

Polygraph

VOLUME 30

2001

NUMBER 3

Contents

Fiscal Year 2000 Report to Congress on the Department of Defense Polygraph Program	153
False Rape Allegations Eugene J. Kanin	163
The Rank Order Scoring System: Replication and Extension with Field Data Donald J. Krapohl, Donnie W. Dutton, and Andrew H. Ryan	172
Spies and Lies John F. Sullivan	182
Electrodermal Activity: A Literature Review Chauncey E. Farrell	185
A Concise History of the Comparison Question James F. Waller	192
Psychological Set: Its Origin, Theory and Application James Allan Matte and Robert Nelson Grove	196
A Brief Rejoinder to Matte & Grove Regarding "Psychological Set" Donald J. Krapohl	203
A Belated Look at Symptomatic Questions Donald J. Krapohl and Andrew H. Ryan	206
A Response to Krapohl & Ryan's "Belated Look at Symptomatic Questions" Cleve Backster	213
Comments on Krapohl & Ryan Criticism of Capps, Knill & Evans Research James Allan Matte	216
Final Comment on the Belated Look at Symptomatic Questions Donald J. Krapohl and Andrew H. Ryan	218
Reply to Rejoinder by Donald J. Krapohl and Andrew H. Ryan James Allan Matte	220
Book Review	223
Closing Editorial	225

Published Quarterly

© American Polygraph Association, 2001

P.O. Box 8037, Chattanooga, Tennessee 37414-0037

A Belated Look at Symptomatic Questions

Donald J. Krapohl and Andrew H. Ryan

Abstract

Backster (1962) suggested that unresponsiveness during polygraph testing was often caused by the examinee being apprehensive about issues outside the scope of the test questions. These concerns drew the examinee's attention away from the diagnostic questions, thereby diminishing differential responsivity, and giving rise to inconclusive polygraph outcomes. Backster named this phenomenon the "super-dampening concept", indicating that all responses during testing were dampened by the examinee's diversion of attention to outside concerns. Backster recommended the use of symptomatic questions, which he reported were useful to identify when the examinee was more worried about an outside issue than the test issues. In the present project, 100 field polygraph cases were subjected to 7-position scoring and an automated form of the Rank Order Scoring System (ROSS, Honts & Driscoll, 1987) to investigate the relationship between dampened responsiveness to diagnostic questions, as represented by polygraph scores, and responses to symptomatic questions. Consistent with previous research, reactions to symptomatic questions had no correlation with the strength of polygraph scores in either the manual 7-position scorings or the automated ROSS. The predicted super-dampening effect was not found. The future of symptomatic questions is discussed.

Key words: outside issue, Rank Order Scoring System, super-dampening concept, symptomatic question, validity

One of the abiding questions for the field of polygraphy is why examinees occasionally fail to react to polygraph questions to which they are lying. Even when the test conditions are ideal, the examination is competently conducted, the examinee appears to be suitable, and the test issues are clearly defined, there are instances where examinees do not react physiologically during polygraph testing. There are possibly as many theories for this phenomenon as there are writers on the topic. Proposed causes for this unresponsiveness have been adrenal exhaustion (Reid & Inbau, 1966), emotional

lethargy (McInerney, 1961), and even a fatalistic attitude or a lack of a sense of guilt (Arther, 1977). Despite these suggestions from the field, there is no research known to the present writers that directly answers this question.

Cleve Backster (1962) proposed a theory based on attention. Called the "super-dampening" concept, Backster asserted that examinees were often distracted by issues outside of the polygraph session, which can limit their attention to the polygraph questions. He wrote:

Acknowledgements

We are grateful for the assistance of DoDPI Instructor William Gaschler for scoring the cases used in this project, and to Kay Williams for data management. We also appreciate the methodological guidance and critical manuscript review of Dr. Dean Pollina and Dr. Stuart Senter of DoDPI.

Donald J. Krapohl is a Researcher with DoDPI, and Andrew Ryan Ph.D. is the Chief of DoDPI's Research Division. The opinions expressed in this article do not necessarily represent those of the Department of Defense or the US government. Reprint requests should be directed to: Donald J. Krapohl, PO Box 10411, Ft. Jackson, SC 29207, USA, or e-mail to dkrapohl@aol.com.

Based on all indications to date it seems apparent that the major portion of "inconclusive" polygraph examination results are not caused by a subject who is a non-reactor at that time, but are caused by the sensory perception "set" of the subject being diverted or tuned in on the wrong protective frequency. It now seems apparent that this is caused by the subject's anticipation of a question involving some outside factor. With such a person the outside issue, about which he is so apprehensive, is much more important to him - or more directly affects his well-being - than does the reason for the polygraph examination, thus causing a "super-dampening" of all responses that ordinarily would have occurred (Backster, 1962, p 65).

Based on his observations, Backster concluded that suppressed responsivity, and the resulting inconclusive outcome, is largely attributable to outside issues competing for an examinee's attention. As a remedy, Backster advocated the inclusion of additional questions, called "symptomatics," among the other polygraph questions. The more common phrasing of symptomatic questions are: "Are you completely convinced that I will not ask you a question during this chart that has not already been reviewed?" and "Is there something else you are afraid I will ask you a question about even though I told you I would not?" Backster warned that when responses occurred to these symptomatic questions, "the subject's outside issue is causing an inter-play on his charts" (Backster, 1962, p 67). In other words, it was an indication that the examinee was overly concerned that the examiner would inquire about other topics beyond those on the reviewed question list, and that the examinee might not react to the other diagnostic questions.

Symptomatic questions are now found in all of the major techniques that are offsprings of the Backster Zone Comparison Technique (ZCT): the Department of Defense Polygraph Institute (DoDPI) ZCT, Matte Quadri-Track, and the Utah ZCT. Though the logic of the super-dampening is credible, empirical data to support the use of symptomatic questions, or even verify the

underlying super-dampening concept is surprisingly meager. Capps, Knill and Evans (1993) set out to test the hypothesis that the mere inclusion of the symptomatic question would reduce the incidence of inconclusive results. They directed three field polygraph examiners to alternately conduct their live criminal examinations with symptomatic questions, and the next examination without symptomatic questions. Capps et al (1993) collected 75 field cases with symptomatic questions, and 75 without symptomatic questions. Based on the original examiner scores, they found that among the symptomatic question cases, there were 4 inconclusive results from manual scoring, whereas the cases without symptomatic questions had 12 inconclusive results. Using an early version of the PolyScore algorithm, the numbers were 5 and 8, respectively. Capps (1993) subjected the manual scores to a chi-square statistical treatment, and concluded that there was a significant effect for the presence of symptomatic question. They reported that their data supported Backster's claim that symptomatic question would significantly reduce the number of inconclusive calls. These findings have been touted by proponents as evidence of the efficacy of the symptomatic question (Matte, 1996; 2000).

It should be noted that in the Capps et al (1993) study there were only a relative handful of inconclusive findings by the original examiners, merely 4 and 12, and as such, subject to more variability than if the sizes had been more substantial. This does not disprove their conclusions, of course, but very small sample sizes are notoriously unstable, and generalizing their findings to the hundreds of thousands of polygraph examinations conducted each year would be premature. Moreover, had the PolyScore decisions been used instead of the examiner calls, the outcome would have been very different. Using the Capps (1993) data for the PolyScore decisions, we conducted a chi-square for with-symptomatics and without-symptomatics for two types of decisions: conclusive, and inconclusive. Those results did not indicate a significant difference in decisions for the presence or absence of the symptomatic question ($X^2 (1)=0.758, p>.05$). In contrast to the manual scores, the PolyScore decisions were unaffected by the symptomatic questions.

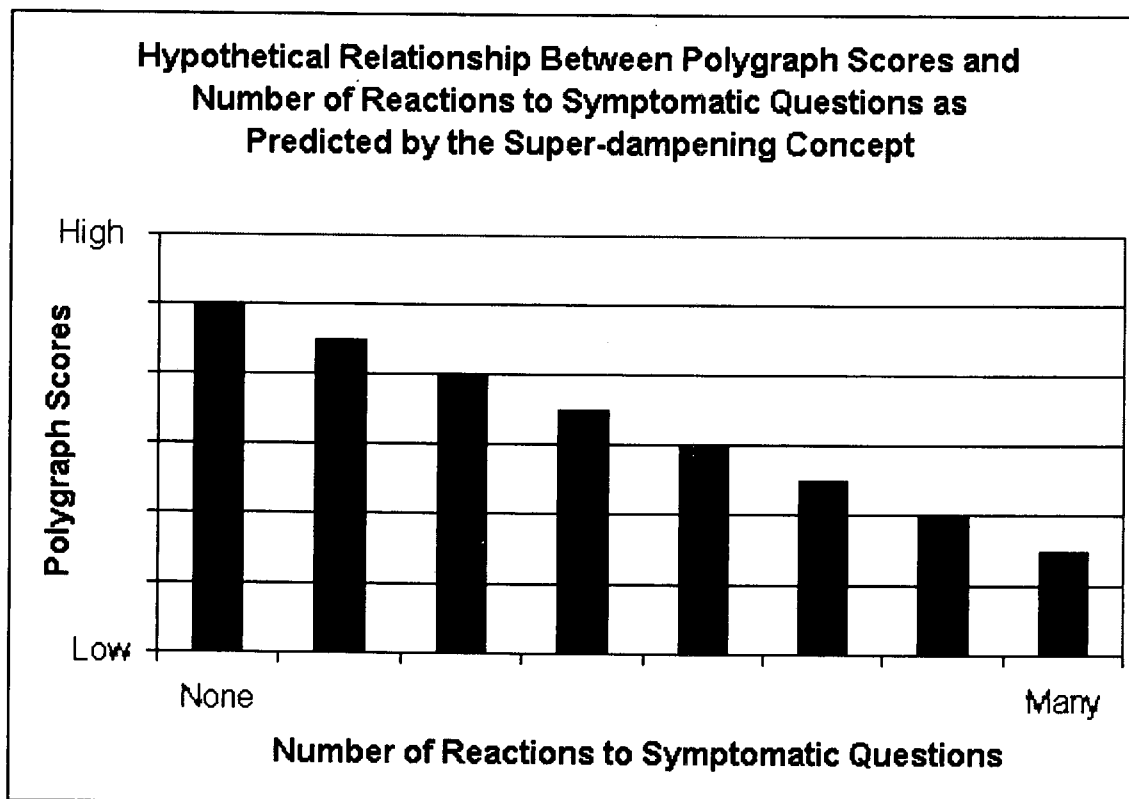
Since this automated algorithm could not be influenced by anything other than the physiological data, we are more inclined to accept these results than those produced by the examiners who interacted with the examinee, conducted the polygraph session, may have known the study hypothesis, and did the manual scoring. Conclusions based on these manual scorings may be confounded by the conditions under which the scores were produced. As a final note, the Capps (1993) manual scoring data were reanalyzed by Honts, Amato, and Gordon (2000), who found that even if one accepted the original examiner scores, the effect of the symptomatic question was quite modest, accounting for only 3% of the variance.

The only other direct test of the symptomatic question was conducted by Honts, Amato, and Gordon (2000). They polygraphed volunteer subjects in a 2 (guilty, innocent) X 2 (outside issue present, absent) X 2 (outside issue question present, absent) between-subjects mock-crime design, with 24 subjects in each condition. In his study, Honts et al found that the symptomatic

question showed no validity for detecting or reducing the effects of outside issues. Statistical analyses of the raw physiological data found little useful information in the reactions to the outside issue questions. He concluded that the power of the symptomatic question to detect outside issues, or reduce their effects, was inconsequential.

The super-dampening concept is still compelling, however. Given that an examinee's attentional resources are finite, and some portion is required for the examinee to produce physiological responding to test questions, it would seem reasonable that there should be some sort of relationship between the strength of polygraph scores and the presence or absence of outside issue concerns. It would also seem plausible that symptomatic questions might identify those examinees who harbor those outside concerns. One would expect that, as response intensity increases to the symptomatic questions, differential responsivity of the diagnostic questions should decrease. Below is the expected relationship in graphic form (See figure 1).

Figure 1.



It may be their intuitive appeal that contributed to the acceptance of the super-dampening concept and symptomatic questions long before they were subjected to independent verification. Since the symptomatic question was created to provide a physiological gauge for how distracting an outside issue was for the examinee, it became our interest to assess the association between the physiological responses to symptomatic questions and polygraph scores. If field data uncovered a relationship, it could provide substantiation for the Backster super-dampening concept, and support the continued or even expanded use of symptomatic questions. The Capps et al (1993) study was designed to determine whether the mere presence of symptomatic questions affected inconclusives, and though we do not know how to interpret their conclusions for the reasons stated earlier, our immediate question pertaining to the relationship between inconclusives and responses to symptomatic questions could not be answered by their methodology. The Honts et al (2000) data suggest that symptomatic questions did nothing meaningful in that regard. However, it is not known whether their findings would generalize beyond the laboratory. We therefore set out to scrutinize field data for evidence that responses to symptomatic questions correlated with reduced differential responding to the relevant and comparison questions.

Methodology

Cases

One hundred field cases were randomly drawn from the DoDPI confirmed case database. Cases in this database were all confirmed by confession, medical tests, or other irrefutable evidence. Because evidence separate from the polygraph decision was the criterion for inclusion in the DoDPI database, there were cases in which the original polygraph decisions were inconclusive or in error. The criteria for selection for the present study were that the cases had to be single-issue DoDPI ZCT examinations in which three relevant and three probable-lie exclusionary comparison questions were used. Symptomatic questions were placed in positions 3 and 8. All were field cases conducted by federal, state or local polygraph examiners. Half were

from confirmed deceptive cases, and the other half were confirmed nondeceptive cases. No demographic variables were considered in the selection process.

Human Scoring

An experienced polygraph examiner scored the 100 cases using the 7-position scoring system (DoDPI, 2001). He also recorded his subjective assessments of the presence and absence of responses to the outside issue questions, by channel and question. The examiner was kept blind to ground truth, base rates, and case facts.

Data Analysis

It was possible to test the effect of the super-dampening concept by looking at tabulations of responses to the symptomatic question and the absolute value of scores to the cases, since the concept predicts an inverse relationship, as seen in Figure 1. A Pearson's r was conducted for the absolute value of the polygraph scores by the number of reactions noted by the scorer to the symptomatic questions. Statistical significance was set at .05.

Automated Scoring

As a means of independent analysis, these same 100 cases were analyzed using a slight variant of the Rank Order Scoring System (ROSS) (Honts & Driscoll, 1987; Krapohl, Dutton & Ryan, 2001). The variation from standard procedure was that the symptomatic questions were also ranked, which is not normally done with the ROSS. The Kircher features were used in the rank assignment: respiration line length, EDA amplitude, and blood volume amplitude. The Kircher features were measured and recorded by a software package (Extract, ver 3.0). Ranking of the Kircher features was also automated, using the functions found in Microsoft Excel.

The physiological features for the symptomatic, relevant, and comparison questions were ranked in order of magnitude, within channel and within chart. For the EDA and blood volume, the largest responses received the highest ranks. For respiration line length (RLL), the shortest line lengths received the highest ranks, because RLL is inversely related to physiological arousal

(Timm, 1982). Ranks were summed separately for relevant, comparison, and symptomatic questions.

Data Analysis

The sum of the ranks of relevant questions was subtracted from the sum of the ranks of the comparison questions, creating a measure of differential arousal. The super-dampening concept holds that larger responses to symptomatic questions should predict dampening of differential responses. A Pearson's r was applied to the absolute value of the difference scores and the summed ranks

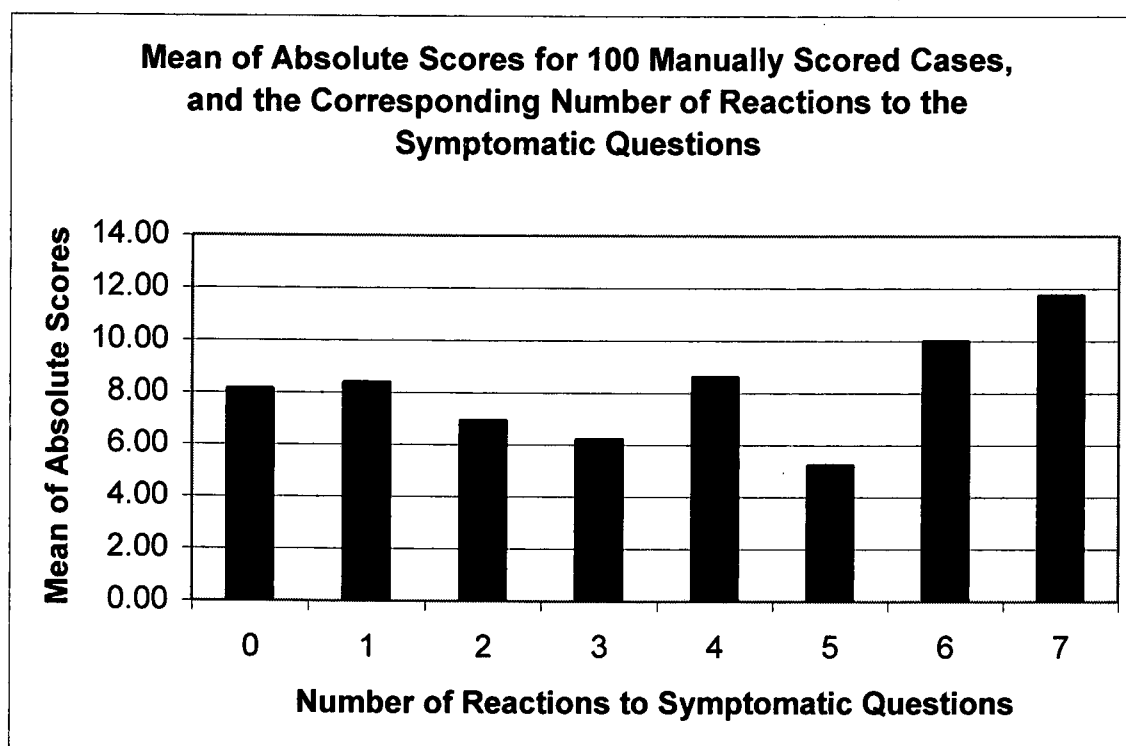
of the symptomatic questions. Statistical significance was set at .05.

Results

Human Scorer

The correlation for the number of reactions to symptomatic questions and the total of absolute scores for all cases was weak ($r=0.07$), and not significantly greater than chance. Figure 2 shows the absolute values of scores across the number of reactions to symptomatic questions.

Figure 2.



Automated Scoring

The correlation between the absolute difference scores of the ROSS and the average ranks of the symptomatic questions was also quite weak ($r=.06$), which was not significantly greater than chance. Similar analyses were conducted for the nondeceptive and deceptive cases separately. Nondeceptive cases showed a poor and non-significant relationship between the two variables ($r=.04$). Deceptive

cases also showed a small correlation ($r=-.18$), but that coefficient also failed to achieve statistical significance.

Discussion

The human scorer and the automated analyses of the present data converge on the finding that the symptomatic question did not predict suppressed polygraph scores in field

cases. The present data also corroborate the laboratory findings of Honts, Amato and Gordon (2000), who found no value in the symptomatic question. Notwithstanding Capps' (1993) conflicted data, there exists no published evidence that the symptomatic question has demonstrated empirically it can do as advertised. Against this backdrop, it is our view that the existing evidence reopens the debate as to whether chart time is being well invested with these questions.

There is a strong sentiment among advocates of symptomatic questions that they provide some value. Many field examiners can point to individual cases where symptomatic questions did uncover an outside issue, and once the issue was resolved, the case concluded with a valid decision. Of course, one can never know from those isolated cases whether the symptomatic questions only identified the presence of an outside issue, or whether, in a circular manner, they actually caused the concern that gave rise to the reactions to the symptomatic questions in the first place. Some experienced and competent field examiners have complained that the inclusion of symptomatic questions raises suspicion in many examinees rather than the alleviating it. This is especially true in non-US cultures, where some examiners have abandoned symptomatic questions because of the complications they introduce into an otherwise straightforward examination process. And when one considers what they are designed to do, it might be argued that the symptomatic question changes an explicitly single-issue (one crime) polygraph examination into an implicitly multiple-issue (this crime plus other crimes) polygraph examination.

To provide balance, let us also state that no published data indicate that symptomatic questions significantly interfere with polygraph examinations. We note that there may be isolated cases where examiners perceived some benefit from using symptomatic questions, just as others perceive

they create problems. Reasonable people can disagree on this issue, but we must acknowledge that anecdotes or selective personal recollections will not provide the best answer, nor will reference to authority: Only data can move us forward. The preponderance of the independent evidence now suggests that symptomatic questions probably do nothing reliably, neither good nor bad, across large numbers of examinations. As such, there are implications for the polygraph community. A collective decision is in order as to whether there is greater value in shortening the test question sequence, continuing with the traditional practice, using the symptomatic question on a case-by-case basis, or developing and validating a question that can demonstrate value. One type of question with at least face validity in the current environment is a countermeasure question. We would suggest, from the trends in the published literature, that countermeasure attempts are more problematic to polygraph validity than are outside issues, and they may be detected or deterred with a direct question. Considerable validation work is needed for whatever question might supercede the symptomatic question before we would be prepared to make a recommendation.

The conclusions of this paper will no doubt challenge long-accepted, if not empirically established doctrine for polygraph examiners trained to use symptomatic questions. Because of its controversial nature, the present authors provided an advance copy of this paper to Cleve Backster, the innovator of many polygraph concepts including those investigated here, and invited him the opportunity to comment. His response is published contiguous to this article. In addition, we will take the additional step of making our raw data available to anyone wishing to conduct a reanalysis. It is our goal to use the best scientific information to help the discipline develop a "best practices model" for polygraphy.

References

- Arther, R.O. (1977). Non-reactors. *The Journal of Polygraph Science*, 12(2).
- Backster, C. (1962). Methods of strengthening our polygraph technique. *Police*, 6(5), 61-68.
- Capps, M.H., Knill, B.L., & Evans, R.K. (1993). Effectiveness of the symptomatic questions. *Polygraph*, 22(4), 285-298.
- Department of Defense Polygraph Institute (2001). Test Data Analysis (Manual Scoring Process) 7-Position Numerical Analysis Scale. Ft. Jackson, South Carolina.
- Honts, C.R., Amato, S., & Gordon, A. (2000). Validity of the Outside-Issue Questions in the Control Question Test. Final Report on Grant No. N00014-98-1-0725. Boise State University.
- Honts, C.R., & Driscoll, L.N. (1987). An evaluation of the reliability and validity of rank order and standard numerical scoring of polygraph charts. *Polygraph*, 16(4), 241-257.
- Krapohl, D.J., Dutton, D.W., & Ryan, A.H. (2001). The Rank Order Scoring System: Replication and extension with field data. *Polygraph*, 30(3).
- Matte, J.A. (1996). *Forensic Psychophysiology Using the Polygraph: Scientific Truth Verification - Lie Detection*. Williamsville, NY: J.A.M. Publications.
- Matte, J.A. (2000). *Examination and Cross-Examination of Experts in Forensic Psychophysiology Using the Polygraph*. Williamsville, NY: J.A.M. Publications.
- McInerney, C.A. (1961). Emotional lethargy. Paper presented at the Eighth Annual Meeting of the American Academy of Polygraph Examiners, Washington, D.C.
- Reid J. E., & Inbau, F.E. (1966). *Truth and Deception: The Polygraph ("Lie Detector") Technique*. Williams & Wilkins: Baltimore.
- Timm, H.W. (1982). Analyzing deception from respiration patterns. *Journal of Police Science and Administration*, 10(1). 47-51