspect, on the other hand, who may have thought the test could be beaten, becomes more concerned about the relevant questions. Because the numbers test is believed to enhance the differential reactivity of innocent and guilty suspects, the U.S. Army Criminal Investigative Command requires that it be used when administering PDD tests during the conduct of criminal investigations.

Several studies support this view. Using 30 verified field cases, Senese (1976) had examiners analyze the physiological data recorded on the first test, which had been administered prior to a numbers test. A month later, the examiners analyzed the physiological data recorded on the second crime series test, which had been administered after a numbers test. He found that the inconclusive rate and error rate were lower following the numbers test. Unfortunately, this study did not include a control group which had not been administered a numbers test. It is therefore impossible to determine whether the increase in effectiveness was caused by the numbers test or by a reduction in the subjects' general level of anxiety as they became more familiar with the examination setting.

In a study designed to avoid that problem, Bradley and Janisse (1981) conducted three numbers tests prior to a control question test in a mock crime study. They manipulated the outcomes of the numbers tests so that the selected numbers were detected on none, one, two, or three of the tests. The accuracy of the subsequent control question test was generally proportional to the number of correctly called numbers tests.

However, most studies have found little, if any effect, of the numbers test on the accuracy of the PDD test. Gustafson & Orne (1963) found that administration of a numbers test had virtually no effect on the accuracy of subsequent tests. Barland & Raskin (1975) and Timm (1982) used similar designs in a mock crime paradigm. Some subjects were administered a numbers test that was correctly called; others had the numbers test deliberately miscalled to weaken their confidence in the PDD test; a control group had no numbers test administered. Both studies found the PDD tests to be equally effective with all three groups.
In some studies, a numbers test decreased the accuracy of subsequent tests. Ellson et al. (1952) conducted two numbers tests, one after the other. Subjects who were told that their first test had been correctly detected showed a decrease in accuracy on the second test. The accuracy decreased from 82% on the first test to 27% on the second test, compared to an accuracy of 70% on the second test for a control group who had not been informed about how they did on the first test. Gustafson & Orne (1963) found a similar effect with subjects who had a psychological need to be detected. When informed that their deception on the numbers test had been detected, they apparently lost interest and the subsequent test was less accurate.

Bradley & Janisse (1979) found an inverted U-shaped relationship between demonstrations of PDD test accuracy and the accuracy of a subsequent test. As predicted by Davis (1961), the accuracy of the PDD test decreased when subjects were led to believe that the PDD test was either extremely accurate or extremely inaccurate.

Kirby (1981) conducted one of two types of numbers tests between the first and second asking of the questions. He found that the accuracy of blind evaluations of the psychophysiological data following the numbers test was somewhat lower than the initial test. He suggested that the numbers test should perhaps not be used when the first test showed clearly truthful or deceptive patterns. However, his design failed to control another variable that might have affected the results. It could be that the first test is more accurate regardless of whether a numbers test is employed or not.

Following the established convention (Decker, 1978), the Department of Defense Polygraph Institute (DoDPI) has taught that the numbers test, when used, is usually placed after the first test of the main PDD test series. The rationale is that if the first test shows no differentiation between relevant and control questions, the examiner should stimulate the examinee by any of several methods, one of which is a numbers (stim) test. Other examiners have suggested that, if the numbers test is so important, it should be the initial test conducted, to minimize the possibility of having an inconclusive test on the main test.

The purpose of this study is to determine whether the numbers test is more effective, in terms of reducing inconclusive results, when conducted before or after the first test of the main test series in a PDD examination of suspected criminals.
Procedure

Subjects

The subjects were all the suspects of at least one criminal offense that were examined by the U.S. Army Criminal Investigation Command (USACIDC) in the three months from 1 January to 1 April 1990, for a total of 482 persons. All were U.S. citizens, and included a mixture of U.S. Army personnel, Department of Army (DA) civil service employees, and their dependents. Information is not readily available regarding the proportion of subjects that during the conduct of the investigation, were suspects, victims, or witnesses; the types of crimes; the number of males and females, or their ages.

Examiners

There were 37 examiners, all of whom were trained at the Department of Defense Polygraph Institute (DoDPI) or its predecessor, the U.S. Army Polygraph School, and were certified as competent to administer PDD examinations by the USACIDC. Two of the examiners were female. The examiners’ ages ranged from 31 to 47, with a mean of 41. Their experience as examiners ranged from 1 to 16 years, with a mean of 3 years. The number of specific issue examinations they had conducted at the outset of the study ranged from 1 to 1,549, with a mean of 245 criminal examinations. Following standard procedures, all examinations were reviewed by the CID Command Quality Control Office in Baltimore, MD.

Equipment

All examinations were conducted with field standard five-channel Lafayette and Stoelting polygraphs. Two respiratory channels, one electrodermal, and two cardiovascular channels were recorded.

Respiration was recorded by pneumatic bellows on the examinees’ thorax and abdomen. One channel, usually the thoracic, was electronically amplified; the other used a mechanical pneumograph component.

The electrodermal channel recorded skin resistance by means of two stainless steel electrodes placed on the second and fourth fingers of the subject’s hand, contralateral to the cardiovascular cuff. No electrolyte was used.

The cardiovascular channels consisted of a single cardiovascular cuff, usually located on the examinee’s upper arm, that was inflated to about 70 mmHg, and provided input to both cardiovascular channels by means of a Y-shaped connector. One of the cardiovascular channels mechanically recorded the pressure changes; the other transduced and electronically amplified the signal prior to recording.
Procedure

The manner in which the numbers test was explained to the examinee was left to the discretion of the individual examiner. Although all examiners learned a standard procedure at the DoDPI, with experience they typically develop an individualized style of administration. No attempt was made to standardize the administration of the numbers test just for this study. A typical numbers test might be conducted as follows.

"John, I am now going to conduct a test in which I want you to answer 'No' to all of the questions. The main purpose of this test is to determine if you are physiologically capable of being administered a PDD examination today. The way in which this test will be conducted is that I want you now to pick a number between 3 and 8, that is not either 3 or 8. Now, have you picked your number? Good. Now, please tell me your number." (The examinee answers that his number is 5.)

"Okay, John. As I recall, you are right handed, correct? Now, I want you to take this pen, and with your left hand write the number you picked on this piece of paper, somewhere in the middle. Please write the number a couple of lines high. Now, I am going to surround the number you wrote with numbers that I will write. I am now going to place this paper on the wall in front of you."

"We will now go over the way in which this test will be conducted. Keep in mind that I want you to answer 'No' to all the questions. The test will go this way: Regarding the number you wrote on that piece of paper, did you write the number 2? (The examinee answers 'No.') Did you write the number 3? (Answer: No.) Did you write the number 4? (Answer: No.) Did you write the number 5?" (Answer: No.)

"John, what is that answer? (The subject answers that the response was a lie.) That's right, John; that is a lie. Now, both you and I know that you wrote the number 5, correct? In fact, I saw you take your left hand and write that number on that piece of paper. I want you to answer 'No' nonetheless, and I will tell you why after this test is over. Did you write the number 6? (Answer: No.) Did you write the number 7? (Answer: No.) Did you write the number 8? (Answer: No.) Very good, John. Now I am going to start at the top of the list and go to the bottom, and I want you to follow along with me with your eyes, answering each question with a 'No.' John, do you have any questions? Now I will start the test."

At this point the test is administered. Immediately after the numbers test, the examiner shows the chart to the subject, pointing out the responses which occurred at the selected number. The examiner asks the subject what caused those responses. Most subjects reply that they do not know, whereupon the examiner
replies "John, these responses occurred when you lied about the number you wrote. Now, lying about a number is normally not a big thing, is it? But it was to you, John, because you saw how big those responses were. Now the other questions on the main test are more important, aren’t they? John, you can imagine how large those responses would be if you are lying to me about any or those, can’t you? That’s the reason you need to be 100% truthful with me here today."

USACIDC examiners assigned to the USACIDC third and sixth regions, located in the United States, Puerto Rico, and Panama, conducted the numbers test before the first test of the main test series. Examiners from USACIDC’s second and seventh regions, located in Korea and Europe, conducted the numbers test in the traditional location following the first crime issue test. All other parts of the examinations were conducted according to the standard USACIDC procedures. During the pretest interview the examiner advised the examinees of their rights; obtained their military, medical, and educational history; asked them to explain what they knew about the crime under investigation; and reviewed the test questions.

The PDD examination consisted of three repetitions of the test questions using either a zone comparison test (ZCT) or a modified general questions test (MGQT) format. The decision as to which examination procedure to use was left to the discretion of the examiner, who took into account the number of issues to be resolved and the nature of the case. In a few cases, where the examination results, after three tests were inconclusive, a fourth test was conducted. In one case in which the numbers test was the first test, the suspect confessed after the first crime test had been conducted.

Results

Of the 482 PDD examinations in this study, the results of the first series (3-4 crime issue tests) were inconclusive in 92 (19%) of the cases. Table 1 compares the PDD outcomes as a function of the location of the numbers test. When the numbers test was first, 21% of the exams were inconclusive, 65% were deception indicated (DI), and 14% were no deception indicated (NDI). These compare to 17% inconclusive, 62% DI, and 21% NDI when the numbers test was conducted after the first crime question test.
Table 1
Frequency of Examiners' Trichotomous Decisions

<table>
<thead>
<tr>
<th></th>
<th>DI</th>
<th>NDI</th>
<th>INC</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Numbers Test 1st</td>
<td>149</td>
<td>33</td>
<td>49</td>
<td>231</td>
</tr>
<tr>
<td>Numbers Test 2nd</td>
<td>156</td>
<td>52</td>
<td>43</td>
<td>251</td>
</tr>
<tr>
<td>Total</td>
<td>305</td>
<td>85</td>
<td>92</td>
<td>482</td>
</tr>
</tbody>
</table>

Whenever we obtain two samples, we expect to see differences between the two sets of numbers. For example, if we were to repeat this study, we would expect that the next two samples would differ somewhat from the present samples. These chance differences are called sampling errors. The question thus arises as to whether the differences between the samples are so small that they are likely to be sampling errors, or whether they are so great that they are unlikely to have arisen simply by chance.

One way to determine the significance of the differences is to use the chi square ($\chi^2$) test. Two chi square tests were applied. The first test assessed the significance of the difference in the inconclusive rates. To do this, the first two columns from table 1 were combined, as shown in table 2.

Table 2
Frequency of Examiners' Decisions Versus Inconclusive Results

<table>
<thead>
<tr>
<th></th>
<th>DEC</th>
<th>INC</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Numbers Test</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First Location</td>
<td>182</td>
<td>49</td>
<td>231</td>
</tr>
<tr>
<td>Second Location</td>
<td>208</td>
<td>43</td>
<td>251</td>
</tr>
<tr>
<td>Total</td>
<td>390</td>
<td>92</td>
<td>482</td>
</tr>
</tbody>
</table>

Dec = number of decisions.
Inc = number of inconclusives.

The chi square test was not significant, $\chi^2(1, N = 482), = 1.05, p = .31$, which means that there is no evidence that the location of the numbers test affected the inconclusive rate. The second chi square test looked at all of the data in table 1. It also was not significant, $\chi^2(2, N = 482), = 3.98, p = .14$. The location of the numbers test had no detectable effect on the number of DI, NDI, or inconclusive outcomes.
The above analysis used only the examiner's decisions at the end of the first series. It is possible that more detailed analysis of the numerical scores upon which the decisions were based could reveal differences caused by the location of the stim test. The Crimes Record Center (CRC) forwarded copies of the examiners' score sheets to the DoDPI for detailed analysis. We used only the ZCT score sheets for these additional analysis because of the greater likelihood that they would be single issue tests in which the subjects would be either lying to all test questions or to none of them.

We input 218 of the 221 sets of ZCT score sheets into the computer. The three remaining score sheets could not be located. To see if this subsample was different from the total population, we repeated the $X^2$ test on the distribution of examiner decisions as a function of the location of the numbers test (Table 3). As before, there was no significant difference in the distribution of test outcomes as a function of whether the numbers test was before or after the first crime chart; $X^2(2, N = 218) = 2.76$, $p = .25$.

Table 3
Frequency of Examiners' Decision on ZCT examinations

<table>
<thead>
<tr>
<th></th>
<th>DI</th>
<th>NDI</th>
<th>INC</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Numbers test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>location</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First</td>
<td>92</td>
<td>11</td>
<td>27</td>
<td>130</td>
</tr>
<tr>
<td>Second</td>
<td>53</td>
<td>14</td>
<td>21</td>
<td>88</td>
</tr>
<tr>
<td>Total</td>
<td>145</td>
<td>25</td>
<td>48</td>
<td>218</td>
</tr>
</tbody>
</table>

We next proceeded to examine the numerical scores that underlie the decisions. Table 4 presents the numerical scores at the end of the third chart.

Table 4
Numerical Scores at End of Three Tests on ZCT Examinations

<table>
<thead>
<tr>
<th></th>
<th>DI</th>
<th>NDI</th>
<th>INC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Numbers test</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>location</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First</td>
<td>-9.3</td>
<td>11.8</td>
<td>3.1</td>
</tr>
<tr>
<td>Second</td>
<td>-6.8</td>
<td>14.2</td>
<td>3.1</td>
</tr>
</tbody>
</table>
DI outcomes produced bigger scores when the numbers test was given first, whereas NDI outcomes had larger scores when the numbers test was given second. An analysis of variance found that these differences were not significant (F = 1.03, p = .35), meaning that the location of the numbers test had no apparent effect on the total numerical scores.

Did the numbers test affect the scores on the first crime series test? To answer this question, we looked at the scores at the end of the first test separately for DI and NDI subjects.

Table 5
Numerical Scores at End of First Crime Series Test

<table>
<thead>
<tr>
<th></th>
<th>DI</th>
<th>NDI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Numbers test</td>
<td></td>
<td></td>
</tr>
<tr>
<td>location</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before</td>
<td>-3.0</td>
<td>3.2</td>
</tr>
<tr>
<td>After</td>
<td>-1.8</td>
<td>4.4</td>
</tr>
</tbody>
</table>

DI subjects had a more deceptive first test when the numbers test had been administered first. This difference was significant (pooled variances T = -2.274, p < .025). Unfortunately, subjects called NDI also tended to produce a less truthful first test when the numbers test had been conducted beforehand, although this difference was not significant, meaning that it could have been caused by sampling error.

Discussion

The results indicate that the location of the numbers test has no apparent practical effect on the distribution of examiners’ decisions. This, in turn, suggests that the location of the numbers test could reasonably be left to the examiner’s discretion, as it does not seem to effect the inconclusive rate or the type of decision.

This finding must be viewed with some caution, however. The lack of ground truth and experimental control in real-life cases makes it difficult to draw firm conclusions, just as trying to generalize from laboratory findings to field conditions is also risky. Although no errors are known to have occurred, there is no way of knowing how many errors may have gone undetected. Moreover, possible differences in examiner expertise, the types of crimes being investigated, or the proportion of guilty and innocent suspects (as distinct from DI, NDI, and inconclusive outcomes) between the regions confound the interpretation of the analyses. A major confound (nonrandom assignment of examiners to the treatment condition) should be controlled in future studies.
The question of what effect the inclusion of the numbers test has on PDD accuracy and inconclusive rates could not be answered by this study, because one was administered on every test. It may be that a numbers test is a waste of time when conclusive results are being obtained, which is usually the case. The bulk of the literature indicates that it has little effect on the accuracy of the examination outcomes, at least in mock crime analogs. It is recommended that a field study, designed jointly by the DoDPI and the field user, be conducted to assess the advisability of requiring a numbers test in all criminal examinations.
References


