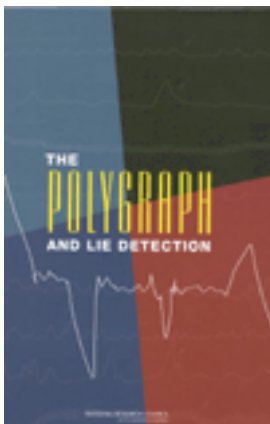


## Free Executive Summary

### The Polygraph and Lie Detection



Committee to Review the Scientific Evidence on the Polygraph, National Research Council

ISBN: 978-0-309-08436-9, 416 pages, 6 x 9, hardback (2003)

This free executive summary is provided by the National Academies as part of our mission to educate the world on issues of science, engineering, and health. If you are interested in reading the full book, please visit us online at <http://www.nap.edu/catalog/10420.html>. You may browse and search the full, authoritative version for free; you may also purchase a print or electronic version of the book. If you have questions or just want more information about the books published by the National Academies Press, please contact our customer service department toll-free at 888-624-8373.

*The polygraph, often portrayed as a magic mind-reading machine, is still controversial among experts, who continue heated debates about its validity as a lie-detecting device. As the nation takes a fresh look at ways to enhance its security, can the polygraph be considered a useful tool? The Polygraph and Lie Detection puts the polygraph itself to the test, reviewing and analyzing data about its use in criminal investigation, employment screening, and counter-intelligence. The book looks at: The theory of how the polygraph works and evidence about how deceptiveness—and other psychological conditions—affect the physiological responses that the polygraph measures. Empirical evidence on the performance of the polygraph and the success of subjects' countermeasures. The actual use of the polygraph in the arena of national security, including its role in deterring threats to security. The book addresses the difficulties of measuring polygraph accuracy, the usefulness of the technique for aiding interrogation and for deterrence, and includes potential alternatives—such as voice-stress analysis and brain measurement techniques.*

**This executive summary plus thousands more available at [www.nap.edu](http://www.nap.edu).**

Copyright © National Academy of Sciences. All rights reserved. Unless otherwise indicated, all materials in this PDF file are copyrighted by the National Academy of Sciences. Distribution or copying is strictly prohibited without permission of the National Academies Press <http://www.nap.edu/permissions/>. Permission is granted for this material to be posted on a secure password-protected Web site. The content may not be posted on a public Web site.

## Executive Summary

**F**or as long as human beings have deceived one another, people have tried to develop techniques for detecting deception and finding truth. Lie detection took on aspects of modern science with the development in the 20th century of techniques intended for the psychophysiological detection of deception, most prominently, polygraph testing. The polygraph instrument measures several physiological processes (e.g., heart rate) and changes in those processes. From the charts of those measures in response to questions on a polygraph test, sometimes aided by observations during the polygraph examination, examiners infer a psychological state, namely, whether a person is telling the truth or lying.

Polygraph testing is used for three main purposes: event-specific investigations (e.g., after a crime); employee screening, and preemployment screening. The different uses involve the search for different kinds of information and have different implications. A question asked about a specific incident (e.g., “Did you see the victim on Monday” or “Did you take the file home yesterday?”) often has little ambiguity, so it is clear what facts provide the criterion for a truthful answer.

For employee screening, there is no specific event being investigated, and the questions must be generic (e.g., “Did you ever reveal classified information to an unauthorized person?”). Both examinee and examiner may have difficulty knowing whether an answer to such a question is truthful unless there are clear and consistent criteria that specify what activities justify a “yes” answer. Examinees may believe they are lying when providing factually truthful responses, or vice versa. Polygraph

tests might elicit admissions to acts not central to the intent of the question and these answers might be judged either as successes or failures of the test. In this regard, we have seen no indication of a clear and stable agreement on criteria for judging answers to security screening polygraph questions in any agency using them.

The use of polygraph testing for preemployment screening is even more complicated because it involves inferences about future behavior on the basis of information about past behaviors that may be quite different (e.g., does past use of illegal drugs, or lying about such use on a polygraph test, predict future spying?).

The committee's charge was specifically "to conduct a scientific review of the research on polygraph examinations that pertains to their validity and reliability, in particular for personnel security screening," that is, for the second and third purposes. We have focused mainly on validity because a test that is reliable (i.e., produces consistent outcomes) has little use unless it is also valid (i.e., measures what it is supposed to measure). Virtually all the available scientific evidence on polygraph test validity comes from studies of specific-event investigations, so the committee had to rely heavily on that evidence, in addition to the few available studies that are relevant for screening. The general quality of the evidence for judging polygraph validity is relatively low: the substantial majority of the studies most relevant for this purpose were below the quality level typically needed for funding by the National Science Foundation or the National Institutes of Health.

## SCIENTIFIC EVIDENCE

### Basic Science

Almost a century of research in scientific psychology and physiology provides little basis for the expectation that a polygraph test could have extremely high accuracy. Although psychological states often associated with deception (e.g., fear of being judged deceptive) do tend to affect the physiological responses that the polygraph measures, these same states can arise in the absence of deception. Moreover, many other psychological and physiological factors (e.g., anxiety about being tested) also affect those responses. Such phenomena make polygraph testing intrinsically susceptible to producing erroneous results. This inherent ambiguity of the physiological measures used in the polygraph suggests that further investments in improving polygraph technique and interpretation will bring only modest improvements in accuracy.

Polygraph research has not developed and tested theories of the underlying factors that produce the observed responses. Factors other than truthfulness that affect the physiological responses being measured can

vary substantially across settings in which polygraph tests are used. There is little knowledge about how much these factors influence the outcomes of polygraph tests in field settings. For example, there is evidence suggesting that truthful members of socially stigmatized groups and truthful examinees who are believed to be guilty or believed to have a high likelihood of being guilty may show emotional and physiological responses in polygraph test situations that mimic the responses that are expected of deceptive individuals. The lack of understanding of the processes that underlie polygraph responses makes it very difficult to generalize from the results obtained in specific research settings or with particular subject populations to other settings or populations, or from laboratory research studies to real-world applications.

### **Evidence on Polygraph Accuracy**

Scientific evidence relevant to the accuracy of polygraph tests for employee or preemployment screening is extremely limited. Only one field study, which is flawed, provides evidence directly relevant to accuracy for preemployment screening. A few additional laboratory studies are relevant to preemployment or employee screening, but they are more analogous to specific-incident investigations than to screening because the deceptive examinee is given a precise recent incident about which to lie.

Of the 57 studies the committee used to quantify the accuracy of polygraph testing, all involved specific incidents, typically mock crimes (four studies simulated screening in the sense that the incidents were followed by generic screening-type questions). The quality of the studies varies considerably, but falls far short of what is desirable. Laboratory studies suffer from lack of realism, and in the randomized controlled studies focused on specific incidents using mock crimes, the consequences associated with lying or being judged deceptive almost never mirror the seriousness of these actions in real-world settings in which the polygraph is used. Field studies have major problems with identifying the truth against which test results should be judged. In addition, they suffer from problems associated with heterogeneity and lack of control of extraneous factors and more generally, they have lower quality than could be achieved with careful study design. Moreover, most of the research, in both the laboratory and in the field, does not fully address key potential threats to validity. For these reasons, study results cannot be expected to generalize to practical contexts.

Estimates of accuracy from these 57 studies are almost certainly higher than actual polygraph accuracy of specific-incident testing in the field. Laboratory studies tend to overestimate accuracy because laboratory conditions involve much less variation in test implementation, in the charac-

teristics of examinees, and in the nature and context of investigations than arises in typical field applications. Observational studies of polygraph testing in the field are plagued by selection and measurement biases, such as the inclusion of tests carried out by examiners with knowledge of the evidence and of cases whose outcomes are affected by the examination. In addition, they frequently lack a clear and independent determination of truth. Due to these inherent biases, observational field studies are also highly likely to overestimate real-world polygraph accuracy.

**CONCLUSION: Notwithstanding the limitations of the quality of the empirical research and the limited ability to generalize to real-world settings, we conclude that in populations of examinees such as those represented in the polygraph research literature, untrained in countermeasures, specific-incident polygraph tests can discriminate lying from truth telling at rates well above chance, though well below perfection. Because the studies of acceptable quality all focus on specific incidents, generalization from them to uses for screening is not justified. Because actual screening applications involve considerably more ambiguity for the examinee and in determining truth than arises in specific-incident studies, polygraph accuracy for screening purposes is almost certainly lower than what can be achieved by specific-incident polygraph tests in the field.**

The accuracy levels in the four screening simulations in our sample, which include a validation study of the Test for Espionage and Sabotage (TES) used in the employee security screening program of the U.S. Department of Energy (DOE), are in the range reported for other specific-incident laboratory studies. The one field study of actual screening presents results consistent with the expectation that polygraph accuracy in true screening situations is lower.

### Countermeasures

Countermeasures pose a potentially serious threat to the performance of polygraph testing because all the physiological indicators measured by the polygraph can be altered by conscious efforts through cognitive or physical means. Certain countermeasures apparently can, under some laboratory conditions, enable a deceptive individual to appear nondeceptive and avoid detection by an examiner. It is unknown whether a deceptive individual can produce responses that mimic the physiological responses of a nondeceptive individual well enough to fool an examiner trained to look for behavioral and physiological signatures of countermeasures. The available research provides no information on whether innocent examinees can increase their chances of achieving nondeceptive

outcomes by using countermeasures. (It is possible that classified information exists on these topics; however, this committee was not provided access to such information and cannot verify its existence or relevance.)

**CONCLUSION: Basic science and polygraph research give reason for concern that polygraph test accuracy may be degraded by countermeasures, particularly when used by major security threats who have a strong incentive and sufficient resources to use them effectively. If these measures are effective, they could seriously undermine any value of polygraph security screening.**

POLYGRAPH USE FOR SECURITY SCREENING

The proportion of spies, terrorists, and other major national security threats among the employees subject to polygraph testing in the DOE laboratories and similar federal sites presumably is extremely low. Screening in populations with very low rates of the target transgressions (e.g., less than 1 in 1,000) requires diagnostics of extremely high accuracy, well beyond what can be expected from polygraph testing. Table S-1 illustrates the unpleasant tradeoffs facing policy makers who use a screen-

TABLE S-1 Expected Results of a Polygraph Test Procedure with an Accuracy Index of 0.90 in a Hypothetical Population of 10,000 Examinees That Includes 10 Spies

S-1A If detection threshold is set to detect the great majority (80 percent) of spies			
Test Result	Examinee's True Condition		
	Spy	Nonspy	Total
"Fail" test	8	1,598	1,606
"Pass" test	2	8,392	8,394
Total	10	9,990	10,000
S-1B If detection threshold is set to greatly reduce false positive results			
Test Result	Examinee's True Condition		
	Spy	Nonspy	Total
"Fail" test	2	39	41
"Pass" test	8	9,951	9,959
Total	10	9,990	10,000

ing technique in a hypothetical population of 10,000 government employees that includes 10 spies, even when an accuracy is assumed that is greater than can be expected of polygraph testing on the basis of available research. If the test were set sensitively enough to detect about 80 percent or more of deceivers, about 1,606 employees or more would be expected “fail” the test; further investigation would be needed to separate the 8 spies from the 1,598 loyal employees caught in the screen. If the test were set to reduce the numbers of false alarms (loyal employees who “fail” the test) to about 40 of 9,990, it would correctly classify over 99.5 percent of the examinees, but among the errors would be 8 of the 10 hypothetical spies, who could be expected to “pass” the test and so would be free to cause damage.

Available evidence indicates that polygraph testing as currently used has extremely serious limitations in such screening applications, if the intent is both to identify security risks and protect valued employees. Given its level of accuracy, achieving a high probability of identifying individuals who pose major security risks in a population with a very low proportion of such individuals would require setting the test to be so sensitive that hundreds, or even thousands, of innocent individuals would be implicated for every major security violator correctly identified. The only way to be certain to limit the frequency of “false positives” is to administer the test in a manner that would almost certainly severely limit the proportion of serious transgressors identified.

**CONCLUSION: Polygraph testing yields an unacceptable choice for DOE employee security screening between too many loyal employees falsely judged deceptive and too many major security threats left undetected. Its accuracy in distinguishing actual or potential security violators from innocent test takers is insufficient to justify reliance on its use in employee security screening in federal agencies.**

Polygraph screening may be useful for achieving such objectives as deterring security violations, increasing the frequency of admissions of such violations, deterring employment applications from potentially poor security risks, and increasing public confidence in national security organizations. On the basis of field reports and indirect scientific evidence, we believe that polygraph testing is likely to have some utility for such purposes. Such utility derives from beliefs about the procedure’s validity, which are distinct from actual validity or accuracy. Polygraph screening programs that yield only a small percentage of positive test results, such as those in use at DOE and some other federal agencies, might be useful for deterrence, eliciting admissions, and related purposes. However, in populations with very low base rates of the target transgressions they



should not be counted on for detection: they will not detect more than a small proportion of major security violators who do not admit their actions.

We have thought hard about how to advise government agencies on whether or how to use information from a diagnostic screening test that has these serious limitations. We note that in medicine, such imperfect diagnostics are often used for screening, though only occasionally in populations with very low base rates of the target condition. When this is done, either the test is far more accurate than polygraph testing appears to be, or there is a more accurate (though generally more invasive or expensive) follow-up test that can be used when the screening test gives a positive result. Such a follow-up test does not exist for the polygraph. The medical analogy and this difference between medical and security screening underline the wisdom in contexts like that of employee security screening in the DOE laboratories of using positive polygraph screening results—if polygraph screening is to be used at all—only as triggers for detailed follow-up investigation, not as a basis for personnel action. It also underlines the need to pay close attention to the implications of false negative test results, especially if tests are used that yield a low proportion of positive results.

A belief that polygraph testing is highly accurate probably enhances its utility for such objectives as deterrence. However, overconfidence in the polygraph—a belief in its accuracy that goes beyond what is justified by the evidence—also presents a danger to national security objectives. Overconfidence in polygraph screening can create a false sense of security among policy makers, employees in sensitive positions, and the general public that may in turn lead to inappropriate relaxation of other methods of ensuring security, such as periodic security re-investigation and vigilance about potential security violations in facilities that use the polygraph for employee security screening. It can waste public resources by devoting to the polygraph funds and energy that would be better spent on alternative procedures. It can lead to unnecessary loss of competent or highly skilled individuals in security organizations because of suspicions cast on them by false positive polygraph exams or because of their fear of such prospects. And it can lead to credible claims that agencies that use polygraphs are infringing civil liberties for insufficient benefits to the national security. Thus, policy makers should consider each application of polygraph testing in the larger context of its various costs and benefits.

## ALTERNATIVES AND ENHANCEMENTS TO THE POLYGRAPH

**CONCLUSION:** Some potential alternatives to the polygraph show promise, but none has yet been shown to outperform the poly-



**graph. None shows any promise of supplanting the polygraph for screening purposes in the near term.**

The polygraph is only one of many possible techniques for identifying national security risks among federal employees. Other techniques attempt to detect deception from facial expressions, voice quality, and other aspects of demeanor; from measurements of brain activity and other physiological indicators; and from background investigations or questionnaires. Computerized analysis of polygraph records has the potential to improve the accuracy of test results by using more information from polygraph records than is used in traditional scoring methods. This potential has yet to be realized, however, either in research or in practice.

We considered the potential to increase the capability to identify security risks by combining polygraph information with information from other screening techniques, for example, in serial screening protocols such as are used in medical diagnosis. There are good theoretical reasons to think appropriate procedures of this sort would improve detection of deception, but we found no serious investigations of such multicomponent screening approaches.

### RESEARCH RECOMMENDATIONS

There has been no serious effort in the U.S. government to develop the scientific basis for the psychophysiological detection of deception by any technique, even though criticisms of the scientific grounding of polygraph testing have been raised prominently for decades. Given the heavy reliance of government on the polygraph, especially for screening for espionage and sabotage, the lack of a serious investment in such research is striking.

The limitations of the polygraph, especially for security screening, justify efforts to look more broadly for effective tools for deterring and detecting security violations. These might include modifications in the overall security strategies used in federal agencies, such as have been recommended by the Hamre Commission for DOE, as well as improved techniques for deterring and detecting security violations focused on individuals. Research offers one promising strategy for developing the needed tools.

**We recommend an expanded research effort directed at methods for detecting and deterring major security threats, including efforts to improve techniques for security screening.**

This effort should pursue two major objectives: (1) to provide federal agencies with methods of the highest possible scientific validity for pro-

tecting national security by deterring and detecting major security threats; and (2) to make these agencies fully aware of the strengths and limitations of the techniques they use. If the government continues to rely heavily on the polygraph in the national security arena, some of this research effort should be devoted to developing scientific knowledge that could put the polygraph on a firmer scientific foundation, develop alternative methods, or develop effective ways to combine techniques and methods. National security is best served by a broad research program on detecting and deterring security threats, not a narrow focus on polygraph research.

The research program should be open to supporting alternative ways of looking at the problems of deterrence and detection because there is no single research approach that clearly holds the most promise for meeting national security objectives. Thus, it might support research ranging from very basic work on fundamental psychological, physiological, social, and organizational processes related to deterring and detecting security threats to applied studies on implementing scientifically rooted methods in practical situations.

**A substantial portion of our recommended expanded research program should be administered by an organization or organizations with no operational responsibility for detecting deception and no institutional commitment to using or training practitioners of a particular technique. The research program should follow accepted standards for scientific research, use rules and procedures designed to eliminate biases that might influence the findings, and operate under normal rules of scientific freedom and openness to the extent possible while protecting national security.**

The mandate should be broad and should include both basic and applied research. The program should use standard scientific advisory and decision-making procedures, including external peer review of proposals, and should support research that is conducted and reviewed openly in the manner of other scientific research. Classified and restricted research should be limited only to matters of identifiable national security. Mission agencies might well continue to conduct implementation-focused research on detecting deception, but their work should be integrated with the broader research program proposed here.



# THE POLYGRAPH AND LIE DETECTION

Committee to Review the Scientific Evidence on the Polygraph

Board on Behavioral, Cognitive, and Sensory Sciences

and

Committee on National Statistics

Division of Behavioral and Social Sciences and Education

NATIONAL RESEARCH COUNCIL

*OF THE NATIONAL ACADEMIES*

THE NATIONAL ACADEMIES PRESS

Washington, D.C.

**[www.nap.edu](http://www.nap.edu)**

THE NATIONAL ACADEMIES PRESS 500 Fifth Street, N.W. Washington, DC 20001

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

This study was supported by Contract No. DE-AT01-01DP00344 between the National Academy of Sciences and the U.S. Department of Energy. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the organizations or agencies that provided support for the project.

#### Library of Congress Cataloging-in-Publication Data

The polygraph and lie detection.

p. cm.

Includes bibliographical references and index.

ISBN 0-309-08436-9 (hardcover)

1. Lie detectors and detection—Evaluation. I. Committee to Review the Scientific Evidence on the Polygraph (National Research Council (U.S.))

HV8078 .P64 2003

363.25'4—dc21

2002151541

Additional copies of this report are available from the National Academies Press, 500 Fifth Street, N.W., Lockbox 285, Washington, DC 20055; (800) 624-6242 or (202) 334-3313 (in the Washington metropolitan area); Internet, <http://www.nap.edu>

Printed in the United States of America.

Copyright 2003 by the National Academy of Sciences. All rights reserved.

Suggested citation: National Research Council (2003). *The Polygraph and Lie Detection*. Committee to Review the Scientific Evidence on the Polygraph. Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

## THE NATIONAL ACADEMIES

### *Advisers to the Nation on Science, Engineering, and Medicine*

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Bruce M. Alberts is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Wm. A. Wulf is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Bruce M. Alberts and Dr. Wm. A. Wulf are chair and vice chair, respectively, of the National Research Council.

**[www.national-academies.org](http://www.national-academies.org)**





## COMMITTEE TO REVIEW THE SCIENTIFIC EVIDENCE ON THE POLYGRAPH

STEPHEN E. FIENBERG (*Chair*), Department of Statistics, Carnegie Mellon University  
JAMES J. BLASCOVICH, Department of Psychology, University of California, Santa Barbara  
\*JOHN T. CACIOPPO, Department of Psychology, University of Chicago  
RICHARD J. DAVIDSON, Department of Psychology, University of Wisconsin, Madison  
PAUL EKMAN, Department of Psychology and Human Interaction Laboratory, University of California, San Francisco  
DAVID L. FAIGMAN, Hastings College of Law, University of California, San Francisco  
PATRICIA L. GRAMBSCH, Department of Biostatistics, University of Minnesota, Minneapolis  
PETER B. IMREY, Department of Biostatistics and Epidemiology, The Cleveland Clinic Foundation, and Departments of Statistics and Medical Information Sciences, University of Illinois at Urbana-Champaign.  
EMMETT B. KEELER, RAND Health, Santa Monica, California  
KATHRYN B. LASKEY, Systems Engineering and Operations Research Department, George Mason University, Fairfax, Virginia  
KEVIN R. MURPHY, Department of Psychology, Pennsylvania State University, University Park  
MARCUS E. RAICHLE, Department of Radiology and Neurology, Washington University, St. Louis  
RICHARD M. SHIFFRIN, Department of Psychology, Indiana University, Bloomington  
JOHN A. SWETS, BBN Technologies (emeritus), Tequesta, Florida  
  
PAUL C. STERN, *Study Director*  
ALEKSANDRA SLAVKOVIC, *Consultant*  
SUSAN R. McCUTCHEN, *Research Associate*  
DEBORAH M. JOHNSON, *Senior Project Assistant*

---

\* Served on the committee until May 28, 2002.

**BOARD ON BEHAVIORAL, COGNITIVE,  
AND SENSORY SCIENCES**

ANNE PETERSEN (*Chair*), W.K. Kellogg Foundation, Battle Creek,  
Michigan  
LINDA MARIE BURTON, Center for Human Development and Family  
Research, The Pennsylvania State University  
STEPHEN J. CECI, Department of Human Development, Cornell  
University  
EUGENE K. EMORY, Department of Psychology, Emory University  
ROCHEL GELMAN, Center for Cognitive Science, Rutgers University  
ANTHONY W. JACKSON, Disney Learning Initiative, Burbank,  
California  
PETER LENNIE, Center for Neural Science, New York University  
MARCIA C. LINN, Graduate School of Education, University of  
California at Berkeley  
ELISSA L. NEWPORT, Department of Brain and Cognitive Sciences,  
University of Rochester  
CHARLES R. PLOTT, Division of Humanities and Social Sciences,  
California Institute of Technology  
MICHAEL L. RUTTER, Institute of Psychiatry, University of London  
ARNOLD SAMEROFF, Center for Human Growth and Development,  
University of Michigan  
JAMES W. STIGLER, Department of Psychology, University of  
California at Los Angeles  
JOHN A. SWETS, BBN Technologies (emeritus), Tequesta, Florida  
RICHARD F. THOMPSON, Neurosciences Program, University of  
Southern California  
WILLIAM A. YOST, Office of Research and the Graduate School,  
Loyola University Chicago  
  
CHRISTINE R. HARTEL, *Director*

## COMMITTEE ON NATIONAL STATISTICS

JOHN E. ROLPH (*Chair*), Marshall School of Business, University of Southern California

JOSEPH G. ALTONJI, Thomas DeWitt Cuyler Professor of Economics, Yale University

ROBERT BELL, AT&T Labs-Research, Florham Park, New Jersey

LAWRENCE D. BROWN, Department of Statistics, University of Pennsylvania

ROBERT M. GROVES, Director, Survey Research Center, University of Michigan

HERMANN HABERMANN, United Nations Statistics Division

JOEL L. HOROWITZ, Department of Economics, Northwestern University

WILLIAM KALSBEEK, Survey Research Unit, Department of Biostatistics, University of North Carolina

ARLEEN LEIBOWITZ, School of Public Policy and Social Research, University of California at Los Angeles

THOMAS A. LOUIS, Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University

VIJAYAN NAIR, Department of Statistics, Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor

DARYL PREGIBON, AT&T Labs-Research, Florham Park, New Jersey

NORA CATE SCHAEFFER, University of Wisconsin-Madison

MATTHEW D. SHAPIRO, Department of Economics, University of Michigan

ANDREW A. WHITE, *Director*



# Contents

PREFACE	xiii
EXECUTIVE SUMMARY	1
1 LIE DETECTION AND THE POLYGRAPH	11
The Instrument, the Test, and the Examination, 12	
The Lie Detection Mystique, 18	
Scientific Issues, 21	
Context of Polygraph Tests, 25	
Structure of This Book, 26	
Notes, 27	
2 VALIDITY AND ITS MEASUREMENT	29
Reliability, Accuracy, and Validity, 29	
A Consistent Approach to Measuring Accuracy , 37	
Validity and Utility, 51	
Criterion Validity as Value Added, 58	
Conclusions, 60	
Notes, 62	
3 THE SCIENTIFIC BASIS FOR POLYGRAPH TESTING	65
The Scientific Approach, 65	
Theories of Polygraph Testing, 69	
Inferences from Polygraph Tests, 83	

The State of Polygraph Research, 91	
Conclusions, 100	
Notes, 103	
4 EVIDENCE FROM POLYGRAPH RESEARCH:	
QUALITATIVE ASSESSMENT	106
Overview, 107	
Laboratory Studies, 109	
Experimental Field Studies, 109	
Observational Field Studies, 112	
An Approach for Planned Field Research, 116	
Bias, Conflict of Interest, and Unscientific Decision Making, 117	
Conclusions, 120	
Notes, 120	
5 EVIDENCE FROM POLYGRAPH RESEARCH:	
QUANTITATIVE ASSESSMENT	121
Specific-Incident Polygraph Testing, 121	
Screening Studies, 130	
Special Populations and Conditions, 134	
Countermeasures, 139	
Conclusions, 148	
Notes, 152	
6 ALTERNATIVE TECHNIQUES AND TECHNOLOGIES	154
Autonomic Indicators, 155	
Measurements of Brain Function, 157	
Detection of Deception from Demeanor, 163	
Direct Investigation, 170	
Conclusions, 173	
Notes, 177	
7 USES OF POLYGRAPH TESTS	178
Tradeoffs in Interpretation, 179	
Using the Polygraph More Effectively, 194	
Legal Context, 201	
Conclusions, 208	
Notes, 211	
8 CONCLUSIONS AND RECOMMENDATIONS	212
Scientific Knowledge, 212	
DOE Polygraph Screening Policy, 218	
A Program of Research, 224	

CONTENTS	<i>xi</i>
REFERENCES	232
APPENDIXES	
A Polygraph Questioning Techniques	253
B Use of Polygraph Screening in the U.S. Department of Energy and Other Federal Agencies	259
C The Wen Ho Lee Case and the Polygraph	280
D Physiological Processes Measured by the Polygraph	286
E Historical Notes on the Modern Polygraph	291
F Computerized Scoring of Polygraph Data	298
G Process for Systematic Review of Polygraph Validation Studies	323
H Quantitative Assessment of Polygraph Test Accuracy	340
I False Positive Index Values for Polygraph Testing	354
J Decision Analysis of Polygraph Security Screening	358
K Combining Information Sources in Medical Diagnosis and Security Screening	364
L Biographical Sketches of Committee Members and Staff	375
INDEX	381





## Preface

**T**he polygraph, known more commonly as the “lie detector,” has a long and controversial history as a forensic tool, but it has also been used in a variety of other contexts, including employment screening. The U.S. federal government, through a variety of agencies, carries out thousands of polygraph tests each year on job applicants and current employees, and there are inevitable disputes that are sometimes highly publicized when someone “fails” a polygraph test. The American Polygraph Association, the largest polygraph association consisting of examiners in the private, law enforcement, and government fields, claims that the polygraph has a high degree of accuracy in detecting truthfulness or deception, with research studies published since 1980 reporting average accuracy rates ranging from 80 to 98 percent. Yet others claim that the studies underlying the polygraph represent “junk science” that has no scientific basis. Can experienced polygraph examiners detect deception? Again there is a diversity of claims. The World Wide Web contains a myriad of web pages advertising methods to beat the polygraph, while some people say that if the examinee knowingly lies, the polygraph will detect the lie.

The Committee to Review the Scientific Evidence on the Polygraph was asked by the U.S. Department of Energy to conduct a scientific review of the research on polygraph examinations that pertain to their validity and reliability, particularly for personnel security screening, and to provide suggestions for further research. Over 19 months, the committee held a series of meetings, visited polygraph facilities at several govern-

ment agencies, and examined large numbers of reports and published papers. We explored some historical dimensions of the research literature on the polygraph, including a link to work at the National Research Council (NRC) more than 80 years ago—and we learned how this led to the creation of the comic book character, Wonder Woman. We attempted to listen carefully to people representing both sides in the debate on polygraph accuracy, and we then stepped back and reviewed the evidence ourselves. The members of the committee brought to our deliberations diverse backgrounds and research perspectives, most of which had special relevance to one or more aspects of the research literature and practice of the polygraph. But we shared one thing in common: none of us had previously been engaged in polygraph research, *per se*, and each was intrigued by the claims in support of and against the polygraph.

Examining alternatives to the polygraph was also a key component of the committee's charge. We did this in a variety of ways, through input from agency representatives, visits to research laboratories, participation of committee members in outside workshops, presentations by researchers before the committee, and by reviewing relevant research literature shared with the committee by others or gathered by individual members and staff. We looked for polished alternatives and promising approaches and attempted to assess their scientific bases.

The committee tried to understand how the polygraph was used in different government agencies, for example, which format of polygraph test, what questions, with what instructions, etc. Andrew Ryan of the U.S. Department of Defense Polygraph Institute (DoDPI) served as a liaison to the committee from the government polygraph agencies, and was especially helpful in providing us with documentation and copies of research papers and manuscripts. David Renzelman and Anne Reed, Allen Brisentine, Paul Cully, and Alvina Jones arranged for visits with those in the polygraph programs at the Department of Energy, the National Security Agency, the Federal Bureau of Investigation, and the Central Intelligence Agency, respectively.

We also appreciate the information we received from many people who made presentations before the committee: Gary Berntson (Ohio State University), Senator Jeff Bingaman (D-New Mexico), Emanuel Donchin (University of Illinois), Lawrence Farwell (Brain Fingerprinting Laboratories, Inc.), General John A. Gordon (National Nuclear Security Administration), John Harris (Johns Hopkins University Applied Physics Laboratory), Charles Honts (Boise State University), William Iacono (University of Minnesota), Stephen Kosslyn (Harvard University), Peter Lyons (Office of Senator Peter Domenici), Joseph Mahaley (Department of Energy), George Maschke ([antipolygraph.org](http://antipolygraph.org)), Anne Reed (Department of Energy), Sheila Reed (North Texas State Hospital), David Renzelman (Department of Energy), Drew Richardson (Federal Bureau of Investigation, retired),

Andrew Ryan (Department of Defense Polygraph Institute), and Alan P. Zelicoff (Sandia National Laboratory).

The events of September 11, 2001, and their aftermath briefly interrupted the committee's activities, but more importantly, they reinforced for the committee the important roles that many of the agencies and organizations we had been visiting play in attempting to assure national security.

Late in the committee's deliberative process, one of the committee members, John Cacioppo, resigned from the committee to pursue research he had initiated as a consequence of his work on the committee. John was a major contributor to the committee's work, especially as it related to psychophysiology, and we owe him a great debt even though he was unable to assist us in the final revisions.

This report would not have been completed had it not been for the tremendous efforts of a number of key staff. Paul Stern served as study director and guided us from the outset, helping us to organize our work and to write the report. His insightful observations often forced us to rethink draft conclusions and summaries, and his good humor and gentle prodding made our writing tasks easier to accept. In summary, Paul was a partner in almost all of our tasks. Christine Hartel, director of the Board on Behavioral, Cognitive, and Sensory Sciences (BBCSS) stepped in at a crucial stage of the committee's work and played a lead staff role when Paul was temporarily absent and has provided input and wise counsel throughout. Andrew White, director of the Committee on National Statistics, also participated in our meetings and offered assistance and support. Susan McCutchen worked on the full spectrum of the committee's activities, secured documents for us, organized our research database, and interacted with representatives of the government polygraph agencies. Anne Mavor and James McGee, study directors, and Jerry Kidd, senior program officer on the BBCSS staff, assisted in the initial screening of articles for the committee's literature review. Deborah Johnson provided valuable project assistance, particularly in making arrangements for the committee's meetings and visits to agencies. Barbara Torrey, then executive director of the NRC's Division of Behavioral and Social Sciences and Education (DBASSE), and associate director Miron Straf, who developed the project initially, took a continuing interest in the work of the committee. We have also been fortunate to have the continuing wise counsel of Eugenia Grohman, director, DBASSE Reports Office, throughout the work of the committee. We are also grateful for help received from Nancy A. Obuchowski, at the Cleveland Clinic Foundation, and Kevin S. Berbaum, at the University of Iowa, in acquainting us with existing software for receiver operating characteristic (ROC) analysis. Finally, we thank Aleksandra Slavkovic who provided technical statistical assis-

tance and support to the committee, especially in connection with the empirical analyses reported in Chapter 5 and Appendix F.

This report has been reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise, in accordance with procedures approved by the NRC's Report Review Committee. The purpose of this independent review is to provide candid and critical comments that will assist the institution in making the published report as sound as possible and to ensure that the report meets institutional standards for objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscript remain confidential to protect the integrity of the deliberative process.

We thank the following individuals for their participation in the review of this report: John F. Ahearne, Sigma Xi and Duke University; Gershon Ben-Shakhar, Department of Psychology, Hebrew University, Jerusalem; Roy D'Andrade, Department of Anthropology, University of California, San Diego; Paul Gianelli, School of Law, Case Western Reserve University; Bert F. Green, Jr., Professor of Psychology, Emeritus, Johns Hopkins University; James A. Hanley, Department of Epidemiology and Biostatistics, McGill University, Canada; Barbara C. Hansen, School of Medicine, University of Maryland, Baltimore; Ray Hyman, Department of Psychology, University of Oregon; Sallie Keller-McNulty, Statistical Sciences, Los Alamos National Laboratory; John Kircher, Department of Educational Psychology, University of Utah; James L. McGaugh, Center for the Neurobiology of Learning and Memory, University of California, Irvine; Gregory A. Miller, University of Illinois; William Revelle, Northwestern University; Anthony E. Siegman, McMurtry Professor of Engineering, Emeritus, Stanford University; Robert M. Stern, Pennsylvania State University; Stephen Stigler, Department of Statistics, University of Chicago; and James Woolsey, Shea & Gardner, Washington, DC.

Although the reviewers listed above have provided many constructive comments and suggestions, they were not asked to endorse the conclusions or recommendations nor did they see the final draft of the report before its release. The review of this report was overseen by John Bailar, University of Chicago (emeritus), and Michael Posner, Department of Psychology, University of Oregon. Appointed by the National Research Council, they were responsible for making certain that an independent examination of this report was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of this report rests entirely with the authoring committee and the institution.

Stephen E. Fienberg, *Chair*  
Committee to Review the Scientific  
Evidence on the Polygraph

# **THE POLYGRAPH AND LIE DETECTION**

