NIST TIP White Paper Submission

Critical National Need Idea Title: Evaluating Behavioral Correlates of Deception

Submitting organization: Montclair State University
Contact: Eileen Fitzpatrick
Linguistics Department
Montclair State University
Montclair NJ 07043
973.655.4286 (office)
973.655.7909 (fax)
fitzpatricke@mail.montclair.edu

Contributing organization: Linguistech Consortium
Contact: Joan Bachenko
PO Box 165
Port Murray NJ 07865
908.334.2086
jbachenko@gmail.com

Key words: Deception Detection, Credibility Assessment

Introduction

The ability to recognize deception is a fundamental goal for national security, law enforcement and any organization whose success depends on the credibility of spoken and written communication. Discovering lies can thwart serious threats as well as provide productive directions in the investigation of past events and more accurate predictions of likely future events.

Several approaches to deception detection have been proposed and, in some cases, developed and deployed. These include physiological measurements such as polygraph and computer voice stress analysis, imaging technologies such as functional MRI and thermal imaging, and behavioral indicators that can be found in verbal and non-verbal actions. However, all of these approaches suffer from shortcomings that prevent them from providing a scientifically reliable and widely applicable solution to the problem of deception detection. Three issues in particular hamper current research: (1) The reliance on laboratory experiments to test a method or technology. These experiments, where subjects are often college students, necessarily fail to replicate the high stakes quality of real world situations. (2) The lack of scientifically valid observational studies. Investigation of high stakes deception can only take place when ground truth is known, a requirement that has proven to be a strong barrier to progress in deception research. (3) The lack of a well-defined community of deception researchers. Work in this area is conducted by largely unconnected groups. There are no common benchmarks of success and no shared databases. Finally, while technologies such as polygraphy and computer voice stress analysis seem well established, they are based on the unproven assumption of a connection between lying and specific physiological activities.

A national program for deception detection research can provide cohesion and resources that will greatly advance the scientific contribution and applicability of the work in this area. This should be a high risk, high reward program whose goal is to increase our knowledge of deceptive behavior and to develop methods and technologies that will lead to measurable improvements in national security, law enforcement, criminal justice and commercial sector applications.

The focus of the program will be on evaluating proposed behavioral correlates of deception—verbal and non-verbal. The program should issue Requests For Proposals (1) to provide innovative methods of assembling a large database of real world truthful and deceptive behavior with robust ground truth verification, (2) to implement deception detection systems that work in both real time and after the fact, and (3) to provide cognitively realistic models of deception detection that can be used when implemented systems are unavailable or unsuitable. In addition, the program should provide sponsorship of competitions that evaluate the various techniques and suggest ways of combining techniques for more effective detection of deception.

The remainder of this paper addresses the major selection criteria of the Technology Innovation Program.

A. Maps to Administration Guidance

In 2003, the National Academy of Sciences National Research Council published *The Polygraph and Lie Detection*, a study that probes the theory, history and use of polygraphy in deception detection.[1] The study also examines other techniques of information validation and concluded that, while polygraphy fails to meet scientific standards of validity, a practical alternative to polygraph testing does not yet exist. The NAS authors propose a research program for detecting security threats that incorporates a significant effort in improving and developing methods to recognize deceptive communication.

In 2006, the National Intelligence Board, at the request of the Intelligence Community, issued a report on methods of educing information that strongly emphasized the need for reliable methods of deception detection. The goal of the report was to provide motivation and groundwork for research into eduction techniques that are scientifically justified and consistent with the preservation of human rights. In concluding its review of deception detection methods, the report asserts: “The U.S. government needs to implement an aggressive, focused strategic plan for supporting behavioral research and developing enhanced capabilities to validate information and sources. Such a plan should focus on understanding actual behavior and prioritize projects on the basis of operational needs, operational realities, cost, and potential return on investment” (p. 52).

Despite this prioritization, research and applied efforts in behavioral methods of deception detection are scattered across many agencies and organizations which test different techniques of assessing credibility with no unified means of evaluation.

Among government organizations, work on behavior and deception is carried out by the Department of Homeland Security (DHS), the Defense Academy of Credibility Assessment (DACA), the Center for the Advanced Study of Language (CASL), the Air Force Office of Scientific Research (AFOSR), the Bureau of Alcohol, Tobacco and Firearms (ATF) and the Federal Bureau of Investigation (FBI). Non-governmental organizations include a small number of universities that conduct experimental research in deception and several small companies that provide deception detection services primarily to law enforcement. The lack of a common forum, interest group or association makes communication amongst the groups sporadic or non-existent, even for those within the same sector (e.g. Government). Communication across organizational lines is nearly always conducted by motivated individuals with specific interests or needs.

Table 1 indicates the diversity of approaches within and across sectors. More detail is provided in the section on Evidence of Commitment.

Sector		**Data**	**Technique**
Government	DHS	experimental, empirical	non-verbal behavior, non-invasive sensors
	DACA	empirical	verbal behavior, polygraph
	CASL	experimental, empirical	verbal and non-verbal behavior
	AFOSR	experimental	verbal behavior
	ATF	empirical	verbal behavior: cognitive interviewing
	FBI	empirical	verbal behavior: statement analysis
Academic		experimental	verbal and non-verbal behavior
Commercial		empirical	verbal behavior, polygraph, voice stress

Table 1: Characterization of groups working on deception detection in the United States

Funding levels for research and development projects are modest, limiting the extent and influence of research results. The market for deception detection products is also small except for government purchases of polygraph technology and computer voice stress devices. Progress in this area thus has the potential to add significant depth to our understanding of human cognition, provide added protections for national security and open up new markets for commercial development.

B. Justification for Government Attention

1. *Magnitude and nature of the problem*

Many professionals in law enforcement, government, and intelligence are required to assess veracity on a daily basis. The ability to spot deception is an issue wherever something important is at stake in communication: in police, security, border crossing, customs, and asylum interviews; in congressional hearings; in financial reporting; in legal depositions; and in predatory communications, including internet scams, identity theft, and fraud.

The traditional tool to assess veracity is the polygraph. The scientific literature shows that, when used in the investigation of specific events such as crimes, the accuracy of the polygraph test is well above chance [1,2]. However, the National Academy of Sciences National Research Council's 2003 report on polygraphy [1] found that, for the other two widely used applications of polygraphy, security screening and pre-employment 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." The main problem with the use of polygraphy in screening, the report found, is that the generic nature of the questions (e.g., "Did you ever reveal classified information to an unauthorized person?") makes it hard to know whether an answer is truthful or not without "clear and consistent criteria that specify what activities justify a 'yes' answer." This difficulty is one facet of the general problem in assessing the validity of any deception detection approach: establishing the ground truth against which to compare a claim.

An additional problem with polygraphy is that it is not available in many circumstances that require a quick decision, such as whether to admit an individual to an airplane, grant a visa, or release a potential suspect.

The NAS report considers several alternative means of detecting deception that may supplement or substitute for the polygraph, focusing on emerging technologies that measure brain activity and those that rely on measures of externally observable behaviors. These technologies show promise for marked improvement of deception detection, but have not been sufficiently evaluated. The NAS report states that "Agencies that use such techniques should support independent scientific evaluation so that they can be fully informed when making decisions on whether and how to use the techniques and on how to use the test results they produce." Despite their promise, these techniques come from many unrelated disciplines and the area lacks a scientific paradigm that will allow for consistent evaluation of the technologies across the disciplines and will address the pervasive problem of establishing ground truth in high stakes situations.

2. *Deception Detection: Societal Challenge*

Deception detection is an emerging area of research and applications that is distributed across a relatively small number of researchers in the disciplines of law enforcement and criminal justice, national security, psychology, anthropology, and linguistics. The variety of disciplines involved presents the greatest societal challenge to the area.

The social scientists largely focus on laboratory experiments with subjects, usually students, acting out a scenario or, in the case of psychophysical experiments, stating as true claims they know to be false. Because the basic facts of the experiment are controlled, it is possible to observe subjects' behavior in a uniform test and so establish statistically the features that are most likely to be associated with deception. However, for ethical reasons, laboratory experiments lack high stakes pressures: the subjects have nothing to lose if they are caught lying. As Nancy Kanwisher, an fMRI researcher at MIT, is quoted as saying in the 7/2/07 New Yorker article “Duped:” “Making a false response when instructed to do so is not a lie,” [To know whether the technology works], “you’d have to test it on people whose guilt or innocence hasn’t yet been determined, who believe the scan will reveal their guilt or innocence, and whose guilt or innocence can be established by other means afterward.”

In other words, it would be necessary to run a legal version of a clinical trial, using real suspects instead of volunteers. Many in the law enforcement and security fields use language-based methods--grouped under the heading of 'statement analysis'--to discover deception in an interview or narrative [3,4]. In this case, possible liars have a great deal to lose if they are caught: reputation, money, freedom, job. However, ground truth is often an unknown unless the facts emerge over time through continued investigation, suspect cooperation, or luck. This lack of control over basic facts makes it nearly impossible to conduct a controlled experiment using "real world" data. As a result there are few scientific studies of statement analysis as it is used in real world applications.

In addition to the research in the social science and law enforcement and security fields, there is also a small group of researchers in computer science and engineering working to operationalize the findings from the other fields, but the lack of a common platform of evaluation and a community that enables comparison of results hampers their ability to take the research from the laboratory to technology development.

In short, the field of deception detection, though it includes many areas of promising research, lacks a unified paradigm for the evaluation of this research and a community with a common methodology and common research goals. It needs an approach that will meld real world data with the guarantee of ground truth given by experimentation. The initial goal should be a database of video and audio recordings along with their transcripts and a collection of written material. The recordings should include interviews, depositions, hearings and meetings; written material should comprise statements, reports and letters where the assessment of truth is critical. In every case, it is essential to support the assessment of truth using innovative techniques for maximizing ground truth support for the claims made in the data.

An example of input to the research would be statements spoken or written by individuals under the active questioning of law and/or security professionals, a clear definition of what constitutes ground truth information and how it is to be matched against the statements, and the guarantee of a given amount of appropriate ground truth data, where ‘appropriate’ may mean something different depending on whether the analysis is verbal or non-verbal, and the data is spoken or written. For example, in the security interview, background checks may be done before and during the initial polygraph interview, depending on the category of the interviewee and the level of security. For those individuals on whom background data is collected before the interview, there will be enough data a priori to determine whether this individual’s statements are worth recording.

Prior to the NAS report, polygraphy was used in screening venues with a high degree of confidence, on the assumption that the success of polygraphy in specific event venues would carry over to screening venues. However, the result of the report has been the erosion of confidence in polygraphy in these venues and the subsequent investment in piecemeal alternatives techniques. Success in the evaluation program we propose would result in a high measure of confidence in the techniques that weathered the evaluation and, therefore, a high degree of confidence that malfeasants would not be cleared to operate in secure environments or to commit further crimes.

3. Evidence of commitment

A by-product of the building of an evaluation system for testing techniques of deception detection would be a bringing together of the main researchers working on these techniques. There is currently no field of 'deception research'. It is an emerging area of research and applications that is distributed across a relatively small number of linguists, psychologists, anthropologists, law enforcement and national security professionals with little cross-disciplinary cooperation. The fact that there is no journal exclusively focused on deception is indicative of the piecemeal nature of the research. It is similar to the state of automatic speech recognition and understanding twenty years ago, when researchers from electrical engineering, computer science, natural language processing and artificial intelligence were working and publishing on separate tracks before the DARPA/NIST evaluations brought the research together.

The likely proposers to a competition in this area would include:

- Government organizations specifically engaged in research on deception, including the Department of Homeland Security's Hostile Intent Detection program, the Center for the Advanced Study of Language's program in Detecting Deception across Cultures, the Air Force Office of Scientific Research's Detecting Deception program, the Army Research Institute, the Naval Research Laboratory, and the National Institute of Justice's Wrongful Conviction program.

- Academic institutions engaged in deception detection research, including, but not limited to Columbia University Computer Science, Montclair State University Linguistics Department, SUNY-Buffalo's Center for Unified Biometrics and Sensors, the University of Arizona's Center for the Management of Information, the UC San Francisco Psychology Department, the University of Rochester, the University of San Francisco Psychology Department, and the University of Virginia

- Companies engaged in applications development in deception detection, including but not limited to: Alias Technologies, Deception Discovery Technologies, the Diogenes Company, the Draper Institute, the Laboratory for Scientific Investigation, MITRE's Counter-deception decision support program, and Realear

Most of the deception researchers in these organizations whom we have met through our contacts at CASL and IARPA are aware of the lack of a unified system of evaluation for the field but, without an overriding initiative to support the development of such a system, see no solution to the problem of controlling access to ground truth information in real world environments.

C. Essentials for Technology Innovation Program Funding

1. Stimulates the Nation’s scientific frontiers

The initiative we envision would generate a computational model of deceptive behavior which would enable analytical techniques to evaluate the ability of a given deception detection technique to predict the probability of deception. The development of such an initiative has the potential of expanding social science research out from the purely experimental approach that has been the hallmark of the social sciences to the newer computational science, identified by the President’s Information Technology Advisory Committee as the “third pillar of 21st century science” [5]. The PITAC devotes a substantial portion of its report to the application of computational modeling in the social sciences, and considers some of the reasons for the reluctance of the social sciences to embrace this new paradigm, including the entrenched nature of the experimental paradigm and the lack of examples comparable to the genome decoding effort in the biological sciences. While our proposal is neither as fundamental nor as large-scale as the Human Genome Project, it is a compelling example of a case where the compilation of real world data can succeed where experimentation is failing.

A shared testbed of verbal and non-verbal data with associated ground truth verification can leverage the smaller successes such as the measurement of event related potentials [6], facial and body movements [7], or aspects of language use that are consistently associated with deception [8] by providing a means of evaluating the successes against real world data and suggesting ways of combining the approaches for greater success. The testbed can also, of course, weed out approaches that have no substantial promise of success.

2. Meets a timely need not met by others.

The threats to the Nation that can be nullified through the outing of deception, from terrorist plots to financial fraud, are only increasing and, as the NAS report emphasizes, we have no empirically validated means of identifying this deception.

With respect to funding, the programs that are targeted at basic academic research are small, based on the experimental design paradigm, and oriented towards discovering the behavioral attributes of liars rather than towards the technologies that might capitalize on these behaviors. The academic research at CASL, for example, is funded at the $1million per annum level. The Air Force Office of Scientific Research has funded their Detecting Deception research at $4 million for a 5-year period.

No other organization besides NIST is in a position to act as a repository of data and an evaluation site for the various, often competing, techniques that are being put forward in the deception detection field. As a consequence, there is, to our knowledge, no initiative of the type we describe here. As confirmed by the NAS report, most of the funding sources are putting money into traditional experimental approaches. While the research on these approaches acknowledges the problem of the lack of high-stakes ground truth [9], the field, lacking an orientation towards the computational approach, seems unaware of how to overcome the problem.

3. Delivers the potential for impacts and transformations.

The ability to extract useful information from deceptive interviews would vastly improve the Nation's ability to uphold the rule of law and maintain its security while avoiding the Constitutional problems and questionable and often ineffective techniques of harsh interrogation. The approaches discussed here, if they were to be evaluated and found effective, would enable law enforcement and security personnel to extract such information by non-invasive means, with the potential of transforming the Nation's crime fighting and intelligence abilities in the way that fingerprinting and polygraphy, when applied to specific crimes, have done in the 20th century.

The research initiative proposed here offers a way to test a wide variety of approaches for extracting truthful information from deceptive individuals, and to operationalize, and potentially combine the approaches that are empirically shown to be most effective.
The initiative would also have the effect of bringing social science, law enforcement, and intelligence personnel together in a continuing venue that would allow cross-pollination of approaches. The creation of a community of deception detection researchers would also foster adoption of the computational science paradigm among the social science researchers as they are required to use the evaluative database to test their techniques.

This last benefit holds perhaps the most potential for transformation since it provides a model to the social sciences of how to use the computational paradigm in research on behavior. As the PITAC report states "despite the great opportunities and needs, universities and the Federal government have not effectively recognized the strategic significance of computational science in either their organizational structures [note the initiative's formation of a deception detection research community) or their research and education planning [note the initiative's computational modeling approach to research evaluation). These inadequacies compromise U.S. scientific leadership, economic competitiveness, and national security."

We believe the initiative proposed here addresses a critical national need for law enforcement and intelligence tools to address rising fraud and security threats to our economy and basic safety. Proposals to contribute to a database of truthful and deceptive behavior with ground truth verification will enable the evaluation of a variety of promising, non-invasive techniques for identifying deception and enable the development of technologies that operationalize these techniques, while bringing together the diverse research interests in deception in the use of a common, innovative scientific paradigm.

References

1. Board on Behavioral, Cognitive, and Sensory Sciences and Education Committee on National Statistics. 2003. *The Polygraph and Lie Detection*. The National Academies Press. Available at http://books.nap.edu/catalog.php?record_id=10420

2. Honts, Charles R. 2004. The psychophysiological detection of deception. In P. A. Granhag and L. A. Strömwall. *The Detection of Deception in Forensic Contexts*. Cambridge University Press.

3.Adams, S. 2002. *Communication under stress: indicators of veracity and deception in written narratives*. Ph.D. dissertation, Virginia Polytechnic Institute and State University

4. Smith, N. 2001. Reading between the lines: An evaluation of the scientific content analysis technique (SCAN). *Police Research Series*. London,UK. Available at www.homeoffice.gov.uk/rds/prgpdfs/prs135.pdf

5. President's Information Technology Advisory Committee. 2005. *Computational Science: Ensuring America's Competitiveness*. Available at http://www.nitrd.gov/pitac/reports/20050609_computational/computational.pdf

6. Farwell and Donchin. Farwell, L.A., and E. Donchin 1991 The truth will out: Interrogative polygraphy ("lie detection") with event-related potentials. *Psychophysiology* 28:531-547.

7. Ekman, P. and W. V. Friesen. 1974. Detecting deception from the body or face. *Journal of Personality and Social Psychology*, 20, 288-298.

8. DePaulo, B. M., J.J. Lindsay, B.E. Malone, L. Muhlenbruck, K. Charlton, and H. Cooper. 2003. Cues to deception. *Psychological Bulletin*, 129(1), 74-118.

9. Frank, M. G. 2008. Research methods in detecting deception research. In Harrigan, J., R. Rosenthal, and K. Scherer. *The New Handbook of Methods in Nonverbal Behavior*. Oxford University Press.