U.S. Department of Transportation
Office of the Secretary of Transportation

The Inspector General

Office of Inspector General
Washington, D.C. 20590

August 16, 2002

The Honorable Elaine Kaplan
Special Counsel
U.S. Office of Special Counsel
1730 M Street, NW, Suite 300
Washington, DC 20036

Dear Ms. Kaplan:

This is in response to your letter of February 5, 2002, that referred certain aviation security allegations brought by Bogdan Dzakovic, a former member of the Federal Aviation Administration's (FAA) covert security testing unit (known as the Red Team), to the Secretary of Transportation for further investigation. Mr. Dzakovic is now employed by the Transportation Security Administration (TSA) and is assigned to its General Aviation Directorate. At present, he is temporarily performing duties assisting in the orientation of airport Federal Security Directors.

Secretary Norman Mineta delegated your request to our office for investigation and subsequent response to you. Presented herein are the results of our investigation of the predicate allegations concerning the Red Team and the Threat Image Projection (TIP) software program. Our investigation entailed lengthy interviews with Mr. Dzakovic; review of his filing to your office; interviews of current and former FAA employees, as well as employees of the new Transportation Security Administration (TSA); and extensive examination of pertinent documentation. This response has been prepared to omit any information or data classifiable as Sensitive Security Information (SSI).

You referred the following allegations to the Secretary for investigation:

- Mr. Dzakovic alleged that Admiral Cathal Flynn, then-Associate Administrator for Civil Aviation Security (CAS), and Leo Boivin, then-Manager of the Red Team, suppressed Red Team testing results and directed the Red Team to not conduct follow-up inspections of airports that yielded especially poor testing results. Specifically, for example, Mr. Dzakovic alleged that in 1997 or 1998, he reported to Mr. Boivin that during a testing trip to the San Juan, Puerto Rico

airport, he observed that the airlines suspended many security measures at times when the airport was particularly busy. Mr. Dzakovic alleged that rather than initiating steps to correct this problem, Mr. Boivin instructed him not to return to that airport in the future. Mr. Dzakovic further alleged that contrary to the Red Team's usual practice of submitting written reports after each testing trip, Mr. Boivin told him not to write a report on his visit to the San Juan airport. Mr. Dzakovic suspected that Mr. Boivin attempted to suppress this information because it reflected unfavorably on the airline industry.

- Mr. Dzakovic also alleged that in August 1999, Mr. Boivin ordered him to begin providing prior notification to FAA local field offices in advance of visiting airports for CTX™ Explosives Detection System (EDS) testing, in violation of the Red Team's protocol of unannounced testing. Mr. Dzakovic asserted his belief that his management instituted pre-notification for the purpose of forewarning the airlines and improving testing performance, thus making CTX appear more effective. Mr. Dzakovic related that he followed Mr. Boivin's orders on two testing trips, but after observing a large discrepancy in the resulting data, he independently decided to return to his previous practice of conducting unannounced testing.

- Mr. Dzakovic alleged that Tony Fainberg, then-Director of the CAS Office of Policy and Planning (located in FAA headquarters), abused his authority by instructing Mr. Dzakovic to exclude data from a study conducted on the reliability of TIP. Mr. Dzakovic related that in 1998, FAA sent him to Reno, Nevada, to conduct a small-scale study on the accuracy and reliability of TIP. He believed this study was the only one of its kind conducted and that it was used by FAA to evaluate whether TIP would be a worthwhile investment to improve airport security. Mr. Dzakovic explained that the study compared the ability of screeners to recognize TIP computer-generated weapons on X-ray monitors against their ability to recognize real weapons. Mr. Dzakovic alleged that Mr. Fainberg instructed him to exclude data on one of the screeners from the final results he reported because this individual scored high in her ability to recognize real weapons, but performed poorly with the TIP-generated images. Mr. Dzakovic suspected that Mr. Fainberg directed the exclusion of this data in order to make TIP appear more effective.

Summary of Findings & Recommendations:

Red Team

In brief, our investigation did not disclose evidence to substantiate the specific allegations of cover-up or suppression of Red Team findings by CAS management. Further, while we did find that Mr. Boivin had instructed the Red Team to pre-

notify FAA Federal Security Managers in advance of CTX testing, we concluded that it was done for legitimate purposes, and Red Team records show a slightly higher test failure rate after pre-notification was instituted. Additionally, our investigation did not confirm the alleged improprieties on the part of Mr. Boivin during Red Team testing in San Juan.

However, we did find considerable merit to other concerns Mr. Dzakovic raised relative to coordination between Red Team management and other CAS elements responsible for providing follow-up to Red Team findings. In particular, we found programmatic weaknesses involving the reporting of Red Team findings and regarding corrective action. Further, we found that changes made by FAA as a result of Red Team testing generally did not have the desired effect of creating sustained improved performance by airport screening companies.

We determined that while follow-up testing generally was not within the mission charter of the Red Team, it did, on occasion, return to certain airports for additional tests. We found that the Red Team submitted, to CAS headquarters, a written report of findings for each of its missions. CAS headquarters in turn forwarded Red Team summary reports to CAS field units—having regulatory responsibility for direct, day-to-day oversight of airport security—for follow-on remedial action (e.g., letters of correction (LOCs) to air carriers, and fines). However, the field units typically did not receive LOCs that headquarters unilaterally issued to air carriers, or the carriers' LOC responses back to headquarters. We further found that CAS headquarters did not track the resolution of deficiencies identified through Red Team testing. We concluded that such lack of information dissemination and tracking organizationally hindered CAS' capacity to effect coordinated remedial action.

We note that the difference between the nature of the mission of FAA's regulatory component and that of the Red Team fostered the disjointed process for effecting follow-on action to Red Team findings. Specifically, prior to federalization of security screening, CAS field units had regulatory authority to initiate civil enforcement action (e.g., fines) against carriers. However, to be legally enforceable, compliance testing by the field units—which served as the basis for enforcement action—had to comport with standardized criteria known to the airlines' screening contractors. The FAA told us that based on the Red Team's creative "out of the box" approach to testing, its findings were deemed not suitable for civil enforcement proceedings, because the Red Team's techniques went beyond the standardized criteria applicable to the field regulatory units.

Additionally, we found that while there was a fundamental Concept of Operations document from 1994, the Red Team lacked standard operating procedures clearly

governing the conduct of its operations and the use of Red Team reports to improve airport security.

Significantly, as noted by Mr. Dzakovic, the Red Team consistently found and reported—throughout its existence—high rates of test failure, reflecting often stark localized and systemic security vulnerabilities. Following September 11, 2001, FAA's Red Team stood down and our office was requested by the President and the Secretary, on an interim basis, to conduct similar covert testing nationwide. We, too, have found an alarmingly high incidence of testing failures, consistent with reports we have issued over the last several years showing vulnerabilities in screening of passengers; checked and carry-on baggage, as well as cargo; access to secure areas of airports; and issuance and control of airport identification badges. The persistence of these problems and lack of sustained improvement in aviation security led to the legislation federalizing security screening under TSA, along with the full range of measures that are being planned and implemented by TSA to strengthen screening operations.

As referenced above, in response to the Red Team's findings, FAA did take some follow-on actions, namely letters of correction to air carriers and fines. However, these follow-on actions were not readily visible, and, given the consistently poor results of testing over time, the intended outcome of sustained improvement in airport security was not apparent. The experience of FAA's Red Team is instructive of the critical importance of training, evaluation, and meaningful follow-on corrective action for identified deficiencies.

Based on our observations regarding the manner in which the Red Team was organized and functioned before September 11, 2001, there are a number of steps that TSA can take to improve the effectiveness of its successor to FAA's Red Team. A robust covert testing program, such as the Red Team, is essential for effective oversight of airport security. We have assisted TSA in the process of developing such a program and TSA has already initiated some testing. In addition, TSA is developing means to integrate lessons learned through covert testing into a standardized, formal screener training process. In our view, the challenge for TSA will be to translate the findings of its covert testing program, in a well-managed manner, to substantive enhancements in key areas such as screener training, screener performance/accountability measures, technology applications, and local testing performed by TSA's field regulatory element.

We have recommended to the Under Secretary of Transportation for Security that TSA, in developing its covert testing program, ensure the following provisions are implemented: (a) detailed standard operating procedures addressing operational considerations for testing; (b) procedures for applying the results of covert testing to substantive enhancements in key areas such as screener training, screener

performance/accountability measures, technology applications, and TSA local testing; and (c) a reliable mechanism for providing meaningful feedback to testing team members about actions taken as a result of their efforts.

TSA has concurred with these recommendations and its new covert testing program will include the foregoing elements. This program operational guidance was recently reaffirmed by the Acting Under Secretary of Transportation for Security.

Threat Image Projection (TIP)

TIP is a software program, developed in the early 1990s, that superimposes the images of fictional threat objects on the monitors of X-ray screening machines for the purpose of keeping screeners alert, helping screeners recognize a variety of potential threat objects, and assessing screener performance. We note that TIP was not designed to detect actual explosive devices or threat items.

While our investigation did not disclose evidence that Mr. Fainberg directed the exclusion of testing data during the 1998 study of TIP in Reno, we identified weaknesses in the methodology used for the study. However, we note that at the time of Mr. Fainberg's Reno study, TIP was still in the research and development stage, having since undergone multiple design and technological improvements. Further, we found that in 2000, FAA's Technical Center in Atlantic City, New Jersey, conducted an evaluation of TIP, which demonstrated TIP's usefulness as an additional tool to measure a screener's ability to detect threat objects.

Importantly, we note that TIP was neither intended to be, nor is, a substitute for covert testing with realistic physical threat objects—using innovative testing techniques. Although TIP is undergoing continued modification, to include expansion of the threat image library, development of more sophisticated TIP images is needed. For example, images projected by TIP need to more closely depict what screeners may actually face (e.g., objects partially obscured by clutter in bags). We are continuing to monitor the deployment and future development of TIP.

Background:

Red Team

The Red Team was an FAA Office of Civil Aviation Security (CAS) headquarters-based unit created consequent to the 1988 bombing of Pan Am Flight 103. Its primary mission was to conduct covert airport security penetration testing for the purpose of identifying both localized and systemic vulnerabilities, and to help strengthen FAA's regulatory inspection capabilities. The Red Team functioned as

part of a larger special assessment staff, whose responsibilities included addressing the Red Team's findings within the context of developing enhanced security measures and formulating long-range strategic policy. The Red Team consisted of approximately 4-8 personnel at any given time, and assignment to the unit was considered prestigious within the CAS directorate. Mr. Dzakovic joined the Red Team in 1995, having previously served as a Federal Air Marshal since 1987.

The Red Team's covert penetration testing was separate and apart from FAA's day-to-day regulatory oversight of airport/air carrier security provided by specialists assigned to CAS Field Offices (CASFOs) nationwide, and by CAS Liaison Offices (CASLOs) in international locations[1]. Until enactment of the Aviation and Transportation Security Act of 2001, air carriers were accountable for the compliance of their employees and security screening contractors with FAA security regulations. Testing conducted by the CASFOs and CASLOs was subject to standardized FAA protocols (i.e., rules of engagement in the interest of fairness to the carriers and their contract screeners), and thus was not as rigorous as the Red Team's more creative "out of the box" approach to testing. For example, CASFOs/CASLOs were required to place a simulated explosive device in an uncluttered bag, oriented in such a way that it would be readily visible to X-ray machine operators. Conversely, the Red Team would disguise the testing device within a cluttered bag[2]. The Red Team's techniques are similar to those we have employed in conducting our covert testing.

TIP

As previously noted, TIP is a software program, developed in the early 1990s that superimposes the images of fictional threat objects on the monitors of X-ray screening machines for the purpose of keeping screeners alert, helping screeners recognize a variety of potential threat objects, and assessing screener performance. Early iterations of TIP experienced operational difficulties, some of which were attributable to images being projected at set time intervals. Problems encountered included screeners being able to time when TIP images would appear, images projected when there were no bags in the X-ray machines, and object images larger than the bags being screened. The above difficulties have largely been resolved through technological advancements.

---

[1] CASFOs and CASLOs, which now reside within TSA, are responsible for overseeing airport/air carrier security operations, to include records inspection, testing, investigation of potential deficiencies identified, and effecting compliance/enforcement action (e.g., letters of correction, civil penalties/fines).

[2] Based on the "no rules" nature of the Red Team's testing techniques—which were more in keeping with actions that might be taken by terrorists and were not subject to the standardized protocols that limited the CASFOs and CASLOs—FAA held that Red Team findings were not suitable for civil enforcement proceedings.

The FAA issued functional requirements for TIP in 1996, resulting in three firms developing machines that FAA subsequently approved as TIP-Ready X-ray (TRX) machines. Since that time, TIP has continued to evolve, with additional threat images being added and frequencies of image projections adjusted. Another generation of TIP is currently being developed at FAA's Technical Center in Atlantic City, New Jersey.

Details:

Alleged CAS management suppression of Red Team testing results and direction not to conduct follow-up testing

- As reflected in the below sections, our investigation did not disclose evidence that Admiral Flynn or Leo Boivin suppressed or covered-up any Red Team testing results. During detailed interviews we conducted, both individuals (now retired) adamantly denied the allegations.

  We determined that follow-up inspections generally were not within the mission charter of the Red Team, though some were conducted. Admiral Flynn, Mr. Boivin, and other former Red Team members told us—consistent with the Red Team's 1994 Concept of Operations document—that it was not the prescribed role of the Red Team to perform follow-on testing, but that the Red Team did return to certain airports on occasion as follow-up. For example, we found that subsequent to testing at San Juan in 1998, the Red Team returned to that airport in 1999, 2000, and 2001. In another example, the Red Team performed testing at London-Heathrow in 1996, followed by a series of testing missions there in 1999, and it again returned in 2001[3].

  We also found that the Red Team prepared and submitted to CAS headquarters a written report of findings for each of its testing trips and that CAS headquarters provided reports, albeit of a summary nature, to CASFOs/CASLOs for any follow-up actions those offices deemed appropriate. We further address the issue of Red Team reports below.

- Regarding the specific allegation of improprieties on the part of Mr. Boivin relative to a 1998 Red Team testing mission at the San Juan airport, Mr. Dzakovic provided us with the following account of events:

  Mr. Dzakovic advised that in April 1998, an airline failed a pass-through positive passenger baggage match (PPBM) that had been tested at the San Juan

[3] International testing by the Red Team was subject to approval of the host government. A number of countries serviced by U.S. flag carriers disallowed such testing.

airport[4]. He related that when Carrie Hancasky, a fellow Red Team member, approached the airline station manager about the test failure, she was informed that there had been a power outage in the automated passenger/baggage tracking system (known as Sabre), thus, no baggage match procedures were being conducted at the time. According to Mr. Dzakovic, the airline manager then commented to Ms. Hancasky to the effect of, "When it gets this busy, we can't do security." Mr. Dzakovic told us he called Mr. Boivin to apprise him of the situation and the airline manager's remark to Ms. Hancasky, and was directed by Mr. Boivin to discontinue testing and not document the manager's remark. Mr. Dzakovic related that he did write a report for this trip, citing the testing findings, but, in accordance with Mr. Boivin's instruction, omitted the airline manager's remark to Ms. Hancasky.

We interviewed Ms. Hancasky, who advised that automated PPBM procedures were not in effect due to the Sabre system power outage, but that the airline claimed to have conducted manual PPBM checks. Ms. Hancasky did not recall the airline manager comment that security was not being performed because it was too busy. Ms. Hancasky told us she had not spoken with Mr. Boivin during the San Juan trip, but was informed by Mr. Dzakovic that Mr. Boivin had directed that additional testing not be conducted at that time.

Mr. Boivin told us he did not recall this specific testing mission, but denied that he would have directed the omission of the airline manager's reported remarks from the mission report. He further asserted that based on the potential seriousness of such a remark, he would have instructed the Red Team members to report this incident to the local CASFO for investigation and potential enforcement action.

We reviewed the Red Team's final report for this testing mission, and confirmed that the airline manager's alleged remark was not present in the report. Further, the report reflects that testing was suspended pending resolution of the situation involving the airline's out-of-service Sabre system and an issue of whether Ms. Hancasky's checked bag ever left Miami. The report also notes that details of the situation were provided to the San Juan CAS Field Unit (CASFU—subordinate to the larger CASFO). We learned that the CASFU investigated and concluded that Ms. Hancasky's bag never left Miami; however, the CASFU subsequently issued a Letter of Correction to the airline for failing to follow proper manual PPBM procedures.

[4] The pass-through PPBM test scenario involved a Red Team member, Carrie Hancasky, booking a flight from Miami to Barbados with a stop in San Juan. Ms. Hancasky deplaned in San Juan and did not re-board the aircraft. The airline was required to note the passenger's failure to re-board and offload the passenger's checked baggage.

Alleged management direction to provide advance notification of CTX testing to FAA field personnel, in violation of Red Team testing protocol

- We determined that although Mr. Boivin required Red Team members to notify FAA Federal Security Managers (FSMs)[5] in advance of CTX testing, our analysis of Red Team testing data showed no appreciable difference in testing results before and after FSM notifications began in August 1999; in fact, we found test failure rates to be slightly higher after the notifications commenced.

  Mr. Boivin told us he instituted FSM pre-notification as a cost-saving measure based on the Red Team's experience of arriving at airports unannounced, only to find CTX machines not operational because they were out of service for maintenance or there was no trained, qualified operator on-site.[6] He advised that pre-notifications of FSMs were made exclusively for CTX tests and not for any other form of Red Team testing.

  Mr. Boivin explained that pre-notification consisted of contacting an FSM and inquiring as to whether there was a CTX machine in place, whether it was operational, and whether trained operators were available. He advised that FSMs were not to be informed of exact dates and times of testing. Several Red Team members we interviewed corroborated Mr. Boivin's statements.

Other relevant Red Team findings

- We found a marked lack of feedback to Red Team members about actions resulting from their testing, fostering a belief on the part of Mr. Dzakovic and other team members that CAS management had not pursued corrective action for test failures. This is particularly evidenced by Mr. Dzakovic's assertion that no material corrective action had apparently been taken after the Red Team's 1996 testing at Frankfurt, Germany, in which all tests conducted resulted in failures.

  We found that although follow-on actions were taken at Frankfurt, as in other instances, it was not communicated to Red Team members. Specifically, Robert Blunk, the then-CAS Regional Division Manager for Europe, Africa, and

[5] Federal Security Managers (FSMs) were CAS specialists assigned as on-site liaisons to large airports and resident air carriers for the purpose of monitoring and coordinating security activities, as well as assisting in the development and implementation of comprehensive security plans.

[6] Until 1997, CTX usage by the airlines was voluntary. From 1997 to September 2001, CTX usage was required for passenger baggage selected for additional screening by the Computer-Assisted Passenger Pre-screening System (CAPPS), and then only if there was a machine available, a qualified operator, and the machine was functional. In many airports, the number of CTX units available for testing was limited.

the Middle East, told us that as a result of the Red Team's 1996 testing, inspectors from his office worked with the affected air carriers to correct the identified deficiencies. Further, according to Mr. Blunk, as a result of the Red Team's testing, he had meetings with the then-senior most official in the German Ministry of Transportation to stress the need for a multi-layer approach to security. Additionally, Mr. Blunk and Admiral Flynn told our office that the Red Team's test results from Germany were used to push for more advanced technology for the detection of explosives in checked baggage.

We obtained information reflecting that in early 2001, Lynne Osmus, the then-Deputy Associate Administrator for CAS, initiated steps to improve feedback to, and solicit input from, Red Team members, as well as enhance dissemination of Red Team findings.

However, as previously noted, while FAA did take action based on Red Team reports, these actions did not have the desired cumulative effect of improved sustained performance by the screening companies. Specifically, we found—through review of Red Team reports and in our own testing—that many of the vulnerabilities reported by the Red Team, as well as by our audits, still existed immediately following September 11, 2001.

- Our review of Red Team records disclosed that the Red Team prepared and submitted to CAS headquarters a written report of findings for each of its testing trips and that CAS headquarters provided reports, albeit of a summary nature, to CASFOs/CASLOs for any follow-up actions those offices deemed appropriate.

  However, we found that although Red Team summary reports were consistently provided to CASFOs/CASLOs, those offices typically did not receive copies of the Letters of Correction (LOCs) sent to the carriers by CAS headquarters' Air Carrier Liaisons, or the carriers' LOC responses back to CAS headquarters. In our view, receipt of these documents by the CASFOs and CASLOs—charged with responsibility for providing direct, day-to-day security oversight—would have enhanced their ability to provide effective follow-up. Further, we believe that coordination and tracking of the resolution of deficiencies identified through Red Team testing, by CAS headquarters, would have led to consistently improved performance by the screening companies.

- We found that the Red Team lacked standard operating procedures clearly governing its operations. Specifically, we found that the only document addressing the mission, operation, and activities of the Red Team was a loosely formulated "Concept of Operations" issued by Mr. Boivin in 1994. This charter document did not describe, with any specificity, such operational considerations as the types of tests to be conducted; criteria for the selection of testing

locations; content/format of reports; report dissemination; and responsibilities for follow-up.

Alleged management manipulation of TIP study data

- When interviewed, Mr. Dzakovic related to us that he had been tasked, as a member of the Red Team, to assist in Mr. Fainberg's study of TIP in November 1998 and January 1999. Mr. Dzakovic advised that Reno, Nevada, was chosen as the test location because it had been the location of the first field deployment of TIP. Mr. Dzakovic said that the study involved using the five screeners with the best TIP scores along with the five screeners having the worst TIP scores, and then comparing those results with screener scores for detection of physical testing objects. Mr. Dzakovic explained that for TIP to be proven effective, the five screeners having the best TIP scores should also be the five screeners with the best scores for detecting physical objects.

  Mr. Dzakovic explained that during the comparison of scores, it was discovered that a screener having one of the worst TIP scores performed exceptionally well in detecting physical threat objects. Mr. Dzakovic alleged that while in Reno, Mr. Fainberg, then-Manager of the FAA/CAS headquarters' Office of Policy and Planning, directed him to exclude this screener from the body of data because, otherwise, her performance would negatively skew the data. Mr. Dzakovic expressed his belief that Mr. Fainberg excluded this screener from the data pool because Mr. Fainberg wanted to make TIP appear more effective than it really was. Mr. Dzakovic believed that this study in Reno was the only evaluation conducted to justify the deployment of TIP.

- We subsequently interviewed Mr. Fainberg, who denied excluding any testing data from his analysis. He asserted that, statistically, one screener deviating from the norm would not have had an appreciable effect on the final conclusion. Therefore, Mr. Fainberg maintained that he would have had no motive in excluding the alleged data.

- David Hobbs, a former Red Team member who served as team leader for the 1998 Reno study, told us that, to his knowledge, none of the screeners tested were excluded from the body of data. He advised that he was unaware of anything unusual during the Reno testing.

- Eric Neiderman, Project Manager and Engineering Research Psychologist at FAA's Technical Center, told us that a validation study of TIP was conducted at the FAA's Technical Center in 2000. Mr. Neiderman said that he was aware of Mr. Fainberg's analysis of TIP, but related that Fainberg's study was simply a paper study that was not definitive. Mr. Neiderman explained that

Mr. Fainberg's testing was not meant to be anything more than a basic assessment, a piece of data to demonstrate the relationship between TIP and other testing data.

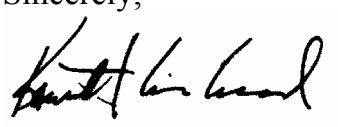
Mr. Neiderman stressed that TIP is continuing to evolve. He advised that there have been a number of improvements to TIP since the first generations were fielded, including about twice as many threat images that can be projected into bags. Additionally, Mr. Neiderman advised that TIP technology is continuing to advance as more images are installed and newer technology is identified and tested.

- In reviewing Mr. Fainberg's 1998 study, we identified weaknesses in its methodology. For example, the TIP data had not been independently recorded contemporaneous to the Red Team's covert testing with physical objects. Rather, Mr. Fainberg relied on TIP data collected by the airport screening company well in advance of the study.

- We did not uncover evidence to substantiate Mr. Dzakovic's allegation that Mr. Fainberg excluded a data point from the testing results. Further, TIP was still in the research and development phase during Mr. Fainberg's 1998 study of TIP in Reno. Since that time, TIP has continued to evolve, including additional threat images being added and frequencies of image projections adjusted. Another generation of TIP is currently being developed at FAA's Technical Center.

Based on our findings, we do not anticipate further investigative action in this matter. If I can answer any questions or be of further assistance, please feel free to call me at (202) 366-1959, or my Deputy, Todd J. Zinser, at (202) 366-6767.

Sincerely,



Kenneth M. Mead
Inspector General