

October 21, 2004



# Acquisition

## Award of the Air Force F-15 Trainer Support Contract (D-2005-005)

**This special version of the report has been revised to omit source selection information or contractor proprietary data.**

Department of Defense  
Office of the Inspector General

*Quality*

*Integrity*

*Accountability*

### **Additional Copies**

To obtain additional copies of this report, visit the Web site of the Inspector General of the Department of Defense at <http://www.dodig.osd.mil/audit/reports> or contact the Secondary Reports Distribution Unit, Audit Followup and Technical Support at (703) 604-8937 (DSN 664-8937) or fax (703) 604-8932.

### **Suggestions for Future Audits**

To suggest ideas for or to request future audits, contact Audit Followup and Technical Support at (703) 604-8940 (DSN 664-8940) or fax (703) 604-8932. Ideas and requests can also be mailed to:

ODIG-AUD (ATTN: AFTS Audit Suggestions)  
Inspector General of the Department of Defense  
400 Army Navy Drive (Room 801)  
Arlington, VA 22202-4704

DEPARTMENT OF DEFENSE

**hotline**

**To report fraud, waste, mismanagement, and abuse of authority.**

Send written complaints to: Defense Hotline, The Pentagon, Washington, DC 20301-1900  
Phone: 800.424.9098 e-mail: [hotline@dodig.osd.mil](mailto:hotline@dodig.osd.mil) [www.dodig.osd.mil/hotline](http://www.dodig.osd.mil/hotline)

### **Acronyms**

ACA	Associated Contractor Agreements
ASC	Aeronautical Systems Center
OEM	Original Equipment Manufacturer
PPIRS	Past Performance Information Retrieval System
RFP	Request for Proposal
RPS	Rapid Prototyping System
SPO	System Program Office
TFE-21	Trainer Flight Equipment-21
TSA II	Training Systems Acquisition Two
TSSC	Training Systems Support Center



INSPECTOR GENERAL  
DEPARTMENT OF DEFENSE  
400 ARMY NAVY DRIVE  
ARLINGTON, VIRGINIA 22202-4704

October 21, 2004

MEMORANDUM FOR ASSISTANT SECRETARY OF THE AIR FORCE  
(FINANCIAL MANAGEMENT AND COMPTROLLER)

SUBJECT: Report on the Award of the Air Force F-15 Trainer Support Contract (Report No. D-2005-005)

We are providing this report for information and use. We conducted the audit in response to a request from Senator John McCain. We considered management comments on a draft of this report when preparing the final report. The Department of the Air Force comments conformed to the requirements of DoD Directive 7650.3 and left no unresolved issues. Therefore, no additional comments are required.

We appreciate the courtesies extended to the staff. Questions should be directed to Mr. Henry F. Kleinknecht at (703) 604-9324 (DSN 664-9324) or Mr. Patrick J. Nix at (703) 604-9332 (DSN 664-9332). The team members are listed inside the back cover. See Appendix C for the report distribution.

By direction of the Deputy Inspector General for Auditing:

David K. Steensma  
Assistant Inspector General  
for Contract Management

**Special Warning**

This special version of the report has been revised to omit source selection information or contractor proprietary data.

# Office of the Inspector General of the Department of Defense

Report No. D-2005-005

October 21, 2004

(Project No. D2003CH-0113)

## Award of the Air Force F-15 Trainer Support Contract

### Executive Summary

**Who Should Read This Report and Why?** Acquisition and contracting personnel within DoD and the Military Departments should read this report because it concerns the technical/management and past performance evaluations used to support a “best value” contract award (task order) by the Air Force for acquisition and support of the F-15 fighter-aircraft training device.

**Background.** This audit was performed in response to a request from Senator John McCain. The Senator asked that we review issues raised by a number of employees at L-3 Communications (the incumbent contractor). After the Air Force issued the contract to The Boeing Company for acquisition and support of the F-15 training device, L-3 Communications employees claimed that the contract was awarded based on an improper evaluation. Specifically, the employees claimed that the Air Force disregarded past performance and cost/price as selection criteria and, instead, issued the task order based solely on an inaccurate assessment of the technical/management criteria.

The Air Force Training Systems Management Directorate, Ogden Air Logistics Center, Hill Air Force Base, Utah, awarded task order QP01 under contract F33657-01-D-2074 to The Boeing Company on November 27, 2002, for F-15 training device acquisition and support. The “best value” award for a 10-year period was based on technical/management, past performance, and cost/price. Technical/management and past performance evaluations were considered equal in importance and, when combined, were significantly more important than cost/price. Based on an initial evaluation of five proposals, only The Boeing Company and L-3 Communications were considered for the “best value” decision. The total evaluated price was \$\* for The Boeing Company and \$\* for L-3 Communications. Because a Multiple Award Task Order contract was used, L-3 Communications cannot protest the decision of the Air Force to award the task order to The Boeing Company.

**Results.** The Air Force Training Systems Management Directorate integrated product team did not effectively conduct its technical/management evaluation and used a questionable methodology to evaluate past performance to support its decision to award The Boeing Company the task order contract for acquisition and support of the F-15 training device. Specifically, the Air Force Training Systems Management Directorate integrated product team:

- did not have detailed formal guidance on how to effectively use the Delphi technique (evaluation process) to develop technical/management criteria to minimize potential biases that could influence the criteria and incorrectly attempted to apply numeric values (scientific or knowledge-based measurements) to the technical/management subfactor evaluation criteria;

---

\* The dollar value has been deleted because it is source selection information or contractor proprietary data.

- did not achieve consensus on the technical evaluation results and failed to rank the offerors in association with the evaluation criteria;
- technical/management evaluators lowered the ratings for L-3 Communications for weaknesses that did not relate to the subfactor evaluation criteria; and
- the methodology used to assess the past performance of the offerors made past performance a nonfactor in the award decision because offerors with outstanding past performance and offerors with no prior performance received equal ratings.

As a result, the Air Force Order Award Authority did not have the most reliable information to support the “best value” decision to award The Boeing Company a 10-year task order for \$\* versus awarding the task order to L-3 Communications for \$\*, a difference of \$31.4 million. The Air Force Training Systems Management Directorate needs to establish procedures to effectively evaluate proposals from offerors and appropriately consider past performance, and the Air Force needs to determine whether the task order for F-15 trainer support should be recompleted before exercising the appropriate follow-on contract option year. See the Finding section for detailed recommendations.

**Management Comments and Audit Response.** The Associate Deputy Assistant Secretary of the Air Force for Contracting and the Commander, Ogden Air Logistics Center concurred with the finding and recommendations. The Commander agrees that the Center did not appropriately use the Delphi technique, however the Commander believed that the use of the technique was not necessary because the criteria and procedures discussed with industry during Industry Days were strictly adhered to and resulted in a fair evaluation process. The Commander stated that he had no plans to mandate the use of the Delphi technique for either establishing evaluation criteria or to achieve consensus among the evaluators in the future. However, the Commander stated that should his position on the use of the technique change, he would establish detailed formal procedures at that time and incorporate those procedures into the Training Systems Acquisition Two User’s Guide. Furthermore, the Commander agreed to modify the User’s Guide to require the use of a neutral rating for offerors without past performance. The Commander disagreed that ratings of L-3 Communications were unfairly lowered and that the Air Force Order Award Authority did not have reliable information to support the “best value” decision. However, the Commander agreed to determine before an extension of the order is approved that such an extension was the most advantageous method of fulfilling the Government’s need.

The planned corrective actions are acceptable. However, we believe the use of the Delphi technique is an excellent tool to ensure an “objective selection” process. We also stand behind our conclusions that the technical evaluators lowered the ratings assigned to L-3 Communications for weaknesses that did not relate to the evaluation criteria, and that the Air Force did not have the most reliable information to support the “best value” decision.

Additional management comments on the final report are not required. See the Finding section of the report for a discussion of the management comments on the finding and recommendations and our audit response, and the Management Comments section of the report for the complete text of the comments.

---

\* The dollar value has been deleted because it is source selection information or contractor proprietary data.

# **Table of Contents**

---

<b>Executive Summary</b>	i
<b>Background</b>	1
<b>Objective</b>	2
<b>Finding</b>	
F-15 Training Device Acquisition and Support	3
<b>Appendixes</b>	
A. Scope and Methodology	24
Prior Coverage	24
B. Evaluation of Weaknesses Identified during the Technical/Management Evaluation	25
C. Report Distribution	28
<b>Management Comments</b>	
Department of the Air Force	31

---

## Background

This audit was performed in response to a request from Senator John McCain. The Senator requested that we review issues raised by a number of employees at L-3 Communications (the incumbent contractor). After the Air Force Training Systems Management Directorate, Ogden Air Logistics Center, Hill Air Force Base, Utah (Training Systems Management Directorate) issued a contract (task order) to The Boeing Company (Boeing) for acquisition and support of the F-15 fighter-aircraft training devices, L-3 Communications employees claimed that the task order was awarded based on an improper evaluation. Specifically, the employees claimed that the Air Force disregarded past performance and cost/price as selection criteria and, instead, issued the task order based solely on an inaccurate assessment of the technical/management criteria.

**Initially Intended as a Sole-Source Contract to Boeing.** In response to an industry request dated February 7, 2002, the contracting officer for F-15 Systems Program Office (SPO) Wright-Patterson Air Force Base, Ohio, clarified the Air Force position to award a sole-source contract to Boeing, which was the original equipment manufacturer (OEM) for F-15 aircrew and maintenance trainers. The response to industry showed that a presolicitation notice was issued on June 6, 2001, to acquire the services that had “historically” been procured from the OEM to streamline the F-15 acquisition and sustainment process. However, after the F-15 SPO contracting officer became aware that the trainer support services were previously procured competitively, the sole-source acquisition effort was discontinued. The F-15 trainer requirement was then transferred to the Training Systems Management Directorate, who had awarded the previous contract to L-3 Communications for support of the F-15 training devices in June 1997.

**Training Systems Product Group.** The Air Force Training Systems Product Group team consists of the Training Systems Management Directorate (sustainment); the Air Force Training Systems Product Group, Wright-Patterson (acquisition); and the Air Force Research Laboratory, Mesa, Arizona (research). The team used the Training Systems Acquisition Two (TSA II) contract to procure the F-15 trainer support services. The TSA II contract is the second in a series of indefinite delivery/indefinite quantity type contracts that the team established to satisfy a variety of training systems requirements. The Training Systems Management Directorate awarded task order QP01 under contract F33657-01-D-2074 to Boeing on November 27, 2002, for F-15 training device acquisition and support. The “best value” award for a 10-year period was evaluated on technical/management, past performance, and cost/price where technical/management and past performance criteria were considered equal in importance and, when combined, were significantly more important than cost/price. Based on an initial evaluation of five proposals, only Boeing and L-3 Communications were considered for the best value decision. The total evaluated price was \$\* for Boeing and \$\* for L-3 Communications. Because a Multiple Award Task Order contract was used, L-3 Communications cannot protest the decision of the Ogden Training Directorate to award the task order to Boeing.

---

\* The dollar value has been deleted because it is source selection information or contractor proprietary data.

---

**Air Force Review.** On January 27, 2003, Senator McCain brought the concerns raised with the award to the attention of the Office of the Inspector General, Luke Air Force Base, Arizona. He requested that the Office of the Inspector General, Luke Air Force Base respond directly to his constituents. That request was forwarded to the Office of the Inspector General, Hill Air Force Base because the contract was awarded by the Contracting Directorate, Ogden Air Logistics Center, Hill Air Force Base. In March 2003, the Office of the Inspector General, Hill Air Force Base tasked the Contracting Directorate, which awarded the F-15 trainer support contract to Boeing, to prepare a response to the issues raised by L-3 Communications employees. The Contracting Directorate response restated the information that was previously conveyed to L-3 Communications during its post-award decision briefing. In March 2003, the Commander, Ogden Air Logistics Center also requested the Acquisition Center of Excellence<sup>1</sup> to determine whether the award was made in accordance with the Request for Proposal (RFP) criteria. The Acquisition Center of Excellence reported that the evaluation team followed the selection criteria contained in the RFP. The review results from the Center were not provided to Senator McCain.

A representative from the Air Force Office of the Inspector General, Headquarters, stated its organization was not the best suited to review the concerns raised with the award because its function was more designed to look into complaints of individual misconduct or mistreatment. Furthermore, the reviews performed by the Air Force Inspector General and the Acquisition Center of Excellence were limited and neither of the reviews identified the issues addressed in this report.

## Objective

The audit objective was to determine whether the Air Force decision to award a contract for F-15 trainer support was made in accordance with the technical/management, past performance, and cost/price criteria contained in the RFP and provides the best value to the Air Force. See Appendix A for a discussion of the scope and methodology related to the objective.

---

<sup>1</sup> The Acquisition Center of Excellence, Ogden Air Logistics Center provides advice on acquisition best practices to contracting offices during the solicitation development. The center was not used on the F-15 trainer support task order.

---

## **F-15 Training Device Acquisition and Support**

The Training Systems Management Directorate integrated product team did not effectively conduct the technical/management evaluation and used a questionable methodology to evaluate past performance to support its decision to award Boeing the task order contract for acquisition and support of the F-15 training device. Specifically, because of a lack of detailed formal guidance regarding conducting the technical/management evaluation, the Training Systems Management Directorate integrated product team:

- failed to effectively use the Delphi technique (evaluation process) to develop technical/management criteria and document how those procedures were used to minimize potential biases that may influence the criteria;
- failed to appropriately use the Delphi technique to evaluate proposals from offerors by applying numeric values “0” to “5” (scientific or knowledge-based measurements) to the subfactors and not achieving consensus or ranking the offerors in association with the evaluation criteria; and
- lowered ratings for L-3 Communications for weaknesses that did not relate to the subfactor evaluation criteria, lowered ratings for the same weakness under multiple subfactor criteria, and lowered ratings for weaknesses that had been resolved through evaluation notices.

Additionally, the methodology used by the Training Systems Management Directorate integrated product team to assess the past performance of the offerors effectively made past performance a nonfactor in the award decision because offerors with outstanding past performance and offerors with no prior performance received equal ratings. As a result, the Air Force Order Award Authority did not have the most reliable information to support the “best value” decision to award Boeing a 10-year task order for \$\* versus awarding the task order to L-3 Communications for \$\*, a difference of \$31.4 million.

## **Proposal Evaluation**

The TSA II User Guide, dated May 22, 2002, which was developed by the Training Systems Product Group, provides guidance for managing acquisitions under the TSA II contract from requirements definition through solicitation, order award, and execution of the order through final close out. The guide states task orders will be awarded based on an assessment of the program technical/

---

\* The dollar value has been deleted because it is source selection information or contractor proprietary data.

---

management approach, proposed total cost/price, past performance, and other factors determined appropriate to make an award decision.

**Supplemental Guidance Issued for F-15 Trainer Support Task Order.** In accordance with the guide, the Training Systems Management Directorate assembled an integrated product team to develop the RFP for the F-15 trainer support task order. The integrated product team operated from January through August 2002 and consisted of 35 individuals:

- 10 from the Ogden training directorate,
- 2 from the Ogden contracting directorate,
- 9 from the ASC training directorate (1 contractor),
- 5 from ASC F-15 System Program Offices (SPOs) (1 contractor), and
- 9 from the Air Combat Command.

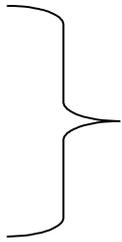
The team conversed by e-mail; held conference calls; and conducted industry days, briefings, and discussions with prospective offerors (using the team Web site and one-on-one meetings) to define the requirements and develop the criteria for evaluating the proposals. The team also developed “General Rules of Engagement” and “Rating Guidelines” to provide supplemental guidance for use by the team during the technical evaluation.

The “Evaluation Factors for Award” section of the RFP states that the offeror proposal judged on technical/management, past performance and cost/price to represent the “best value” to the Government would be awarded the order. The contracting officer for the task order stated that the criteria for the technical/management evaluation was developed by the integrated product team. The technical/management evaluation criteria identified six subfactors:

- Concurrency,
- Commonality,
- Baseline Requirements,
- Program Management and Staffing,
- Transitioning, and
- TFE-21 Conversion to Rapid Prototyping System (RPS).

The technical/management and past performance criteria were equal in importance and, when combined, were significantly more important than cost/price.

Table 1 shows an illustration, based on our interpretation of Air Force guidance, of how technical/management and past performance would constitute at least 30 percent each and cost/price no more than 40 percent.

<b>Table 1. Evaluation Factors and Illustration of Weighting</b>	
<b>Factor 1: Technical/Management Evaluation</b>	<b>Weighting</b>
Subfactor 1: Concurrency	
Subfactor 2: Commonality	
Subfactor 3: Baseline Requirements	
Subfactor 4: Program Management & Staffing	
Subfactor 5: Transitioning	
Subfactor 6: TFE-21 Conversion to RPS	
<b>Factor 2: Past Performance</b>	≥30%
<b>Factor 3: Cost/Price</b>	≤40%

**Technical/Management.** The integrated product team intended to use the Delphi technique to evaluate how well proposals from the offerors met the technical/management evaluation criteria. The Delphi technique was developed by the Rand Corporation and later enhanced by the U.S. Government as a group decision-making tool. The tool allows a group of experts to come to some consensus of opinion when the decisive factors were subjective, and not scientific or knowledge-based. The TSA II User Guide states, when the Delphi technique is used, that the offerors are ranked based on their capability to meet the evaluation criteria. The guide also includes an evaluation form that provides an example of how the ranking should be accomplished. However, the form does not show offerors being ranked, but does show offerors receiving numeric scores, including fractional numbers.

Based on that guidance, the integrated product team developed procedures to implement the Delphi technique and incorporated those procedures in the “Evaluation Factors for Award” section of the RFP. Those procedures state that each offeror would be assigned a numerical rating from “0” to “5” for each technical/management subfactor. The integrated product team also provided supplemental “General Rules of Engagement” and “Rating Guidelines” that authorized evaluators to assign offerors fractional numbers to the 0.5 level and provided broad guidelines designed to create consistency among evaluator ratings.

**Past Performance.** The TSA II User Guide restricts past performance reviews to the performance evaluations and contractor performance assessment reports for previous TSA II orders. However, the guide permits additional performance information to be obtained from the program managers or contracting officers for previous TSA II orders in absence of TSA II order performance evaluations or contractor performance assessment reports. The guide does not provide a methodology for how to assess the previous work performance of a contractor.

---

Based on that guidance, the integrated product team developed the procedures to evaluate past performance for this RFP. The “Evaluation Factors for Award” section of the RFP states that past performance would be limited to delivery orders awarded to the offerors within the TSA II contract. Specifically, the Government team would consider the relevancy the “past and present work performed under all TSA II delivery orders as well as the semi-annual performance reports and CPAR [contractor performance assessment report] ratings” of the offeror. The section further included the methodology that the integrated product team developed to assess the past performance of the offeror. The section states that the performance of an offeror would be rated as either “acceptable” or “not acceptable.” Performance of an offeror was considered “acceptable” unless the offeror received a performance rating of less than “green” (satisfactory) on a previous TSA II task order. Offerors with less than satisfactory performance on a previous task order would be found “not acceptable” and will not be considered for the award. Only those offerors that had received previous task orders under the TSA II contract were at risk of not being considered for the award because offerors lacking previous awards received “acceptable” ratings.

**Cost/Price.** The TSA II User Guide states that the total proposed cost/price of an offeror will be used to make an award decision. Based on that guidance, the integrated product team developed the procedures for evaluating the cost/price proposals received from the offerors. The RFP states that, for award purposes, a cost/price proposal was evaluated based on firm-fixed-price contract line items for program management, contractor logistics support, ramp up and transition of the Training Systems Support Center (TSSC) and Training Flight Equipment-21 (TFE-21) modification. Other contract line items were evaluated based on labor rates for specified hours relating to future modifications, maintenance calls, overtime, and surge. The integrated product team provided estimated quantities for all rates for normalization and rate comparison. The offeror provided quantities and labor categories for fixed-price level-of-effort contract line items for the software development center and TSSC functions.

## **Technical/Management - Delphi Technique to Develop Evaluation Criteria**

The integrated product team, lacking detailed formal guidance, failed to effectively use the Delphi technique (evaluation process) to develop technical/management criteria and document how those procedures were used to minimize potential biases that may influence the criteria. The Delphi technique recognizes the value of expert opinion, experience, and intuition and allows the use of limited information, when full scientific knowledge is not available, to achieve general consensus on an issue. The Delphi technique is particularly appropriate when decision-making is required in a political or emotional environment or when decisions affect strong factions with opposing preferences, as was the case with the F-15 trainer support source selection. For example, a representative from the Ogden training directorate stated that representatives from the F-15 SPO were impressed with the capabilities of the Boeing developed RPS and wanted the trainer effort awarded sole source to Boeing. Use of the Delphi technique, is designed to filter out any biases when developing evaluation criteria.

---

One study of the Delphi technique, conducted by the Carolla Development Corporation, produced a white paper discussing basic procedures for developing criteria. The procedures include:

- **Pick a facilitation leader.** Leader should be an expert in research data collection and not a stakeholder.
- **Select a panel of experts.** Panelists should have intimate knowledge of area and may be department managers, project leaders, and even stakeholders.
- **Identify a strawman criteria list from the panel.** In a brainstorming session, build a list of criteria that all think appropriate for the subject project, no “correct” criteria.
- **The panel ranks the criteria.** For each criterion, the panel ranks it as 1 (very important), 2 (somewhat important), and 3 (not important). Each panelist ranks the list individually, and anonymously if necessary.
- **Calculate the mean and deviation.** For each criterion in the list, find the mean value and remove all items with a mean greater than or equal to 2.0. Place criteria in rank order and discuss reasons for items with a high standard deviation. The panel may insert removed items back into the list after discussion.
- **Rerank the criteria.** Repeat the ranking process among the panelist until the results stabilize.

The integrated product team had no records of whether the Delphi technique was followed, who the facilitation leader or panel experts were, and, most importantly, how the evaluation criteria was developed and ranked by the group to achieve consensus. Failure to follow and document the appropriate Delphi technique procedures raises questions about whether biases influenced the selected evaluation criteria. Consequently, the shortcomings of the framework used by the integrated product team to develop the criteria may have determined the outcome.

**Questionable Subfactor Criteria.** Three of the six subfactors the integrated product team developed to determine how well a proposal submitted by an offeror satisfied the technical requirements were met, in part, through the accomplishment of acts that required Boeing involvement.

**Subfactor 1: Concurrency.** The concurrency criteria, which represented 20 percent of the total available technical/management rating points, was used to assess the processes proposed by an offeror for identifying, acquiring, and fielding the data updates needed to keep the trainers concurrent (identical) to the aircraft. The concurrency criteria emphasized each offeror securing advance Associated Contractor Agreements (ACAs) with Boeing and its subcontractors for the needed data and linked the technical score of an offeror to the accomplishment of acts controlled by a competitor. The integrated product team required offerors to submit supporting documentation to convey the maturity or level of progress of

---

the agreements and to communicate an understanding of the types of information exchange needed. The integrated product team used the information received to assess the viability and risk associated with the approach of an offeror.

L-3 Communications had its rating lowered for this criteria because L-3 Communications had not finalized its ACA with Boeing when its proposal was reviewed. Also, a majority of the Boeing subcontractors refused to enter into ACAs with L-3 Communications until a contract was awarded and a firm requirement existed. A number of the technical evaluators considered the progress that L-3 Communications made to establish ACAs with Boeing and its subcontractors a weakness and lowered ratings of the contractor accordingly.

**Subfactor 2: Commonality.** The commonality criteria, which represented 13 percent of the total available technical/management rating points, was used to assess the plan of an offeror for achieving the Air Force objective to “decrease long-term training systems costs, maximize use of common software and develop common system architectures to the maximum extent possible.” The commonality criteria emphasized that each offeror incorporate common core software from the RPS system (Boeing developed software) in modifications and design of new systems. The plans of the offerors were required to explain how RPS software would be used to accomplish that objective and provide a roadmap for modifying or replacing trainers that reduced the duplication of effort and facilitated common hardware/software use for all F-15 trainers.

However, the cost reduction objective could be accomplished without incorporating RPS software in the training devices. Offerors could use any method, not limited to using the Boeing developed RPS software, to decrease long-term training systems costs. Although the integrated product team created an RPS data library in an attempt to put all offerors on a level playing field, the information provided may not have been sufficient to allow offerors (other than Boeing) to effectively bid to the requirement. The technical scores given by the integrated product team to all other offerors for the “Commonality” subfactor, are significantly lower than that given to Boeing. See Figure 1. “Technical/Management Evaluation Scores of the Offerors.”

A representative from L-3 Communications stated that his company and other offerors requested access to the RPS data library at the Industry Day conferences. However, the integrated product team initially took the position that because the system contained a mix of releasable and Boeing proprietary information in a classified system that no access would be given. The L-3 Communications representative stated that only after L-3 Communications and other offerors expressed concern that a valid technical approach could not be achieved without access to the data, did the integrated product team permit the offerors to review the RPS library. A member of the integrated product team stated the library consisted of three documents (three volumes each) and was available for review over a six-week period. However, according to the L-3 Communications representative, the data contained in the library could not be reproduced, and only included samples of the software code. The L-3 Communications representative further stated that the data available was for an outdated version of RPS. Therefore, according to the L-3 Communications representative, it was not

---

possible to determine the overall executive structure or the system level design of RPS.

**Subfactor 6: TFE-21 Conversion to RPS.** The TFE-21 conversion to RPS criteria, which represented 22 percent of the total available technical/management rating points was also used to assess the approach of an offeror for incorporating the RPS software (Boeing developed software) into the trainers. The criteria required an evaluation of the approach of the offeror for incorporating the RPS software into the TFE-21 maintenance training devices. Furthermore, through a memorandum, dated August 19, 2002, the integrated product team added the additional requirement to include a price for an ACA with Boeing to accomplish the conversion task. The requirements of the subfactor were met when an offeror proposed a viable modification program that was likely to lead to maintainable training devices that met or exceeded TFE-21 system performance characteristics and when the offeror identified how it would use existing RPS data and processes to minimize the time to field future operational flight plan changes in the devices. According to an L-3 Communications representative, it was very difficult to articulate to Boeing what services and data were needed to accomplish this conversion task and complete an ACA due to the RPS system architecture being largely unknown.

Although we were unable to determine which member or members of the integrated product team were responsible for the significant changes to the draft evaluation criteria, the final criteria were changed to add the “TFE-21 Conversion to RPS” subfactor and to make it the highest weighted evaluation criteria. Keeping the trainers identical to the aircraft is the critical task affecting warfighter capability. However, the task could be accomplished without incorporating RPS software into the training devices. Furthermore, incorporating RPS into the trainers may not be the best way to reduce long-term training costs. Although the integrated product team created an RPS data library in an attempt to put all offerors on a level playing field, the information provided may not have been sufficient to allow offerors (other than Boeing) to effectively bid to the requirement. The technical scores given by the integrated product team to all other offerors for the “TFE-21 Conversion to RPS” subfactor are significantly lower than that given to Boeing. See Figure 1.

**Procedures Needed.** The Ogden Training Directorate needs to establish detailed formal procedures to effectively use the Delphi technique to develop technical/management criteria and document how those procedures were used in future evaluations to minimize potential biases that may influence the criteria.

## **Technical/Management - Delphi Technique to Evaluate the Proposals**

The Ogden Training Directorate, lacking detailed formal guidance, failed to appropriately use the Delphi technique to evaluate proposals from offerors. The Ogden Training Directorate used 14 members of the integrated product team (6 from the Ogden Training Directorate; 4 from the Air Combat Command; 2 from the ASC F-15 SPOs, 1 contractor; and 2 from the ASC training directorate, 1 contractor) to evaluate the technical/management proposals from the offerors.

---

Each evaluator was generally assigned to three of the six technical/management subfactor teams and tasked with determining how well the proposals met the technical criteria. The subfactor teams reviewed proposals in October 2002 and operated for the most part from Ogden Air Logistics Center and Wright-Patterson Air Force Base. The evaluators reviewed the proposals from the offerors, identified the strengths (the approach proposed by an offeror exceeded or positively impacted a technical requirement) and weaknesses (the approach proposed by an offeror failed to meet or adversely impacted a technical requirement) of each proposal. Additionally, instead of ranking proposals (Delphi technique), each evaluator assigned the offerors numerical values (scientific or knowledge-based measurements) based on the level of compliance that the evaluator perceived a proposal from an offeror had to the subfactor criteria. Furthermore, the subfactor teams did not achieve consensus on the numerical values assigned to each offeror or rank offerors in association with the evaluation criteria.

**Numerical Ratings.** The technical evaluators attempted to use numerical ratings (scientific or knowledge-based measurements) for the technical/management subfactors to identify which proposal best satisfied the technical requirements. The integrated product team developed criteria for its technical/management evaluation and classified those criteria into six subfactors: “Concurrency,” “Commonality,” “Baseline Requirements,” “Program Management and Staffing,” “Transitioning,” and “TFE-21 Conversion to RPS.” The TSA II User Guide stated each contractor’s capability to meet the subfactor criteria should be ranked. However, the guide contained only general guidance on how that should be accomplished. The TSA II User Guide included a sample evaluation form that showed contractors being scored in 0.5 increments. Interpreting that guidance, the integrated product team instructed the technical evaluators to assign the proposals submitted by the offerors numerical ratings from 0 to 5, in 0.5 increments, to represent how well the proposals met the subfactor criteria.

Figure 1 shows the technical/management evaluation scores given to each offeror.

Criteria	EC <sup>1</sup>	SWF <sup>2</sup>	*	L-3	Boeing	*	*
Subfactor 1: Concurrency	M	9 (20%)		2.3 20.7	3.8 34.2		
Subfactor 2: Commonality	M	6 (13%)		2.3 13.8	3.9 23.4		
Subfactor 3: Baseline Requirements	M	6 (13%)		3.0 18.0	3.7 22.2		
Subfactor 4: Program Management & Staffing	M	7 (16%)		2.0 14.0	2.7 18.9		
Subfactor 5: Transitioning	M	7 (16%)		2.8 19.6	1.4 9.8		
Subfactor 6: TFE-21 Conversion to RPS	M	10 (22%)		2.9 29.0	3.9 39.0		
<b>Total</b>				<b>115.1</b>	<b>147.5</b>		
<sup>1</sup> The Evaluation Class (EC) designates whether the criteria is Mandatory (M) or Desirable (D) to supporting the F-15 Trainers. <sup>2</sup> The System Weight Factors (SWFs) were used to convert the offeror's ratings (upper left corner) into values (lower right corner) representative of the importance the Air Force placed on the accomplishment of the area evaluated through each subfactor. The percentage shown represents the significance each subfactor has on an offeror's overall technical/management score. * Data omitted because it is source selection information or contractor proprietary data.							

**Figure 1. Technical/Management Evaluation Scores of the Offerors**

Based on their evaluation, the integrated product team concluded that the proposal submitted by Boeing was “the overall technically superior proposal in fulfilling the Government’s requirements.” Boeing received a score of 147.5 out of the 225 possible rating points, while L-3 Communication received a score of 115.1 points.

However, due to the technical complexity of the solicitation requirements and the lack of scientific or knowledge-based measurements, it was not possible to effectively differentiate between the 11 rating increments<sup>2</sup> and achieve consensus on that rating. The Delphi technique is a tool that a group of experts can use to come to some consensus of opinion when the decisive factors are subjective. The best that can be expected from the Delphi technique is for a group of experts to come to agreement on whether the proposal of an offeror meets the subfactor criteria or not, or whether the approach proposed by an offeror is better than that of a competitor. If it were possible to accurately identify which of the 11 equal intervals represents the level of compliance of an offeror with the stated criteria, the integrated product team would not need to use the Delphi technique.

<sup>2</sup> The technical evaluators were tasked to assign the offerors one of 11 possible rating increments (0, 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5) to represent how well their proposal met subfactor requirements.

**Achieving Consensus.** The scientific or knowledge-based ratings used by the integrated product team made it impossible to achieve consensus on the technical/management subfactor scores. Technical evaluators stated that their teams discussed the weaknesses and strengths noted with each proposal submitted by the offerors, but were unable to agree on a common set of strengths and weaknesses or which of the 11 possible ratings should be used to represent how well the proposals from the offerors met the subfactor requirements. Table 2 shows that individual evaluator ratings were averaged to determine the team summary for four of the six subfactors for both L-3 Communications and Boeing.

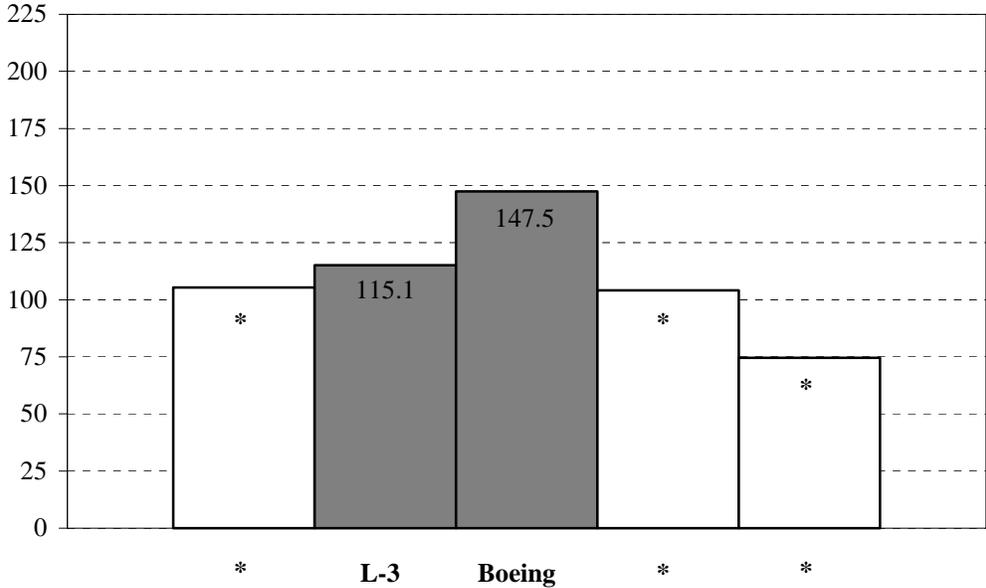
<b>Table 2. Evaluation Ratings of the Offerors</b>							
	Evaluator						Team
<u>L-3 Communications</u>	<u>A</u>	<u>B</u>	<u>C</u>	<u>D</u>	<u>E</u>	<u>F</u>	<u>Summary</u>
Concurrency	3.0	3.0	2.0	2.0	2.0	2.0	2.3 <sup>1</sup>
Commonality	2.5	2.0	3.0	2.0	2.0	2.0	2.3 <sup>1</sup>
Baseline Requirements	3.0	3.0	3.0	3.5	2.5	3.0	3.0 <sup>1</sup>
PM & Staffing	1.5	2.5	2.0	2.5	2.8 <sup>2</sup>	2.5	2.0
Transitioning	4.0	3.0	3.0	3.0	2.5	1.5	2.8 <sup>1</sup>
TFE-21 Conversion	3.0	3.0	3.0	2.0	3.0	3.0	2.9
	Evaluator						Team
<u>Boeing</u>	<u>A</u>	<u>B</u>	<u>C</u>	<u>D</u>	<u>E</u>	<u>F</u>	<u>Summary</u>
Concurrency	3.5	3.5	4.0	4.0	4.5	3.5	3.8 <sup>1</sup>
Commonality	3.5	3.5	4.0	4.0	4.5	4.0	3.9 <sup>1</sup>
Baseline Requirements	3.5	3.0	3.5	3.5	4.5	4.0	3.7 <sup>1</sup>
PM & Staffing	2.8 <sup>2</sup>	2.5	2.5	3.0	2.5	3.0	2.7 <sup>1</sup>
Transitioning	2.0	1.5	1.5	1.5	1.5	1.0	1.4
TFE-21 Conversion	4.0	3.5	4.0	4.0	4.5	4.0	3.9
<sup>1</sup> Denotes team summary was derived by average.							
<sup>2</sup> Evaluator failed to follow guidance for establishing a rating in 0.5 increments.							

Differences in evaluators experience levels and their prior experiences with the offerors impacted the ability of the teams to achieve consensus on ratings. For example, one evaluator was new and lacked experience evaluating the sufficiency of proposals. At the time of his assignment, he informed the evaluation team chief that he felt uncomfortable with his ratings affecting the technical ratings given to offerors. Although the evaluator felt he possessed the technical knowledge necessary to evaluate proposals, the evaluator stated he was more comfortable identifying the strengths as opposed to the weaknesses of the proposals submitted by the offerors. This was reflected in his ratings, none of the weaknesses he noted with the proposal from L-3 Communications related to the respective subfactor criteria.

Another evaluator had a bad experience with L-3 Communications on a prior contract. According to the evaluator, an L-3 Communications employee refused to meet a contract deliverable. The employee’s position did not sit well with the evaluator and caused the evaluator to inquire about how fast the Air Force could replace L-3 Communications. Based on our assessment of the validity of the weaknesses the evaluator noted with the L-3 Communications proposal, it is clear that this encounter biased the evaluator’s assessment of the proposal submitted by L-3 Communications. Only 2 of the 9 weaknesses (22 percent) that the evaluator noted related to the respective subfactor criteria.

The impact that inexperience and personal biases had on the validity of the weaknesses noted and the inability of evaluators to accurately place offerors in one of the 11 possible numerical rating intervals caused the subfactor evaluation groups to be unable to achieve consensus. These types of problems are exactly what the Delphi technique is designed to mitigate and avoid.

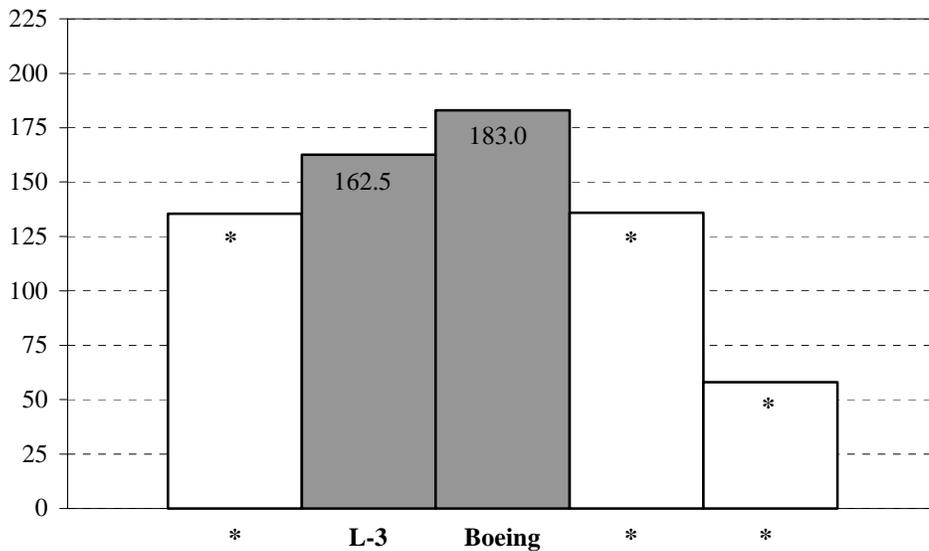
**Ranking Offerors in Association With Evaluation Criteria.** Using its “scientific or knowledge-based” approach, the integrated product team failed to effectively interpret the overall technical/management scores of the offerors. Boeing was determined technically superior because its proposal received the highest technical score. Figure 2 shows that Boeing received 147.5 of the 225 possible rating points while L-3 Communications received 115.1 points.



\* Data omitted because it is source selection information or contractor proprietary data.

**Figure 2. Air Force Technical/Management Ratings**

To demonstrate the effects of ranking the technical/management evaluations, we converted the subfactor scores into numerical rankings. Specifically, we substituted the technical scores each offeror received from the subfactor groups with values from “1” to “5.” We assigned a value of “5” to the offeror the integrated product team rated the highest, a value of “4” to the offeror the integrated product team rated the next highest, and so forth. For example, Boeing, the offeror that rated the highest for the “Concurrency” subfactor, had its “Concurrency” subfactor score replaced with a numerical ranking of 5. The rankings were then multiplied by the corresponding system weight factor and the results summed to derive the total effective rankings of the offerors. Although the results were impacted by the shortcomings of the technical/management criteria and proposal evaluation, Figure 3 shows the difference in proposals from the offerors when ranking is used.



\* Data omitted because it is source selection information or contractor proprietary data.

**Figure 3. Illustration of Technical/Management Rankings**

The Delphi technique is not designed to provide more information than that provided in Figure 3. If the Ogden Training Directorate wants scientific technical evaluations, the Ogden Training Directorate needs to develop more scientific measurements to distinguish between different rating scores because the Delphi technique is not appropriate.

**Procedures Needed.** The Ogden Training Directorate needs to establish detailed formal procedures that appropriately use the Delphi technique to evaluate proposals of offerors to achieve consensus and rank the offerors in association with the evaluation criteria.

## Technical/Management – Weaknesses Noted

Technical/Management ratings for L-3 Communications were lowered for weaknesses that did not relate to the subfactor evaluation criteria, were lowered for the same weakness under multiple subfactor criteria, and were lowered for weaknesses that had been resolved through evaluation notices. During the audit we discussed the weaknesses and strengths we believed were invalid with the technical evaluators. For the most part, we found the strengths evaluators noted with the proposals for L-3 Communications and Boeing were consistent with the evaluation criteria. The only major exception noted, was that some evaluators increased the ratings given to Boeing for the “Transitioning” and “Concurrency” subfactors for proposing to relocate the TSSC to the F-15E Weapon Systems Training Simulation Facility on Seymour Johnson Air Force Base, North Carolina, despite knowing when they were evaluating proposals that the facility was unavailable.

However, as Table 3 shows, based on our interpretation of the published criteria, only 19 (39 percent) of the 49 weaknesses noted for L-3 Communications were valid while 21 (91 percent) of 23 of the weaknesses identified for Boeing were appropriate. We found that L-3 Communications and Boeing basically had the same number of valid weaknesses identified by the evaluators, 19 and 21 respectively. See Appendix B for a detailed list of the weaknesses we found inappropriate and reason.

<u>Subfactor</u>	<u>L-3 Communications</u>			<u>Boeing</u>		
	<u>Yes</u>	<u>No</u>	<u>Percent Valid</u>	<u>Yes</u>	<u>No</u>	<u>Percent Valid</u>
Concurrency	4	11	27	2	0	100
Commonality	7	7	50	2	2	50
Baseline Requirements	2	2	50	3	0	100
PM & Staffing	2	4	33	5	0	100
Transitioning	1	4	20	9	0	100
TFE-21 Conversion to RPS	<u>3</u>	<u>2</u>	<u>60</u>	<u>0</u>	<u>0</u>	-
<b>Total</b>	<b>19</b>	<b>30</b>	<b>39</b>	<b>21</b>	<b>2</b>	<b>91</b>

**Weakness Not Related to Criteria.** Evaluators downgraded the ratings given to L-3 Communications and Boeing for weaknesses unrelated to the criteria identified for the respective subfactor. For example, one evaluator downgraded the rating given to L-3 Communications for the “Concurrency” subfactor because he found that its plan to update only the software of TFE-21 trainers placed the approach proposed by the offeror for the TFE-21 conversion at high risk. The weakness does not relate to the criteria that the integrated product team stated

---

would be used to evaluate how well a proposal from a offeror met the “Concurrency” subfactor requirements. In the F-15 trainer support RFP, the integrated product team stated the “Concurrency” subfactor was met when the proposal described how it met or exceeded the concurrency requirements. Specifically, for both avionic and maintenance training devices:

- (1) How well will the offeror’s approaches to obtaining data from the aircraft prime contractor and/or other contractors facilitate meeting the 60-day concurrency window? The proposal will be evaluated for a clear description of how the offeror determines when an ACA is needed, what the content of the ACA should be, and how it establishes ACAs. The offeror will be evaluated on their demonstrated understanding of the essential need to have an ACA established with the prime aircraft contractor and other contractors to meet the requirements of the solicitation. To assess risk, the offeror will be evaluated on the maturity of documentation supporting its ACAs (actual, unexecuted ACA; letter of intent, etc.). The proposal will be evaluated for type and form of data, delivery schedule of data, and additional support and/or tools required to use data. The offeror’s approach will be evaluated to determine viability and mitigation of risk.
- (2) How well does the offeror integrate this data into an efficient design process?
- (3) How well does the offeror’s overall approach to designing, developing and fielding updates facilitate meeting the 60-day concurrency window?
- (4) How well does the offeror’s proposal communicate how the offeror will meet concurrency requirements for multiple WST OFP configurations in the field simultaneously, including the potential requirement to periodically switch WST configurations back and forth (e.g., Suite 4E+ to Suite 5 and back)?

None of the criteria state that the plan of the offeror for updating the TFE-21 trainers would impact its “Concurrency” subfactor rating. Furthermore, in the post award brief, the integrated product team identified that weakness as a deficiency which adversely impacted the rating given to L-3 Communications for the “TFE-21 Conversion to RPS” subfactor.

Other evaluators lowered the ratings given Boeing and L-3 Communications for issues related to the “Cost/Price” factor. Two evaluators lowered the ratings given to Boeing for the “Commonality” subfactor because they believed the labor rates Boeing proposed were high. Another evaluator lowered the rating he gave L-3 Communications for the “Commonality” subfactor because he believed hours proposed by L-3 Communications were inflated to compensate for the low labor rate that it proposed. Furthermore, another evaluator lowered the rating given to L-3 Communications for “Commonality” subfactor because the evaluator believed the hardware replacement and conversion upgrades it proposed were costly. These issues do not relate to the technical/management rating criteria for the subfactors.

---

**Same Weaknesses Under Multiple Criteria.** The technical/management evaluation ratings for L-3 Communications were lowered for the same weaknesses under multiple subfactor criteria. For example, one evaluator (assigned to rate proposals against the “Concurrency,” “Commonality,” and “Baseline Requirements” subfactor requirements) downgraded the ratings given to L-3 Communications for all three subfactors for not taking steps to enter into ACAs with the contractors necessary to support the trainers in a timely manner. Only the “Concurrency” subfactor identifies the ACAs of an offeror as one of its evaluation criteria.

In another example, a number of evaluators collectively lowered the ratings they gave L-3 Communications for the “Concurrency,” “Baseline Requirements,” “Program Management and Staffing,” and “Transitioning” subfactors for proposing to move the TSSC just off of Luke Air Force Base, Arizona, when the current facilities became no longer available. The RFP states that the plan of an offeror for relocating the assets and functions of the TSSC will be evaluated through the “Transitioning” subfactor. We find it improper to lower the ratings given to an offeror for multiple subfactors for the same perceived weakness.

**Weaknesses Resolved Through Evaluation Notices.** The technical/management evaluation ratings of L-3 Communications were reduced for weaknesses resolved through evaluation notices. For example, evaluators assigned to rate proposals against the “Program Management and Staffing” subfactor requirements, downgraded the ratings given to L-3 Communications because their total TSSC staffing was management top heavy and the staffing at Seymour Johnson Air Force Base, North Carolina, was in question. The integrated product team sent L-3 Communications evaluation notices, providing them an opportunity to clarify what was stated in their proposal.

In its response, L-3 Communications stated that each TSSC manager was qualified to perform tasks in their respective areas of specialization and routinely performed engineering and support tasks in addition to their management functions. Furthermore, if the TSSC workload occasionally fluctuated to a lower level, L-3 Communications would reduce its noncore workforce accordingly. This approach ensured retention of key management and technical staff and provided the greatest flexibility for the program. In addition, a reduced technician staffing level at Seymour Johnson Air Force Base should not increase risk. It appears that responses from L-3 Communications resolved the concerns of the integrated product team.

Had the Delphi technique been properly used, the process would have filtered out the problems with specific technical evaluators. Therefore, we are not making a recommendation relating to this area.

## **Past Performance - All Offerors Equal**

The methodology used to assess past performance effectively made it a nonfactor in the award decision because offerors with outstanding past performance and no prior performance received equal ratings. In accordance with the TSA II User Guide, the integrated product team identified which offerors had been awarded

orders in the past under the TSA II contract. The team contacted the contracting officers for those orders and identified the completed performance ratings that each contractor had received. Using that information, the team concluded the past performance of the three offerors that had previously received orders was “acceptable.” Under the integrated product team rating method, the team also found the past performance of the two offerors without experience was also “acceptable.” However, the integrated product team methodology for rating past performance gives offerors without previous performance an advantage, because no credit was given for above average performance. Therefore, previous performance could only adversely impact the ability of an offeror to receive the order. Where there is a risk for poor performance, there should also be credit given for above average performance.

Using performance data from the Past Performance Information Retrieval System (PPIRS) relating to the orders awarded prior to the F-15 trainer support task order under the TSA II contracts, we determined that the integrated product team could have developed a methodology to credit those contractors with above average past performance. Table 4 shows all three offerors that received orders previously had performed above average and provides an illustration of how past performance could be rated.

<b>Table 4. Illustration of Past Performance Ratings for TSA II Contracts</b>				
Contractor	Order	Relevancy	PPIRS Rating for Quality of Product/Service <sup>†</sup>	Proposed Rating
*	*	*	*	*
L-3 Communications	1	High	Very Good	4.5
L-3 Communications	2	Moderate	Very Good	4.0
L-3 Communications	3	Moderate	Satisfactory Plus	3.5
*	*	*	*	*
*	*	*	*	*

<sup>†</sup> Offerors received satisfactory or higher on all the other areas assessed.  
 \* Data omitted because it is source selection information or contractor proprietary data.

The illustrated ratings were derived using a five-point rating scale we developed based on the relevancy of the previous performance an offerors to the requirements of the solicitation and the level of performance of the offerors. For example, an offeror whose work was highly relevant to the requirements of the solicitation and performed very good in the past would be rated a “4.5,” whereas an offeror whose work was only moderately relevant and performed very well would have received only a “4.0” rating. The Ogden Training Directorate should develop a past performance rating system that credits offerors with above average performance and does not provide an advantage to offerors without previous performance.

## Cost/Price - Evaluation

Three members of the integrated product team (two from the Ogden Contracting Directorate and one from the Ogden Training Directorate) evaluated the total cost/price proposed by the offerors. The source selection documentation shows that the cost/price proposals from the offerors were evaluated in accordance with the TSA II User Guide and the RFP. Table 5 shows the total cost/price proposed by L-3 Communications, was approximately \$31.4 million less than that proposed by Boeing (\$\* versus \$\*, respectively).

<u>Cost Categories</u>	<u>Boeing</u>	<u>L-3 Communications</u>
Trainer Support	\$ *	\$ *
Trainer Flight Equipment-21 Modification	*	*
Fixed Price Level of Effort Contract Line Item Numbers	*	*
Fixed Price for Future Modifications	*	*
Cost Reimbursable for Future Modifications	*	*
Fixed Price for Overtime, Surge, & Maintenance Calls	*	*
Cost Reimbursable No Fee Contract Line Item Numbers	<u>*</u>	<u>*</u>
<b>Total Price</b>	<b>\$ *</b>	<b>\$ *</b>

The F-15 trainer order award briefing dated November 20, 2002, also shows a “Quantified Risk Assessment” cumulative cost comparison between L-3 Communications and Boeing. The cost comparison, based on cost numbers taken from the offerors’ proposals that were applied to the approach suggested by the offerors, showed the total Boeing cost at \$\* versus \$\* for L-3 Communications. However, the differences in the suggested approaches of the offerors (necessary items versus desirable items) make the comparison of little value. Furthermore, both the chairman of the selection team and the order award authority stated this cost comparison that was not done in accordance with the RFP was not a factor in the decision to award the task order to Boeing.

## Conclusion

Based upon the evaluation of technical/management and past performance factors, the Training Systems Management Directorate concluded the award should be made to Boeing. The Air Force Order Award Authority stated that the technical superiority, quality, and risk associated with the approach proposed by Boeing provided the best overall value to the Government despite its \$31.4 million additional cost. However, the integrated product team did not appropriately use

\* The dollar value has been deleted because it is source selection information or contractor proprietary data.

---

the Delphi technique and there were problems affecting the subfactor ratings given to L-3 Communications. Furthermore, the sufficiency of the information provided to differentiate the levels of performance, the lack of consensus reviewing proposals from the offerors, and the interpretation of the overall technical/management scores of the offerors raises concern. In addition, emphasis put on incorporating software developed by Boeing into the trainers and the requirement to obtain an agreement from Boeing in advance of contract award to obtain the data needed to support the trainers favors Boeing. Based on the results of this audit, the Training Systems Management Directorate needs to determine whether the task order for F-15 trainer support should be recompeted before exercising the appropriate follow-on contract option year.

## **Management Comments on the Finding and Audit Response**

**Delphi Technique for Developing Evaluation Criteria.** The Commander, Ogden Air Logistics Center stated they did not use the Delphi technique to develop the evaluation factors and subfactors. Instead, the Center used an equally valid methodology, which employs consideration of the risks involved in the acquisition so that critical factors and subfactors may be developed accordingly.

**Audit Response.** The Delphi technique, if used correctly, is an excellent tool to ensure an “objective selection” process in developing evaluation criteria, particularly when the decisions affect strong factions with opposing preferences, as was the case with the F-15 trainer support source selection. The use of the Delphi technique would have eliminated any questions related to the selection of evaluation criteria.

### **Concerns Raised By Prospective Offerors During the Solicitation**

**Development.** The Commander stated that the evaluation factors for award were issued in draft form to allow prospective offerors the opportunity to respond with any concerns regarding the approach. The Center did not receive any indication during that phase that any offerors had concerns with the evaluation approach.

**Audit Response.** Several offerors and members of the integrated product team took issue with the evaluation factors and questioned the emphasis put on incorporating software developed by Boeing into the trainers and the requirement to obtain an agreement from Boeing in advance of contract award at the time the Air Force was developing the solicitation.

**Risk Assessment to Develop Evaluation Criteria.** The Commander stated they used risk assessment to determine the risk areas of planned acquisition and developed the evaluation criteria around the risks identified and the technical requirements that were most likely to be discriminators during the evaluation.

**Audit Response.** There was no documentation to support any risk assessment used by the Air Force to develop the evaluation criteria, and we were unable to determine which member or members of the integrated product team actually developed the criteria.

---

**Results of Bid Protest Review.** The Commander commented that he believed the Government Accountability Office would have supported the source selection decision had it been protested. Accordingly, he recommended that we seek a peer review of the audit by the bid protest section of the Government Accountability Office.

**Audit Response.** Because the Ogden Air Logistics Center used a Multiple Award Task Order contract for acquisition and support of the F-15 trainer devices, L-3 Communications was not permitted to protest the decision to award the task order to The Boeing Company. Therefore, the Government Accountability Office protest section would not review the award and there is no reason to request such a review.

**Deviation from Air Force Guidance for Awarding Orders Under Multiple Award Contracts.** The Commander stated that although the evaluation technique and modifications to the Delphi Technique in its limited use in the source selection process could be improved upon, the audit report produced no findings or recommendations that demonstrate that Air Force guidance for awarding orders under multiple award contracts was not followed nor that there was any arbitrariness in the way ratings were assigned.

**Audit Response.** Differences in experience levels of the technical evaluators, experiences evaluators previously had with the offerors, as well as variations in team and individual evaluator rating methodologies impacted the ability of the teams to achieve agreement on proposal strengths and weaknesses. As a result, differences existed in the way ratings were assigned.

**Validity of Weaknesses Noted with L-3 Communications.** The Commander stated that the ratings assigned to L-3 Communications were not unfairly lowered for weaknesses that did not relate to subfactor evaluation criteria, for the same weakness under multiple subfactor criteria, or for weaknesses that had been resolved through evaluation notices. The Commander took exception with the audit conclusions about the validity of a number of the weaknesses identified in the report.

**Audit Response.** We thoroughly discussed the weaknesses that were invalid with the technical evaluators. As a result, we concluded that only 19 (39 percent) of the 49 weaknesses were valid. Most of the evaluators that we talked to agreed with our conclusions.

**Past Performance Methodology.** The Commander agreed that the methodology used to evaluate past performance did not allow discrimination among offerors with acceptable past performance. The Commander stated the evaluation methodology was structured to recognize bad past performance. In retrospect, the Commander realizes that allowing more granularity in the past performance ratings would achieve a more precise means of differentiating among the offerors in the past performance area.

**Audit Response.** Adopting a methodology that offers a greater spectrum of past performance ratings would provide a better means to differentiate among offerors in the past performance area.

---

**Information to Support the “Best Value” Decision.** The Commander stated that the Air Force Order Award Authority did have reliable information to support the “best value” decision. The Commander also stated that evaluation criteria were followed, that offerors provided sufficient information in their proposals to conduct evaluations and assign proposal ratings, and that the Order Award Authority was given sufficient and reliable information regarding those ratings to support a “best value” decision.

**Audit Response.** The Ogden Air Logistics Center could have properly used the Delphi technique to preclude any objectivity issues related to developing the evaluation criteria and technical evaluations. Unfortunately, the source selection evaluation factors used by the Air Force raised questions about whether an “objective selection” process was followed.

## **Recommendations and Management Comments**

**We recommend that the Commander, Ogden Air Logistics Center in conjunction with the Commander, Training Systems Product Group, Wright-Patterson, Air Force Base:**

**1. Establish detailed formal procedures to effectively use the Delphi technique to develop technical/management criteria and document how those procedures were used in future evaluations.**

**2. Establish detailed formal procedures that appropriately use the Delphi technique to evaluate proposals from offerors to achieve consensus and rank offerors in association with the evaluation criteria.**

**Ogden Air Logistics Center Comments.** The Commander, Ogden Air Logistics Center concurred and stated that there were no plans to mandate the use of the Delphi technique for establishing criteria or to achieve consensus among the evaluators in the future. However, should the Center change its position in the future, it would establish and incorporate into the Training Systems Acquisition Two User’s Guide detailed formal procedures on how the Delphi technique would be followed. The Commander also agreed to document the evaluation results accordingly.

**3. Either develop an effective past performance rating system for future competitive Training Systems Acquisitions Two orders or eliminate past performance as an evaluation criteria.**

**Ogden Air Logistics Center Comments.** The Commander concurred and agreed to amend the Training Systems Acquisition Two User’s Guide to require use of a neutral rating for offerors without past performance records when past performance is assessed.

---

**4. Determine whether further action on the F-15 trainer support task order is appropriate before extending the contract.**

**Ogden Air Logistics Center Comments.** The Commander concurred stating that before an order extension is approved, he will ensure that an order extension is the most advantageous method of fulfilling the need of the Government.

---

## Appendix A. Scope and Methodology

We reviewed the procedures and documentation used to support the Air Force decision to award the F-15 trainer support contract to Boeing. The award evaluation was performed at Hill Air Force Base, Utah, and Wright-Patterson Air Force Base, Ohio. The dates of the documentation reviewed ranged from June 1997, the date of the previous F-15 trainer support contract award, to March 2003, the date of the Acquisition Center of Excellence review. We reviewed the technical/management, past performance, and cost/price criteria contained in the RFP. We also reviewed applicable guidance related to fair opportunity requirements for orders exceeding \$2,500 issued under multiple task order contracts. Additionally, we interviewed technical/management evaluators from Luke Air Force Base, Ogden Air Logistics Center, and Wright-Patterson Air Force Base to evaluate strengths and weaknesses documented on the technical evaluations. We also interviewed representatives from the Air Combat Command at Langley Air Force Base, Virginia. Furthermore, we reviewed the RFP to determine how the past performance of the offerors was assessed and how their proposed costs were evaluated.

We performed this audit from May 2003 through May 2004 in accordance with generally accepted government auditing standards.

Although an announced objective, we did not review the management control program because our scope was limited to evaluating only the task order requested by Senator McCain.

**Use of Computer-Processed Data.** We did not use computer-processed data to perform this audit.

**Use of Technical Assistance.** We used Quantitative Methods analysts for information and guidance relating to the Delphi technique.

**Government Accountability Office High-Risk Area.** The Government Accountability Office has identified several high-risk areas in DoD. This report provides coverage of the Defense Contract Management area.

### Prior Coverage

No prior coverage has been conducted on the Air Force's source selection of the F-15 trainer support contract during the last 5 years.

## Appendix B. Evaluation of Weaknesses Identified during the Technical/Management Evaluation

### Index B-1. Boeing

<u>No.</u>	<u>Evaluator</u> <sup>1</sup>	<u>Description of Weakness</u>	<u>IG DoD Assessment of the Inappropriate Weaknesses</u>
B1	2/E	High labor rate/high cost.	Relates to “Cost/Price” factor.
B2	2/B	Labor rates are above average.	Relates to “Cost/Price” factor.

### Index B-2. L-3 Communications

<u>No.</u>	<u>Evaluator</u> <sup>1</sup>	<u>Description of Weakness</u>	<u>IG DoD Assessment of the Inappropriate Weaknesses</u>
L1	1/B	Software only update. This places the TFE-21 conversion approach at high risk.	Relates to subfactor 6 <sup>2</sup> .
L2	1/B	Provided estimates for Suite 4+ when the Air Force requested total savings.	Relates to “Cost/Price” factor.
L3	1/B	Over inflated the hours to compensate for low labor rates.	Relates to “Cost/Price” factor.
L4	1/H	Contractor ACA status and expectation creates significant risk.	Prospective contractors refused to sign an ACA until contract award.
L5	1/H	Contractor Statement of Work is weak.	Does not relate to technical/management subfactor criteria.
L6	1/J	Offeror's plan to keep the TSSC in the Phoenix area, while eliminating the need to move personnel, keeps the location far from an “F-15 centered” Air Force base.	Does not relate to technical/management subfactor criteria.
L7	1/J	The potential for refusal by some subcontractors to enter into ACAs with incumbent may thwart the ability of L3 Communications to secure necessary data with which to accomplish any tasking.	Prospective contractors refused to sign an ACA until contract award.
L8	1/J	“Late requests” for ACAs with subcontractors shows poor planning capability on the part of the incumbent.	Multiple subfactor ratings downgraded for this weakness. (See No. L4 and L7.)
L9	1/G	RPS is only mentioned once under ACA and is not part of concurrency process.	Relates to subfactors 2 and 6. <sup>2</sup>
L10	1/G	Proposed receipt of various RPS loads in operational flight plan development process is not mentioned.	Relates to subfactors 2 and 6. <sup>2</sup>
L11	1/G	Concurrency will be done with the contractor remaining in the Phoenix area. This is a short-term solution (personnel retention) to a long-term problem (efficient use of personnel).	Does not relate to technical/management subfactor criteria.

<b>Subfactor/ No.</b>	<b>Evaluator<sup>1</sup></b>	<b>Description of Weakness</b>	<b>IG DoD Assessment of the Inappropriate Weaknesses</b>
L12	2/E	Viability of offeror's approach is extremely low and very high-risk or the hours are inflated to adjust for the low labor rates.	Relates to "Cost/Price" factor.
L13	2/B	Slow action in requesting ACAs.	Relates to subfactor 1 <sup>2</sup> . (See No. L4, L7, and L8.)
L14	2/B	High price tag for future TFE-21 updates.	Relates to "Cost/Price" factor.
L15	2/G	Offeror proposes \$* for TFE-21 upgrade, but assumptions make this impossible.	Relates to "Cost/Price" factor.
L16	2/J	The lack of timeliness in requesting ACA responses leads one to question the inability of the offeror to plan adequately for deadlines.	Relates to subfactor 1 <sup>2</sup> . (See No. L4, L7, L8, and L13.)
L17	2/J	The TFE-21 software conversion to RPS software is provided at * but carries a \$* price tag for hardware replacement.	Relates to "Cost/Price" factor.
L18	2/J	Proposed funding profile for the conversion plan exceeds available funding by \$*.	Relates to "Cost/Price" factor.
L19	3/H	Contractor is remaining in Phoenix, off of the military facility. The risk is that over time, the contractor will become more and more separated from the primary user.	Does not relate to technical/management subfactor criteria.
L20	3/J	The lack of timeliness in requesting ACA responses leads one to question the ability of the offeror to plan adequately for deadlines.	Relates to subfactor 1 <sup>2</sup> . (See No. L4, L7, L8, L13, and L16.)
L21	4/D	Seems like total TSSC staffing is management top heavy.	Subfactor rating downgraded even though evaluation notice was resolved.
L22	4/C	The relocation of the TSSC to a contractor facility and the closure of offices at Luke Air Force Base will present a challenge to maintain the same level of TSSC interface the program has had over the past 10 years.	Does not relate to technical/management subfactor criteria.
L23	4/I	Keeping the TSSC near but not at Luke Air Force Base could pose a problem for security measures since they are not discussed.	Does not relate to technical/management subfactor criteria.
L24	4/F	Manning at Seymour Johnson Air Force Base is of concern.	Subfactor rating downgraded even though evaluation notice was resolved.
L25	5/D	"The primary downside of leaving the F-15 TSSC near Luke Air Force Base is that no other F-15 Training Device resources or Government personnel reside in the area."	Issue reported by the evaluator as both a strength and weakness.
L26	5/I	No mention of security measures, climate control, or fire prevention in the new facility.	Subfactor rating downgraded even though evaluation notice was resolved.

\* The dollar value has been deleted because it is source selection information or contractor proprietary data.

---

<b><u>No.</u></b>	<b><u>Subfactor/ Evaluator</u></b> <sup>1</sup>	<b><u>Description of Weakness</u></b>	<b><u>IG DoD Assessment of the Inappropriate Weaknesses</u></b>
L27	5/F	Concern over having cleared facility in time for move of equipment.	Subfactor rating downgraded even though evaluation notice was resolved.
L28	5/A	Plan for relocating TSSC assets and functions. Medium risk.	Subfactor rating downgraded even though evaluation notice was resolved.
L29	6/G	Offeror bases this proposal on having an unclassified version of the RPS software January 2003, which will not happen.	Does not relate to technical/management subfactor criteria.
L30	6/G	The option to declassify current RPS 4E+ software does not address cost and schedule.	Relates to "Cost/Price" factor.

<sup>1</sup> Evaluator name was removed from this report.

<sup>2</sup> Subfactor 1 = Concurrency  
 Subfactor 2 = Commonality  
 Subfactor 3 = Baseline Requirements  
 Subfactor 4 = Program Management and Staffing  
 Subfactor 5 = Transitioning  
 Subfactor 6 = TFE-21 Conversion to RPS

---

## **Appendix C. Report Distribution**

### **Office of the Secretary of Defense**

Under Secretary of Defense for Acquisition, Technology, and Logistics  
Under Secretary of Defense (Comptroller)/Chief Financial Officer  
    Deputy Chief Financial Officer  
    Deputy Comptroller (Program/Budget)  
Director, Defense Procurement and Acquisition Policy

### **Department of the Navy**

Naval Inspector General  
Auditor General, Department of the Navy

### **Department of the Air Force**

Assistant Secretary of the Air Force (Acquisition)  
Assistant Secretary of the Air Force (Financial Management and Comptroller)  
Auditor General, Department of the Air Force  
Air Force Inspector General  
Commander, Aeronautical Systems Center  
    Director, F-15 System Program Office  
    Director, Training Systems Product Group  
Commander, Ogden Air Logistics Center  
Director, Aerospace Operations

### **Combatant Command**

Inspector General, U.S. Joint Forces Command

### **Other Defense Organizations**

Director, Defense Contract Audit Agency  
Director, Defense Contract Management Agency

### **Non-Defense Federal Organization**

Office of Management and Budget

---

## **Congressional Committees and Subcommittees, Chairman and Ranking Minority Member**

Senate Committee on Appropriations

Senate Subcommittee on Defense, Committee on Appropriations

Senate Committee on Armed Services

Senate Committee on Governmental Affairs

House Committee on Appropriations

House Subcommittee on Defense, Committee on Appropriations

House Committee on Armed Services

House Committee on Government Reform

House Subcommittee on Government Efficiency and Financial Management, Committee on Government Reform

House Subcommittee on National Security, Emerging Threats, and International Relations, Committee on Government Reform

House Subcommittee on Technology, Information Policy, Intergovernmental Relations, and the Census, Committee on Government Reform



# Department of the Air Force Comments



DEPARTMENT OF THE AIR FORCE  
WASHINGTON, DC

OFFICE OF THE ASSISTANT SECRETARY

20 SEP 2004

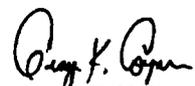
MEMORANDUM FOR DEPUTY INSPECTOR GENERAL FOR AUDITING  
OFFICE OF THE INSPECTOR GENERAL  
DEPARTMENT OF DEFENSE

FROM: SAF/AQC  
1060 Air Force Pentagon  
Washington, DC 20330-1060

SUBJECT: Air Force Response to DoD IG Draft Report of Audit, Project No. D2003CH-0113  
Award of the Air Force F-15 Trainer Support Contract

I concur in the attached response to subject audit. The Air Force concurs in all four recommendations. However, the Air Force maintains that the criteria and procedures documented in correspondence and published in the final RFP, were strictly adhered to and resulted in an efficient, effective, and fair evaluation of all proposals.

Point of contact for the Air Force response is: Ms. Kathleen James, SAF/AQCK, (703)588-7012, email: [Kathleen.james@pentagon.af.mil](mailto:Kathleen.james@pentagon.af.mil)

  
GRAY K. COYNER, Col, USAFR  
Associate Deputy Assistant  
Secretary (Contracting)  
Assistant Secretary (Acquisition)

Attachment:  
20 Sep 04 OO-ALC/CC Memorandum w/Atch



DEPARTMENT OF THE AIR FORCE  
HEADQUARTERS OGDEN AIR LOGISTICS CENTER (AFMC)  
HILL AIR FORCE BASE, UTAH

20 September 2004

MEMORANDUM FOR SAF/AQC

FROM: OO-ALC/CC  
7981 Georgia Street  
Hill Air Force Base, Utah 84056-5824

SUBJECT: DoD (IG) Draft Report, Award of the Air Force F-15 Trainer Support Contract. (Project No. D2003CH-0113) 29 Jun 04

1. This is in response to the DoD IG request that the Commander, Ogden Air Logistics Center, in conjunction with the Director, Training Systems Product Group, Wright-Patterson Air Force Base, provide Air Force comments on subject draft report.
2. The stated objective of this audit was to determine whether the Air Force decision to award a contract for F-15 trainer support was made in accordance with the technical/management, past performance, and cost/price criteria contained in the RFP and provides best value to the Air Force. There are a number of items in this report that are outside the scope of the stated objective; e.g., the audit report included findings regarding Air Force's methodology in developing evaluation criteria, inappropriate technical/management evaluation, and methodology to evaluate past performance. However, the Air Force concurs with the overall intent of the findings and recommendations contained in the draft report.
3. The Air Force maintains that the criteria and procedures discussed with industry during Industry Days, documented in correspondence and published in the final RFP documents, were strictly adhered to and resulted in an efficient, effective, and fair evaluation of all proposals. The Order Award Authority (OAA) had sufficient information to make a "best value" decision using the results of the technical/management evaluation, past performance results, and total evaluated price. While the government order selection team and the OAA strived for maximum objectivity, the selection process by its nature is subjective, so professional judgment was implicit throughout the entire process. The discriminating factors in the award decision were based on the technical superiority, quality, and risk at a reasonable and realistic price of the successful offeror.

  
KEVIN J. SULLIVAN, Major General, USAF  
Commander

Attachment:  
AF Response to DoD (IG) Draft Report,  
Project No. D2003CH-0113

Air Force Response to DoD (IG) Draft Report D2003CH-0113

We have reviewed the draft version of the audit report, and offer the following comments and management response to the report recommendations.

**Findings and Management Response**

**1. The integrated product team failed to effectively use the Delphi technique (evaluation process) to develop technical/management criteria and document how those procedures were used to minimize potential biases that may influence the criteria.**

**Response:** The Air Force agrees that it did not use the Delphi Technique as originally developed by the Rand Corporation. In fact, it is in retrospect a misnomer to call the process the Air Force used in assigning numeric values a “Delphi method.” The Rand Corporation model of the Delphi Technique calls for use of the consensus building tool to develop the measures to be used in the evaluation process, in this case, the evaluation factors and subfactors. The Air Force did not employ this model to develop the evaluation factors and subfactors. Instead, it chose to use an equally valid methodology, which employs consideration of the risks involved in the acquisition so that critical factors and subfactors may be developed accordingly. Most importantly, while the Air Force concedes a true “Delphi technique” was not part of our methodology, we followed an acceptable evaluation process in both developing our criteria and in evaluating our proposals. The methodology employed precisely followed what was specified in the solicitation itself, which was issued in draft form to allow prospective offerors the opportunity to respond with any concerns regarding the approach. We did not receive any indication during that phase that any offerors had concerns with our evaluation approach.

The Air Force uses risk assessment to determine the risk areas of planned acquisitions, and develops evaluation criteria around the risks identified and the technical requirements that are most likely to be discriminators during the evaluation. This process was used by the IPT to develop the evaluation criteria for the F-15 Trainer Support contract. The Air Force guidance concerning using risk assessment and discriminators in determining evaluation criteria can be found in the Air Force Source Selection Procedures Guide, March 2000. Had this source selection decision been protested to the Government Accountability Office (GAO), we believe that GAO would have supported our selection of evaluation criteria in this case, as well as the evaluation that was conducted here. Accordingly, we recommend that the DOD IG seek a peer review of this audit by the GAO, particularly by the bid protest section of the GAO. The Air Force would be willing to participate in discussions with your office and GAO concerning this audit. In any event, the DOD IG should not attempt to establish new standards to assess the validity of source selection decisions. Rather, the DOD IG should apply the standards adopted by GAO over the past 80 years.

**2. The integrated product team failed to appropriately use the Delphi technique to evaluate proposals from offerors by applying numeric values “0” to “5” (scientific or knowledge-based measurements) to the subfactors and not achieving consensus or ranking the offerors in association with the evaluation criteria.**

**Response:** The Air Force agrees that the Delphi Technique was not used precisely as identified in the audit and developed by the Rand Corporation. However, the procedures used were in accordance with the described procedures outlined in the TSA II User’s Guide and the RFP. The basis for the evaluation and ultimate decision by the Task Order Award Authority in no way deviated from the evaluation approach set forth in the RFP and in accordance with the Air Force Source Selection Guide (March 2000) and regulatory guidance for awarding orders under multiple award contracts. Further, while the evaluation technique and modifications to the Delphi Technique in its limited use in the source selection process could be improved upon, the audit report itself produces no findings or recommendations that demonstrate the procedures specified in accordance with the Air Force Federal Acquisition Regulation Supplement (AFFARS) 5352.216-9001, Awarding Orders Under Multiple Award Contracts, were not followed, nor that there was any arbitrariness in the way ratings were assigned. Further, by demonstration in the report itself, even had the explicit Delphi Technique been used in its entirety, there is no basis to conclude that the outcome of the evaluation and decision made would have been different. Perhaps the biggest “failure” in the Air Force approach was simply one of inappropriately citing a technique that was not used in its purest form, that is, the Delphi Technique. However, not having followed such a technique does not render the entire process that was employed, and clearly spelled out in the RFP, invalid.

**3. The integrated product team lowered ratings for L3 Communications for weaknesses that did not relate to the subfactor evaluation criteria, lowered ratings for the same weakness under multiple subfactor criteria, and lowered ratings for weaknesses that had been resolved through evaluation notices.**

**Technical/Management – Weaknesses Noted.** The Air Force disagrees that L-3’s ratings were unfairly lowered based on weaknesses unrelated to the evaluation criteria, considering the same weakness under multiple subfactors, and weaknesses that had been resolved through evaluation notices.

**Weaknesses Not Related to Criteria.** The report cites a weakness in the TFE-21 section that was considered a risk to the L-3 Concurrency approach. The report does not specify the weakness under discussion, but it is believed to be Number L1, which addresses L-3’s failure to include appropriate hardware changes in the TFE-21 conversion. The evaluation team considered the L-3 approach to concurrency, as exemplified in their TFE-21 conversion proposal, to be too focused on software and lacking in appropriate hardware changes. Because excess reliance on software solutions complicates concurrency changes, this approach jeopardizes L-3’s ability to meet the 60-day concurrency window. This specifically falls under the third element of the Concurrency criterion.

For Official Use Only  
Source Selection Information – See FAR 2.101 and 3.104

- 3 -

The report and appendix cite several weaknesses as being related to the Cost/Price factor. The Cost/Price evaluation does not have weaknesses *per se*, and the weaknesses under discussion were written against information provided in the technical proposal. Frequently, issues that might at first appear to be cost issues are actually related to the proposed technical approach; for example, a flawed technical approach could result in a proposal with more hours than would reasonably be expected.

**Same Weaknesses Under Multiple Criteria.** The report criticizes the Air Force for downgrading different aspects of L-3's technical proposal for a single weakness. It should be noted that all offerors were evaluated this way; when a single weakness affected more than one criteria of any offeror's proposal, the weakness was considered under all the criteria affected. This is not inappropriate, in that a flaw that would affect different areas of a contractor's performance should be recognized to present a risk to all the affected areas.

The report takes exception to an evaluator lowering ratings under Concurrence, Commonality, and Baseline Requirements because L-3 had not taken timely steps to enter into ACAs and states that ACAs should be evaluated only under Concurrence. The Air Force disagrees. The RFP explicitly addressed the ACA process in the Proposal Preparation Instructions (PPI) and Evaluation Criteria related to Concurrence. For Commonality, the PPI stated, "The lack of an ACA or other satisfactory method of acquiring data could legitimately cause a weakness in Commonality." The evaluation criterion includes identification of the necessary data and an approach to obtaining and interpreting the data. ACAs are also directly related to a number of items under Baseline Requirements, such as configuration management and threat system databases. The use, non-use, or failure to timely pursue ACAs affect all three areas differently, with consequences that must be assessed to arrive at an accurate measure of a proposal's quality and likelihood to meet the Air Force's needs.

It should be noted that multiple write-ups of a weakness under a single evaluation subfactor did not result in multiple downgrades of a proposal.

**Weaknesses Resolved Through Evaluation Notices.** The report assumes that any response to an evaluation notice (EN) resolves the weakness. This is not the case. An EN is issued when the evaluator does not have adequate information to complete the evaluation. If the response to an EN provides adequate information for the evaluator to complete the evaluation, the EN is closed. Closure of an EN does not mean that the information provided resolved the related weakness.

**4. The methodology used by the Air Force to assess the past performance of the offerors effectively made past performance a nonfactor in the award decision because offerors with outstanding past performance and offerors with no prior performance received equal ratings.**

For Official Use Only  
Source Selection Information - See FAR 2.101 and 3.104

- 4 -

**Past Performance - All Offerors Equal.** The Air Force agrees that the methodology used to evaluate past performance, essentially an "acceptable vs. nonacceptable" methodology, did not allow discrimination among offerors with acceptable past performance.

In this case, all offerors received the same past performance rating because no offerors had an unacceptable past performance history. The evaluation methodology was structured to recognize bad past performance. In retrospect, the Air Force realizes that allowing more granularity in the past performance ratings would achieve a more precise means of differentiating among the offerors in the past performance area. However, even if qualitative differences among past performance ratings had been employed for offerors with acceptable past performance, no past performance ratings assigned would have outweighed the technical merits of the selected offeror's proposal who also had acceptable past performance.

The past performance evaluation methodology was included in the draft RFP posted for offeror comment. If any prospective offeror objected to the evaluation scheme, exception should have been taken before the RFP was issued. As with many of the conclusions of the DOD IG in this audit, the GAO would have likely found this assertion untimely had it first been raised in a post-award protest.

**5. As a result, the Air Force Order Award Authority did not have reliable information to support the "best value" decision to award Boeing a 10-year task order for \$ [REDACTED] versus awarding the task order to L3 Communications for [REDACTED], a difference of \$31.4 million.**

The Air Force disagrees that the Air Force Order Award Authority did not have reliable information to support the "best value" decision. We have provided ample evidence to demonstrate that the stated evaluation criteria were followed, that offerors provided sufficient information in their proposals to conduct evaluations and assign proposal ratings, and that the Order Award Authority was given sufficient and reliable information regarding those ratings to support a "best value" decision.

Further, the Basis for Award in the RFP stated that the Air Force would award to a higher priced offeror if the technical superiority of the proposal outweighed the cost difference. The Order Decision Authority adequately justified the trade-off of cost and performance.

"Total evaluated price" is a source selection term referring to the outcome of an evaluation; it is not the value of the resultant contract. In the case of the F-15 order, the cost/price evaluation included estimates that are not binding amounts on the order. If the estimates are excluded from both cost proposals, the values of the orders associated with both proposals are \$ [REDACTED] for L-3 and \$ [REDACTED] for Boeing, a difference of \$10,576,852.

**Conclusion.** The Air Force disagrees with the report conclusion that the evaluation was flawed. The report cites flaws in the application of the Delphi technique. That the Air Force did not follow a technique defined by the IG after the evaluation was completed does not

Source selection information or contractor proprietary data has been blacked out.

For Official Use Only  
Source Selection Information – See FAR 2.101 and 3.104

- 5 -

mean that the evaluation conducted in accordance with the RFP was flawed. We have adequately demonstrated that the evaluation criteria stated in the RFP were followed, and that all offerors were rated consistently in accordance with the stated criteria. The Air Force is required by FAR 17.207(c)(3) to ensure that any options exercised in the future are the most advantageous method of fulfilling the Government's need, price and other factors considered, before every option is exercised. This type of assessment will be conducted prior to exercise of all future options for the F-15 training system.

**Recommendations**

1. Establish detailed formal procedures to effectively use the Delphi technique to develop technical/management criteria and document how those procedures were used in future evaluations.

Response: Concur. The Air Force agrees that should it choose to mandate use of the Delphi technique for developing source selection criteria in the future, it will document the procedures and evaluation results accordingly. At this time, the Air Force has no plans to mandate use of the Delphi technique for establishing evaluation criteria. We will continue to develop criteria in accordance with the Air Force Source Selection Procedures Guide, March 2000, which takes into account the results of risk assessments and the technical requirements that are most likely to be discriminators during the evaluation. If the Air Force plans to use the Delphi Technique in the future on the TSA II program, formal procedures will be developed and incorporated into the TSA II User's guide as appropriate.

2. Establish detailed formal procedures that appropriately use the Delphi technique to evaluate proposals from offerors to achieve consensus and rank offerors in association with the evaluation criteria.

Response: Concur. If the Air Force decides to establish a policy mandating use of the Delphi Technique to achieve consensus among the evaluators in the future, we will establish detailed formal procedures regarding how that technique will be followed. At this time, however, we have no plans to establish such a policy.

3. Either develop an effective past performance rating system for future competitive Training Systems Acquisitions Two orders or eliminate past performance as an evaluation criteria.

---

For Official Use Only  
Source Selection Information – See FAR 2.101 and 3.104

- 6 -

**Response: Concur.** The TSPG has decided to amend the TSA II Users' Guide to require use of a neutral rating for offerors without past performance records when past performance is assessed

**4. Determine whether further action on the F-15 trainer support task order is appropriate before extending the contract.**

**Response: Concur.** Before any extension of this order is approved, the Air Force will ensure that the extension of the order is the most advantageous method of fulfilling the Government's need, price and other factors considered.

## **Team Members**

The Office of the Deputy Inspector General for Auditing of the Department of Defense, Contract Management prepared this report. Personnel of the Office of the Inspector General of the Department of Defense who contributed to the report are listed below.

David K. Steensma  
Henry F. Kleinknecht  
Patrick J. Nix  
Jason T. Steinhart  
Michael B. Vandesteene