How to \textit{Really} Improve Airport Security

Until human factors are considered in the design of airport security jobs, technological responses to security breaches will fail to achieve their desired end.

\textbf{BY DOUGLAS H. HARRIS}

Current threat detection systems were implemented at airports in the early 1970s in response to increased hijackings of commercial aircraft prior to that time. These systems have been in place relatively unchanged since then. For example, screening procedures continue to focus on detecting metallic objects on the person or in passengers’ carry-on baggage. Human operators continue to play a key role in the detection process by examining X-ray images, resolving metal detector alarms, conducting body scans with metal detection wands, conducting physical searches of baggage, and maintaining order at screening checkpoints.

Now, after the terrorist attacks of September 11, 2001, the effectiveness of these systems has been called into question, even though the hijackings of that day did not involve any screening failures (hijackers carried only authorized items when boarding the aircraft). Since September 11, elected officials, media pundits, and security experts have been critical of airport threat detection systems, using strong language to criticize screening personnel for their inattentiveness, lack of motivation, and inappropriate employment histories, such as criminal records (for example, “Questions About,” 2001; “Weapons Still Elude,” 2001). Subsequent detection failures have led to heightened criticism and punitive actions against screening personnel and their employers.
What is required is a system sufficiently flexible to meet anticipated security threats with timely and appropriate responses.

Elected and appointed government officials have prescribed increased pay, more and better training, background checks, and federal employment to improve the proficiency of threat detection personnel. Seldom, however, has the fundamental problem of these systems been mentioned: that they require people to perform tasks, under performance-degrading conditions, that they are ill-suited to perform in the first place. Any significant improvements, therefore, must first require better integration of human operators and technology within threat detection systems.

In this article, I attempt a scientific assessment of a subject that has been the victim of emotional rhetoric, conventional wisdom, and the bias of special interests. Consider, first, the characteristics of an acceptable threat detection system that must screen more than 1.5 million airline passengers each day. According to a panel of the National Research Council that studied airline passenger security screening (National Research Council, 1996), an ideal screening system would detect both metallic and nonmetallic threat objects with a high degree of accuracy (high detection rate combined with a low false-alarm rate) in less than 6 seconds and provide sufficient information to system operators to permit the speedy and accurate resolution of detection alarms.

Although funding, technology, and human limitations may preclude developing an ideal system anytime soon, here I propose science-based, practical approaches that can and should be taken now to improve threat detection at airports. These approaches encompass the screening of passengers, carry-on baggage, and checked baggage for threat objects (weapons and explosives).

Threats at Airports: The Big Picture

Passengers and baggage are not the only sources of threats to commercial aviation security. Threats can also come from the many processes that support an airport and the passengers and aircraft it serves: catering, maintenance, cleaning, ticketing, baggage handling, air traffic control, retail, food services, parking, car rental and others. For example, members of the cleaning staff stashed guns and grenades in the plane’s washroom to support the hijacking of TWA Flight 847, which led to 17 days of terror in 1985 (Gladwell, 2001). Thus, even perfect threat detection applied to passengers and their baggage would not necessarily result in acceptable levels of security.

Moreover, the history of airport security suggests that tightening one component leads determined terrorists to seek and exploit weaknesses in other components (Gladwell, 2001). But attempts to plug all the potential holes in airport security would so burden and disrupt the already overstressed airport and air traffic system that the terrorists would be handed the victory they seek. Such attempts might also have the unintended result of creating a rigid system, analogous to the Maginot Line built before World War II to protect the eastern border of France but easily outflanked by German invaders. (One might say that the terrorists of September 11 outflanked the screening systems by using permissible weapons – small knives and box cutters – to execute the hijackings.) What is required, instead, is a system sufficiently flexible to meet anticipated security threats with timely and appropriate responses. Such a system would consist of the following integrated and coordinated components:

- security teams of selected, cross-trained, motivated personnel performing appropriate tasks, with continuing measurement and feedback of their performance;
- an arsenal of technologies and procedures that can be quickly configured and deployed to meet probable threats; and
- timely intelligence continuously disseminated to security teams on probable adversaries and threat scenarios.

Appropriate Roles for Human Operators

Current approaches to screening passengers and carry-on baggage require operators to perform tasks for which humans are poorly equipped: the detection of rarely occurring, low-signal-to-noise-ratio signals embedded in the context of varying background configurations. An extensive body of research on human vigilance has led to the unavoidable conclusion that humans are poor monitors and that a variety of interventions and countermeasures directed toward improving monitoring performance in different settings and tasks have demonstrated little benefit (Davies & Parsuraman, 1982; Mackie, 1987). Consequently, a primary objective in the design of security systems should be to remove the operator from this monitoring role. Instead, screening machines should be designed to monitor, detect, and alert operators to potential anomalies, and procedures need to be developed to enable operators to determine whether or not the anomaly is a threat object. That is, the monitoring/detection/alarm function should be allocated to machines and the alarm resolution function to human operators.

Human operators have special – even unique – capabilities needed for alarm resolution, such as pattern recognition, abstract reasoning, spatial visualization, and cognitive flexibility. As a consequence, an appropriate (even necessary) role for human operators is in the resolution of detected signals. However, even in this role, the combined effects of the design of the specific operator tasks, the environment in
which the tasks are performed, and the selection and training of operators will be critical to the realization of effective performance. Using human operators appropriately in threat detection systems requires systematic efforts to match their capabilities with the requirements of the system and to minimize the effects of human limitations.

The CTX-5000, the first explosives-detection system for checked baggage to receive Federal Aviation Administration certification, provides an example of the allocation of functions as just described. Though originally conceived as a fully automated system, the CTX-500 ultimately incorporated the alarm resolution capabilities of human operators to meet the certification standard during field use.

Paradoxically, as threat detection systems become more automated, human integration issues are likely to become even more important to the successful implementation of these systems. Human operators will be performing the more difficult and complex tasks that defy automation, such as alarm resolution. Moreover, the increased automation will introduce new performance issues that will need to be addressed.

For example, an operator directly involved in X-ray signal detection necessarily monitors the performance of the X-ray machine. Degraded images easily detected by the operator lead to appropriate equipment calibration or maintenance. If detection is automatic, on the other hand, the operator will need to have some other means of monitoring and ensuring that the machine is operating properly. Such problems introduced by automation can be resolved if they are recognized and addressed in the system design process (for example, see Parasuraman, Mollo, Mouloua, and Hilburn, 1996; Rasmussen, 1986; Reason, 1990; Wiener, 1988).

Pay and Performance

The low wages paid to screeners are frequently cited as a principal contributor to poor operator performance. It is argued that higher wages would attract “better” people to these jobs and, in turn, lead to better performance. However, increased pay alone is not likely to increase operator performance directly. A substantial body of evidence suggests that any linkage between pay and performance is tenuous at best and is probably insignificant in the face of much more powerful determinants of operator performance, such as job design, performance measurement and feedback, and the match between operator aptitudes and tasks (for example, see Filipczak, 1996; Guzzo, 1988; Guzzo, Jette, & Katzell, 1985; Hertzberg, 1968; Lawler, 1971, 1981).

The figure below illustrates a framework for understanding the relationship between pay and performance. Attaining job proficiency involves the application of two kinds of factors: those that attract and keep people on the job and those that enhance performance when they are on the job. Compensation (wages and benefits) is one of several factors that can attract and keep people on the job. Higher levels of some factors can compensate for lower levels of other factors; for example, higher pay might make up for poor working conditions or low job or organization status. Also, higher pay might attract sufficient numbers of appli-
The systematic measurement and feedback of operator performance has been missing from threat detection jobs at airports throughout the entire history of airport security in the United States.

how soon they would stop bowling altogether if there were no system of keeping score and no feedback on how many and which pins were knocked down with each roll of the ball. The huge sports industry – players, teams, leagues, fans, tournaments, records, statistics, magazines – is based on measuring performance and providing feedback (scores, standings, winners, champions). The power of this process in military and industrial jobs has been known for decades; when it is absent, performance suffers and personnel turnover increases (Harris & Chaney, 1969; Howell & Cooke, 1989; Ilgen & Klein, 1988; Swezey & Salas, 1992). Operators cannot be expected to maintain high levels of detection performance day after day, month after month without consistent, regular, realistic, accurate measurement and feedback of their performance. Yet, the systematic measurement and feedback of operator performance has been missing from threat detection jobs at airports throughout the entire history of airport security in the United States. (The crude, unrealistic system employed to penalize airlines for poor security performance is inadequate to meet these requirements.) In addition to providing a continuing basis for motivating and improving operators, performance measurement and feedback can also serve to enhance proficiency of the overall threat detection system by

- diagnosing performance weaknesses and tailoring performance-based training to overcome the weaknesses identified,
- assessing the impact of new detection technologies and procedures, and
- providing criteria for validating operator selection and job assignment techniques.

Some promising technologies, such as Threat Image Projection (TIP) and the Screener Proficiency Evaluation and Reporting System (SPEARS), have been available for several years to support on-line performance measurement and reporting. I completed assessments of these technologies and others discussed in this article while I was a member of the National Research Council Committee on Commercial Aviation Security. TIP involves the electronic insertion of images of threat objects into operational detection systems during ongoing baggage screening at checkpoints. The images inserted can be configured to represent any threat object within the contents of any type of bag that might pass through the system and can be inserted on any desired frequency or schedule. SPEARS is designed to generate measures and reports based on data obtained from operator responses to TIP images. These approaches need to be fully developed, field-tested, and implemented in support of operator performance enhancement.

The ultimate success of on-line performance measurement for improving operator performance will have as much to do with the manner in which systems are introduced and employed as with the effectiveness of the systems themselves. Critical considerations include the detailed procedures needed for collecting and analyzing data, the extent to which the measurement system intrudes into operator screening tasks, and the degree of understanding by operators of how the measures are to be obtained and used. Rarely, for example, should the results obtained from any on-line performance measurement be employed as the basis for punitive actions, such as fines, penalties, reprimands, or other punishments. Punitive actions based on performance measures will inevitably lead to subversion of the measurement process, contamination of the results obtained, and resistance to remedial measures that might be introduced to enhance performance.
Selection of Threat Detection Personnel

Assuming that incentives are increased sufficiently to attract people to airport threat detection jobs and that jobs are designed to be compatible with human capabilities, further gains could be realized by selecting people with the highest aptitudes for these jobs. The development of selection techniques would start with a detailed job analysis to identify the tasks that are performed, their frequency, and their importance to successful job performance. Then the critical skills required for satisfactorily performing the frequent and important tasks would be identified. Finally, the aptitudes necessary for the critical skills would be identified, and measures of these aptitudes would be developed and incorporated into a selection test or test battery. Validation, in which test scores are correlated with actual job performance, would indicate how the test should be used for personnel selection.

Recent development and validation of a test designed to predict success on inspection jobs requiring the combined aptitudes of general cognitive ability, abstract reasoning, and spatial visualization (Harris & Spanner, 2001) serve as an illustrative example. The jobs involved application of ultrasonic techniques to detect defects in piping and containment vessels in nuclear power plants. The test was made up of 36 items, each consisting of a task requiring the application of the three aptitudes; it was computer administered and scored and required 36 minutes of testing time. The correlation coefficient between test scores and the performance measures used to qualify candidates for critical inspection jobs was 0.51. The practical significance of this degree of correlation was that 92% of those with test scores of 25 or greater qualified for the ultrasonic inspection jobs, whereas only 25% of those with test scores of less than 25 qualified. This translated into a cost savings of more than $1 million for each 100 qualified ultrasonic inspectors available to the nuclear industry.

Smarter, Focused Threat Detection

Shortly after September 11 and the subsequent increase in airport security, while boarding a flight in Atlanta for the third leg of a trip from Santa Barbara, California, to Daytona Beach, Florida, I was called to the ticket counter and directed to a table for a search of my carry-on computer bag. A lengthy, intensive inspection of the bag ensued, during which pens were uncapped and recapped, rubber erasers were flexed, business cards were shuffled, note tablet pages were riffled, a mini-flashlight was turned on and off, a floppy disk was examined from all sides and edges, and so on. Then the bag, containing the examined items plus the unexamined computer, computer transformer, connecting cord, data wire, and mouse, was returned to me, and I carried it aboard the aircraft. Access to computer and accessories was through not-so-obvious zippers at the ends of the bag.

The story illustrates two problems. First, an unfocused inspection is likely to be ineffective – in this case, an attempt to conduct a thorough examination of everything in the bag led to overlooking the most likely threat objects the bag contained. Second, the owner of the bag, a frequent flyer who had completed two flights already that day in larger aircraft without committing a terrorist act, was probably not a high-risk passenger.

Efforts should be directed toward making airport security systems smarter and more flexible, not necessarily more encompassing or extensive. Intelligence on threats and adversaries can be the basis for configuring a flexible system to meet those threats and to focus enhanced security measures mainly on high-risk individuals. To this end, channels of information and coordination must be established between airport security systems and the intelligence-gathering and -analysis activities of national and local law enforcement agencies.

Another source of intelligence is the Computer-Assisted Passenger Prescreening System (CAPPS). When a ticket is purchased, CAPPS can rate passenger risk on the basis of about 40 characteristics, including address, credit history, destination, travel companions, type of payment, and type of trip. Special technology, such as electronic strip searches, that are too time-consuming to be used with all passengers could be employed to screen those identified as high risk. Other procedures not involving technology, such as personal interrogation and assessment, could also be employed with the few passengers identified as high risk. It appears, in retrospect, that sufficient intelligence was available on September 11 to deter the hijackings, but the security system was not designed to apply it to threat detection at airports.
A Final Note
As this article was being prepared for publication, the U.S. Department of Transportation was initiating the process of selecting and training more than 30,000 airport screeners for deployment over the next 10 months. The stated assumption is that stringent employment standards, combined with a minimum of 40 hours of classroom and 60 hours of on-the-job training and a proficiency examination for each selected screener, will significantly improve threat detection at airports (Aviation and Transportation Security Act, November 19, 2001). This massive and costly effort will have little impact unless it is accompanied by significant changes in the jobs these screeners will be performing.

References

Douglas H. Harris is chairman and principal scientist of Anacapa Sciences, Inc., P.O. Box 519, Santa Barbara, CA 93102-0519. He has specialized in improving human performance on inspection, investigation, and intelligence tasks and has served on the National Research Council’s Committee on Commercial Aviation Security and Panel on Airline Passenger Security Screening.