International Bulletin of Political Psychology

Volume 16 | Issue 14 Article 1

5-26-2010

Spot Off: The GAO Takes On the TSA's Behavior Detection Program

Editor

International Bulletin of Political Psychology

Title: Spot Off: The GAO Takes On the TSA's Behavior Detection Program
Author: Editor
Volume: 16
Issue: 14
Date: 2010-05-26
Keywords: Airport, Aviation, Behavior Detection, Profiling, Security, SPOT, Terrorism

Abstract: The United States Government Accountability Office (GAO) has recently Issued Efforts to Validate TSA's Passenger Screening Behavior Detection Program Underway, but Opportunities Exist to Strengthen Validation and Address Operational Problems (May 2010, GAO-10-763). This IBPP article will describe and comment on the main GAO findings and additional data on which the findings are based. The article will end with some basic challenges to behavior detection as a useful security measure.

Main GAO Findings (from Highlights of GAO-10-763 and all findings pertain to the Screening of Passengers by Observation Techniques (SPOT) deployed by the Transportation Security Administration (TSA):

(1) "TSA deployed SPOT nationwide without first validating the scientific basis for identifying suspicious passengers in an airport environment." (Comment: Just one challenge to SPOT scientific validation is TSA's own comprehensive security plan, which describes 20 layers of security. As described in Figure 1 of the report, these layers vary from "intelligence," "crew vetting," "hardened cockpit door," to "Federal Flight Deck Officers," "canines," and more. How moment-to-moment variations within and among all of these interact with some true or approximate validity of behavior detection may be unknowable).

(2) "A scientific consensus does not exist on whether behavior detection principles can be reliably used for counterterrorism purposes." (Comment: If the GAO is correctly using the term reliably, this means that there's no consensus on whether behavior detection (BD) will even provide an identical answers based on identical information in an identical situation. Without this, attaining validity—demonstrating that the answer provided is, indeed, correct—becomes even more difficult).

(3) "No other large-scale security screening program based on behavioral indicators has ever been rigorously scientifically validated." (Comment: Two of the more common attempts involving mass screening are to prevent or identify (a) theft of merchandise by employees at commercial venues and (b) espionage by government employees. These two attempts have at least one problem in common. The base rate of theft and espionage may be so low that the information used to identify thieves and spies will result in a much, much larger number of people being falsely than truly accused. In addition, the degree of predictive accuracy required by information to avoid this consequence is very unrealistically high. Thus, we are left with a more statistically effective and cost-effective alternative of assuming all individuals do not pose a threat.

(4) "DHS's current plan to assess SPOT is not designed to fully validate …behavior detection…For example, factors such as the length of time BDOs [behavior detection officers] can observe passengers without becoming fatigued are not part of the plan…". (Comment: Regardless of even a perfect linkage between some information and a threat from a passenger, fatigue may decrease the probability that information is even noticed by a BDO. Moreover, if that information is noticed, fatigue may decrease the probability that the BDO interprets it correctly and, if interpreted correctly, the information is acted on in the correct manner).

(5) "TSA…is not fully utilizing the resources it has available to systematically collect and analyze the information obtained by BDOs on passengers who may pose a threat…". (Comment: An important key to any screening program is a continuous evaluation process wherein findings are used to modify that program. Even if the screening program has attained suitable validity when it is initiated, there can be validity drift as the nature of the (a) threat (who and what poses the intentional security concern), (b) vulnerability (all the things that can go wrong and how great the impact of them going wrong will be), (c) risk (continuous interactions of threat and vulnerability), (d) situation (economic, social, cultural, political, etc.) in which all of this is occurring, and (e) information used to identify so-called suspicious passengers changes as to type and meaning. This is why an additional finding is so crucial—that TSA "…lacks outcome-oriented measures to evaluate the program's progress toward reaching its goals".

(6) "TSA's Transportation System Operations Center [TSOC]…generally does not check all law enforcement and intelligence data bases available to it to identify persons referred by BDOs". (Comment: Even if TSA was fully utilizing collection and analysis capabilities mentioned in (5) above, that analysis would be more problematic than need be given that each data base not being currently used could add somewhat to overall behavioral detection validity).

(7) "…most BDOs lack a mechanism to input data on suspicious passengers into a data base used by TSA analysts and…to obtain information from the TSOC on a timely basis". (Comment: These findings relate to (5) and (6) above and illustrate the interdependencies among collection and analysis capabilities and shortfalls. These findings also would seem to detract from the motivation of BDOs to do their job in a mindful and goal-oriented fashion as opposed to an automatic, procedure-oriented, mindless fashion.

Additional GAO Data on which Findings Are Based. (1) Behavioral detection programs often receive political support, because such programs have been developed, implemented, and evaluated by Israeli security authorities who presumably have an even greater challenge. However, as described in Footnote 10 of the GAO report, the Israeli programs (in contrast to SPOT), do not use a list of specific behaviors—viz., facial expression, body language, and appearance—with numerical values for each, do not use thresholds for deciding whether to interview a passenger, and do conduct frequent covert tests of security personnel's attentiveness. (2) According to Footnote 28, the National Research Council has found that "…scientific support for linkages between behavioral and physiological markers and mental states is strongest for elementary states, such as simple emotions; weak for more complex states, such as deception; and nonexistent for highly complex states, such as when individuals hold terrorist intent and beliefs." (3) According to Footnote 35, even in the best of circumstances 9 nonverbal indicators might be a limit for highly trained personnel to detect. (4) One interpretation of the following quantitative data might suggest that increasing resources are being allocated for SPOT without acceptable benefits. According to Figure 3, SPOT expenditures have increased from $47 million in FY 2007 to $198 million in FY 2009 with appropriations of $212 million for 2010. According to Figure 4 and concerning the interval from June 2004 through September 2008, 2 billion passengers were processed through airports using SPOT, 152,000 were identified for secondary referrals, 14,000 for law enforcement referral, and 1100 were arrested. Of these arrests, none were for terrorism-related offenses (see Table 2). Moreover, at least 16 people have been known (after the fact) to have been involved in terrorism plots and have traveled through airports in which SPOT has been employed (see Footnote 92). Yet, 93% of 94 SPOT Instructor evaluations accomplished in March 2009 found instructors to "exceed expectations"—the highest rating possible based on a 3-point scale over 57 items (Table).

Basic Challenges to Behavior Detection. The predictability of complex social behavior in a mass screening mode may not be attainable through time. The scientific and linguistic assumptions on which

behavior detection are based—viz., a philosophy of science based on logical positivism and the language of variables, causality, logic, material premise, and statistical inference so successful in predicting many aspects of the physical world—may be inappropriate for the counterterrorism task at hand. This is why a recent report (see William H. Press, February 2009, "Strong profiling is not Mathematically Optimal for Discovering Rare Malfeasors," PNAS Proceedings of the National Academy of Sciences of the United States of America 106(6)) supports at least some usage of random screening in mass security programs. On the other hand, behavior detection—as with the polygraph—likely will work as a deterrent against people who believe in its validity. So, the response from the Department of Homeland Security (DHS) to 11 specific recommendations from the GAO was to concur on all of them. This may have been a refreshing interagency concurrence. But what else could DHS do? (Comment may be sent to bloomr@erau.edu).

Keywords: Airport, Aviation, Behavior Detection, Profiling, Security, SPOT, Terrorism