

**January 5, 1999**

**P.S. Protest No. 98-22**

## **INTERNATIONAL BUSINESS MACHINES CORPORATION**

**Solicitation No. 102590-98-A-0047**

### **DIGEST**

Protest involving testing and evaluation of flats sorting machines is dismissed in part and denied in part. Protester's assertions concerning test of successful offeror's machine are not established, issues arising in the course of protester's test under testing contract are not for resolution in protest process, and contracting officer acted within her discretion in declining to consider performance of protester's machine outside test parameters.

### **DECISION**

International Business Machines Corporation (IBM) protests various aspects of the award of a contract for Next Generation Flat Sorting Machines (NGFSM) to Mannesmann Dematic Rapistan Systems Corp. (Rapistan).

The Postal Service sorts large volumes of flats mail, such as catalogs, periodicals, and large first class envelopes, by machine and by hand.<sup>1</sup> On February 12, 1998, the Postal Service issued solicitation 102590-98-A-0047, seeking proposals to manufacture and install 175 fully automated NGFSMs which could feed flat mail automatically and process mail by optical character recognition and video coding technologies. The initial purchase was 175 machines, with options for additional quantities of machines sufficient to sort the flats currently handled on 814 Model 881 flat sorting machines.

---

<sup>1</sup> This summary of the circumstances preceding contract award is taken, in large part, from the contracting officer's statement and its supplement.

The Postal Service began investigating the benefits of using fully automated flat sorting machines in early 1997, when it purchased such a machine designed by Alcatel Postal Automation Systems (Alcatel)<sup>2</sup> and installed at the Dominick V. Daniels Processing and Distribution Center in Kearny, NJ. In July, 1997, the Postal Service issued a notice in the Commerce Business Daily to seek prequalified sources for the NGFSM. IBM, Siemens, and Alcatel responded to the notice.

In September, 1997, the Postal Service conducted pre-qualification tests of the IBM, Siemens, and Alcatel machines. The IBM machine, designed by the Swiss company Muller Martini, failed its pre-qualification test in Switzerland because it damaged a large percentage of the test mail. The Siemens machine, tested in Germany, failed because of a high jam and damage rate. Although only the Alcatel machine, tested in New Jersey, passed the initial pre-qualification test, the Postal Service conditionally prequalified both IBM and Siemens.

IBM requested permission to install its machine in a Postal Service P&DC to conduct further pre-qualification testing. The Postal Service made a site in St. Paul, MN, available to IBM beginning January 5, 1998. IBM installed its machine in February and began training Postal Service operators at that time. On March 12, IBM passed the pre-qualification test.

In January, the Postal Service provided a draft Performance Cost Model to the three vendors which described how the Postal Service would evaluate the performance of the competing machines once the NGFSM solicitation was issued. It indicated that the Postal Service would compare each vendor's cost to process 62 million pieces per day for 10 years and would measure missort and misfaced rates from test deck runs and damage rates, throughput, and other factors from live mail and would charge \$.29 for each missort. As amended March 11, 1998, the solicitation contained essentially the same cost model as provided in January.

To compensate the vendors for transporting, installing and supporting their equipment, the Postal Service entered into testing contracts with each of them. (The Postal Service paid IBM approximately \$1.2 million under the IBM Test Contract for supporting its machine throughout the test period.) Each contract included the NGFSM Competitive Test Plan of April 27.<sup>3</sup>

---

<sup>2</sup> Just prior to the contract award, Mannesmann Dematic purchased Alcatel and assigned the manufacture of the Alcatel-designed machine to its subsidiary Rapistan. Generally, this decision refers to Alcatel with respect to actions and events prior to award, and to Rapistan thereafter.

<sup>3</sup> The test plan was amended on June 23, 1998 to include retest procedures. The retest is explained in further detail below.

The total test period was 12 weeks. It included four weeks for installation, four weeks for pretest, and four weeks for the formal test.

During installation, the Postal Service provided live mail and 24-hour access to the facility and allowed the contractors time to train Postal Service mail processing clerks in the operation of their machines. The Postal Service believed that this was more than ample time for training because of its experience that operators loading mail onto automated machines become proficient within a few weeks of practice and future experience does not significantly improve their ability. IBM's proposal provided for only eight hours of training for its machine operators.

The four-week pretest period allowed the vendor to train test assistants to conduct and practice the test procedures while machine operators continued to practice mail loading operations.

The test plan provided that the eight-week installation and pretest period is the contractor's time to make all adjustments and fine tune its machine. Test Plan 4.3.1 provided:

Installation and Pretest times are available to the contractor to make any final adjustments to the equipment. Once the formal tests start, no changes shall be made, except with the prior written approval of the USPS Program Manager. No adjustments or allowances are made in any data recorded before an approved change.

Only preventive maintenance previously specified in the vendor's written maintenance plan was permitted after testing began.

The test plan provided for four weeks of testing to collect operating information.<sup>4</sup> It stressed the importance of accurate sortation to the Postal Service, stating as a test

---

<sup>4</sup> Information concerning the following characteristics was collected:

1. Feeder throughput
2. Operator productivity
3. Rejects
4. Fly-outs
5. Jams/stops
6. Mail damage
7. Multiple feeds
8. Mail stock integrity
9. Machine sort errors
10. Carbon transfer rate
11. On-line video encoding performance
12. Tray dispatch time
13. Staffing requirements
14. Maintenance

(Footnote continued on next page.)

objective the Postal Service's focus on "the labor cost per piece *correctly* processed as the critical element." Test Plan at 1.2 (emphasis added.)

Before testing began, Siemens pointed that Postal Service operators had been using Alcatel's machine for nearly one year, contending that that experience gave Alcatel an unfair advantage in testing, and requesting that the experienced operators be relocated and inexperienced temporary employees be detailed to the Alcatel machine during testing.<sup>5</sup> Because that approach would have been disruptive to postal operations, and because it was believed that Siemens could adequately train operators within the prescribed installation and pretest period, the Postal Service decided not to accede to Siemens' request. However, in the interest of fairness, the pretest period, which previously had been two weeks, was extended to four weeks. When the amendment extending the period was issued IBM responded by e-mail that it did not require the additional two weeks, and that it was prepared to proceed in accordance with the original schedule.<sup>6</sup>

The test was conducted simultaneously between May 4 and May 29 for all vendors. The four-week test consisted of 14 hours of live mail processing to determine throughput, among other performance characteristics. In addition, two test decks were processed to determine missort and misfaced rates.

The IBM machine performed at a missort error rate more than double the Postal Service's performance requirement of an error rate less than 1%.

On or about June 17, the Program Manager informed the contracting officer that testing personnel at St. Paul, MN failed to collect misfaced mail data for IBM during the test deck runs. Consequently, these tests had to be rerun. For the retest, the Postal Service developed new test decks, representative of the normal mix of FSM 881 mail.<sup>7</sup> To avoid

---

(Continued from previous page.)

- 15. Ergonomics and Safety
- 16. Ability to process all USPS flat mail . . . within [specified dimensions].

<sup>5</sup> According to the contracting officer, IBM, like Siemens, knew that the Kearny operators had been processing mail with Alcatel's machine for over one year. IBM asserts that it first learned this at the debriefing.

<sup>6</sup> The substance of Siemens' concern about operator experience was not disclosed to the other offerors. The amendment extending the pretest stated no reason for the extension, and Purchasing's notes of its telephone conversations advising the other offerors of the extension do not reflect the explanation offered by the contracting officer. Those notes and IBM's e-mail message reflect IBM's concern that it had been working extended hours following the denial of a request for an extension which it had previously made, so that the extension, while "useful, . . . was not necessary."

<sup>7</sup> The contracting officer's rebuttal comments elaborated on the reasons for the change. Test Deck A of the original test was composed of dead letter flat mail of the three primary flat mail types in the in ap-

(Footnote continued on next page.)

further errors in test procedures, it was decided to run the retests sequentially, using a single team of test supervisors who would travel to each site.

On June 23 and 24, in individual discussions the offerors were informed of the need for the retest, and advised that called each offeror to inform them that a retest of the test deck runs was necessary and that no modifications should be made to their machines. All three offerors agreed to the retest and agreed not to modify their machines.<sup>8</sup>

---

(Continued from previous page.)

proximately the expected distributions appearing in the mailstream: 60% advertising; 25% periodicals; and 15% First Class. In contrast, Test Deck B had been specially constructed to have pieces representing all the size ranges appearing in the mailstream in equal amounts, so that very small and very thick pieces were equally represented in Test Deck B with middle size catalogs. As a result, Test Deck B had disproportionately large numbers of difficult size thin and very thick pieces when compared to their representation in the mailstream which would identify particular problems a machine might have with certain size pieces.

There were inconsistencies between the error rate data from the initial test and error rates on live mail and error rates were inconsistent between Test Deck A and B, with machines generally displaying higher error rates on the B deck. Consequently, when it was necessary to conduct the retest because it had failed to collect IBM's misfaced data during the initial test, Test Deck B was dropped because it was not representative of mail in the normal mailstream. Instead, the retest test decks were constructed to be similar to Test Deck A, including mail in the same distributions as it occurred in the mailstream. In addition, the Postal Service constructed the retest test decks to ensure that the test decks were identical. (There were five retest decks, A-F. Each retest deck with the same letter designation was identical.

<sup>8</sup> The offerors' machines had to be in the same design configuration as in the four-week formal test because use the formal test data was to be used for all Cost Model inputs except missort and misfaced rates. Engineering staff personnel explained this reasoning to the vendors during technical discussions on June 24 and 25.

On June 1, the Postal Service had issued identical modifications (Mod 2) to the test contracts of Siemens and IBM to extend the period in which their machines would reside at Postal Service facilities so they would be available to test competing optical character readers (OCRs) on the winning offeror's NGFSM machine and to give each the alternative to continue processing mail until the commencement of OCR testing at no cost to the Postal Service or cover the machine and leave it idle. (No Alcatel modification was needed since that machine was postal-owned.) Siemens agreed to the no-cost modification and executed Mod 2 on June 16. IBM, however, sought payment if its machine was to be used for mail processing during the post-test period, and the Postal Service agreed to fund that use. A revised Mod 2 reflecting that payment was provided to IBM on June 24. That revised modification reflected the decision to conduct the retest, and included this statement: "THE CONTRACTOR MAY MAKE NO CHANGES TO THE MACHINE UNTIL THE COMPLETION OF ANY AND ALL RETESTING FOR THE NGFSM COMPETITIVE TEST." IBM executed that modification on June 29.

All three of the test contracts were subsequently modified to incorporate the test plan revisions which reflected the retest. Each of those modifications (Alcatel Mod 4, executed July 16; Siemens Mod 3, executed July 24, and IBM Mod 3, executed July 21: contained the statement: "THE CONTRACTOR MAY MAKE NO MODIFICATIONS TO THE MACHINE UNTIL THE COMPLETION OF ANY AND ALL RETESTING FOR THE NGFSM COMPETITIVE TEST."

(Footnote continued on next page.)

The retest consisted of 14 runs with the five 1,000 piece test decks. Four of the test decks were used three times; the fifth test deck was used twice. The Postal Service arranged each deck in random order before it was transported to the test site. After the test deck pieces were sorted by the machine under test, the pieces were shuffled to avoid significant numbers of pieces destined for the same bin appearing in consecutive order.

The directions to machine operators at each test site to shuffle the second and third runs of each test deck to create a random deck were given orally. No particular procedure for shuffling the test deck was prescribed.

On July 8-10, the Postal Service conducted the Siemens retest. On July 13-15, the Postal Service conducted the Alcatel retest. On one run, Alcatel's machine displayed an unusually high read reject rate. The Postal Service permitted Alcatel to correct the problem by performing a minor repair requiring 15 minutes to accomplish, and reran the test.<sup>9</sup>

On July 16-19, the Postal Service performed the retest of IBM's machine. On the first test deck run, a feature of IBM's machine unrelated to the items being scored did not function properly. The Postal Service permitted IBM to deal with the issue manually and did not count the run. After the Postal Service test team shuffled the deck, IBM began its new run 1. The machine missorted mail at a rate much higher than observed on any other machine previously tested. After observing the IBM machine's problems on the unscored run and its high missort rate on run 1, Postal Service testing personnel stopped testing, took a long lunch, and permitted IBM an opportunity to check its machine. When the Postal Service testing personnel returned, IBM initially stated that its machine was ready to perform run 2. Upon further reflection, IBM agreed with Postal Service personnel that it should perform further analysis and resume testing the next morning.

When testing resumed, IBM scored a relatively low, for IBM, missort rate for run 2. However, the hand counts of pieces sorted exceeded the machine count by 26 with only 7 missorts recorded.<sup>10</sup> On run 5, the second run of test deck B, the hand count exceeded machine count by 20 pieces and only 9 were recorded as errors. The errors

---

(Continued from previous page.)

One of the items the Postal Service pointed out to IBM in its June discussions was its machine's excessive error rate.

<sup>9</sup> This read reject rate did not affect Alcatel's Cost Model score, but reduced the quantity of the mail pieces used for calculating Alcatel's missort and misfeed scores.

<sup>10</sup> That is, the physical count of the number of mailpieces sorted to the destination bins exceeded the number of mailpieces which the machine had counted as having been sorted.

were generally recognized as double pulls (in which the machine fed two mail pieces instead of one, dropping both into the same destination bucket), because the Postal Service's hand count of pieces in buckets exceeded the machine's count of the pieces sorted by the same number as the number missorted. When the machine feeds two pieces and drops the two pieces in the correct bucket for the first of the two pieces the hand count of pieces in buckets will exceed the machine's count by one, and the second piece should be a missort (*i.e.*, in the wrong bucket), unless by happenstance it has an address that belongs in that bucket. (The instances of unrecognized double-pull missorts increased if the test decks were not fully shuffled between tests.)

The Postal Service recognized that the first five runs demonstrated that the machine incurred many double pull pieces which were not being recorded as errors. This indicated that the deck had many pieces destined for the same bin location in sequence because the test decks had not been adequately shuffled. The Postal Service Test Director instructed the local personnel to shuffle the decks more carefully.

In subsequent runs, the IBM machine continued to incur double pull pieces, but nearly all were recorded as errors. After run 10, IBM demanded that the Postal Service stop testing and allow IBM to adjust its machine to stop its high missort rate. The Postal Service believed that the IBM machine was functioning in its normal manner, which produced high missort rates due to large numbers of double pull pieces, but that there no machine failure existed to prevent continued testing. However, the Postal Service allowed IBM to stop testing and allowed it a three hour period to adjust its machine. After the three-hour period, testing resumed, and while IBM's performance improved, the error rate was still over the 1% requirement.

After its poor missort performance in the retest, IBM sent the Postal Service a "White Paper," providing its explanation of why its missort rate was so high, suggesting that only the last four runs of the retest represent "true" performance, proposing potential modifications to reduce missorts, and generally describing other advantages of IBM's machine and management team. IBM asserted that its machine performed poorly in runs 1-10 because it mistakenly loaded an older set of software parameters. It also proposed specific physical modifications to reduce the machine's high missort rate. IBM requested that only the last four runs of its test deck retest be considered in the Performance Cost Model and suggested that additional test deck runs be made to validate the accuracy of the last four runs of the test deck retest.

The contracting officer considered both IBM requests, but ultimately rejected them. She could not ignore IBM's first 10 retest runs because plan for the retest stated that the Postal Service would use 14 runs to determine missort rates. She considered conducting a second retest, but believed the only fair way to conduct such a retest would be to allow all vendors the opportunity to further adjust their machines and retest, entailing another lengthy delay in deployment. In light of the Postal Service's urgent need to begin deployment and the estimated cost of \$357,000 per day for each day of deployment

delay, she decided that was unacceptable to the Postal Service. She also asked Postal Service engineering experts to review the modifications IBM proposed. They did not believe that IBM's proposed modifications could significantly reduce the IBM's machine's missort rate.

The solicitation established three factors for evaluating proposals for award:

1. Performance as measured by the Postal Service's cost model (most significant evaluation factor)
2. Technical evaluation of proposals (second most significant evaluation factor)
3. Price

On August 14, the NGFSM contract was awarded to Rapistan. The Postal Service awarded the contract to Rapistan because it represented the best value to the Postal Service. It virtually tied with Siemens for the best Cost Model Performance, presented a technical proposal essentially equivalent to the other offerors, and offered the lowest price. (Siemens had the second lowest price; IBM's price was third low.) In addition, Rapistan's machine accurately sorted flat mail and currently meets all the Postal Service's performance requirements.

Shortly after the award, Dave Bausch of IBM called the Postal Service Manager of Acquisition Management, leaving a voice mail message which thanked the Postal Service for its professionalism and fair dealing in the conduct of this purchase.

IBM subsequently requested and received a debriefing, at which some of Alcatel's performance data from the retest was disclosed. This protest followed. The protest raises the following issues (an issue concerning the calculation of IBM's score was withdrawn and is not discussed herein):

— Equipment Modification

IBM was unfairly prohibited from modifying its equipment between the end of the initial test and the beginning of the retest when the other participants were operating and modifying their machines. Had IBM been allowed to do so, it would have modified its machine in a way which would have significantly improved its missort error rate by correcting its double-pull problem.

The prohibition on modifications announced at the discussions on June 23 and 24 overlooked modifications vendors might already have made. Only IBM signed a modification (Mod 2) addressing changes before the retest. Rapistan's failure to sign such a modification until after the retest "can only mean that it would not agree to the 'no modification' policy until after the retest's conclusion." The "only reasonable explanation" for the substantial worsening in Rapistan's mail damage



rate between the initial test and the retest “is that Rapistan must have modified its equipment.”

— Operator Training

Skilled operators can significantly improve machine missort rates. The use of postal operators experienced in the use of the Rapistan machine gave Rapistan an advantage in the testing to IBM’s detriment.

— Inadequate Shuffling During IBM’s Retest

The retest of IBM’s machine was unfair because the test decks were not adequately unsorted in runs 2 and 3, artificially reducing IBM’s missort rates, misleading IBM into thinking that it had solved its missort problems with its adjustments following the first run.

— Disparate Treatment During Tests

Since the early runs of IBM’s retest, the last retest conducted, were not fully shuffled, and the retest plan was to be the same for each vendor, it is reasonable to assume that the decks were not shuffled for the other vendors, and that they improperly benefited from that lack. Any shuffling which may have occurred “could not possibly have [been] done . . . equally” since there was no written procedure for shuffling. (A more equitable procedure would have been for the pieces to be numbered and ordered sequentially.)

— Flawed Best Value Analysis

It was irrational for the Postal Service to rely solely on the results of the initial test and the retest in the selection of the flats sorters, when it had other information available which more accurately reflected IBM’s performance. That information included the live mail missort rates achieved during the pretest and initial tests, the results of a “fifteenth run” as part of the retest, as well as the live mail experience with the machine subsequent to the retest after IBM made the modification which it was prevented from making prior to the retest.

The contracting officer’s response to the protest includes the following:

— Rapistan did not modify its machine following the initial test, as it confirmed in a July 31 letter: “There have been no design changes to the [flat sorter] . . . in Kearny, NJ, since the beginning of the competitive test, May 4, 1998, through the completion of the competitive re-test, July 15, 1998.” The difference between the damage rates cited by IBM is explained by the fact that the initial test dam-

age rate involved live mail, while the retest damage rate involved test mail.<sup>11</sup> IBM's rate on the latter also exceeded the rate on the former (although not as much).

— IBM had ample time to train its operators, and was not disadvantaged in either test in that regard. IBM's proposal states that only limited time is necessary to train machine operators; the eight weeks allowed was more than adequate for that purpose. The Postal Service's experience is that once operators "quickly learn" how to load (stack) the mail, they "do not improve any further."

— The retest was fair. The test data indicates that the Rapistan test decks were adequately shuffled, and that its error rates were not artificially low. Operators were directed to randomize the test decks after the first and second runs; while no means was specified, none was required. Contrary to the protester's suggestion, it was not the Postal Service's practice to order test decks numerically in other tests.

— IBM was benefited in the course of the retest because it had the opportunity to correct problems with its machine after run 1 and after run 10. Further, IBM benefited from the "artificially low" missort rates for three runs prior to the discovery of the problem of inadequate shuffling.

— Conducting a further retest, as IBM proposes, would harm Rapistan, which properly prepared its machine; adversely affect the integrity of the Postal Service's purchasing process; and unacceptably delay the deployment of needed flats sorters at an estimated cost of approximately \$45 million.<sup>12</sup>

IBM submitted comments on the contracting officer's statement. It challenges the contracting officer's factual recital in many regard:

— The contracting officer's statement omits the fact that between the September 1997 pre-qualification tests and the issuance of the solicitation, the NGFSM program team recommended a noncompetitive award to Alcatel, a decision to which IBM "voiced strong objections." That recommendation warrants a careful look for "evidence of evaluator bias."

---

<sup>11</sup> Elsewhere, the contracting officer notes that after three runs, a test deck's "pieces become too frayed to be fair representations of typical mail pieces."

<sup>12</sup> The contracting officer asserts that the favorable missort rates which IBM seeks to have considered are not comparable to the retest rates. Some arise from the sortation of live mail, which is "easier to sort because much of it [consists of] consecutive pieces destined for the same ZIP Code," and others involved "unverified . . . runs" subsequent to the retest.

— The post-retest runs were not “unverified.” Run 15 involved the same test deck used in the (scored) run 14. Further subsequent tests involved the same operators used in the retests and quantities of dead mail similar to that used in the retest test decks, and were attended by postal management personnel (although not personnel from postal Headquarters).

— That postal experts did not believe that IBM could correct its machine’s problems overlooks the fact that IBM *had* corrected them.

— “Changes” or “modifications,” and not “design changes,” were prohibited prior to the retest. The contracting officer’s statement qualifies the limitation to “design changes.”

— Rapistan had an advantage in the retest because its machine was installed and operating the longest. IBM had to “recover[] from the earlier system shut-down” while Rapistan and Siemens were able to begin retest preparations upon notification.

— The Postal Service did not point out the double-pull problem at the June discussions, and repeatedly asserted its understanding that IBM had a software problem. It was only the second day of the retest that the extent of the double-pull problem was known. IBM could not have focused on the double-pull problem before the test to correct it because it was not aware of it.

With respect to the issues raised in its protest, IBM’s comments made the following points:

— Because the Postal Service did not apply the test procedures equally to all offerors, the protests must be sustained (citing *Zenith Data Systems, Inc.*, P.S. Protest No. 95-19 *et al.*, November 22, 1995; *Markim Trucking*, P.S. Protest No. 95-38, November 2, 1993, and *Concept Automation, Inc., v. United States*, 41 Fed. Cl. 361 (1998)).

— It was unfair to allow Rapistan to continue to run its machine between the tests, when IBM was not allowed to. Rapistan was not bound by the “no modifications” policy because it did not execute its amendment (which was only prospective in its application) until after the retest. Rapistan’s assertion that it made no “design changes” does not mean that it made no “modifications.”

— Comparisons of Rapistan’s performance data disclosed to IBM and disclosed to the General Counsel in the course of the protest (and disclosed to IBM only in general terms) are of such magnitude as to suggest to IBM that they may have been the result of modifications to Rapistan’s machine. Rapistan had various incentives to modify its machine between the test and the retest even before it learned that a retest would occur.

— Rapistan obtained a benefit in the retest because its operators had more experience, and thus skill. IBM is not contending that its operators were insufficiently trained, but that more experienced operators can run the equipment “so that it performs significantly better.”

— The retests were flawed because shuffling was inadequate; Rapistan’s test data show a similar pattern of hand counts exceeding machine counts. The failure arose from the absence of a procedure for sequencing the test mail such as was prescribed for the testing of the follow-on optical character readers. In IBM’s case, the error “delayed IBM’s recognition . . . that its poor retest performance was due to [a problem loading software]. Its performance thereafter was significantly better, and that performance was confirmed by its subsequent post-test performance.

— As a result, the best value analysis was flawed because it “gave undue weight to the retest data which USPS . . . knew was defective and unreliable.” The Postal Service cannot sustain a decision based on the artificial retest data when it has other data showing its performance is “drastically better.”

Counsel for Rapistan submitted comments on the contracting officer’s statement which made the following points:

— Many of IBM’s protest issues are untimely, because they involve defects in the solicitation which must be raised before the time set for the receipt of proposals (Procurement Manual (PM) 3.6.4 b. and c.). These include the purchase of the Alcatel machine, the use of experienced operators on the Alcatel machine, and problems arising in the course of the retest.

— IBM was treated fairly in the retest. Indeed, from Rapistan’s viewpoint, IBM was shown favoritism when it was allowed two opportunities to make adjustments to their machine in the course of the test. IBM was benefited by any “under-shuffling” of the test decks because it artificially lowered its missort rate, which, in any event, exceeded the missort rate required by the Postal Service.

— The information provided at IBM’s debriefing established that the test decks were adequately shuffled for Alcatel’s tests. The variations in Alcatel’s test scores were not the result of improper modifications, but related to differences in the types of mail (live mail vs. test decks).

— Rapistan made no “design modifications” to its machine either after or before it was told to make no “design changes” to it. “[T]he only adjustments made to

the machine during the pretest period before the retest were of the type IBM itself made” in the course of the retest.<sup>13</sup>

— The allegation that Alcatel would make modifications to the postal-owned machine between test and the retest is “preposterous.” It had no reason to do so because no retest was scheduled, it no longer owned the Kearney machine, and any modification which “failed” would give rise to a Postal Service claim. Further, any change necessary in connection with the production contract first article could better be made on an Alcatel-owned machine.

— IBM knew it had a double pull problem before the retest. The Postal Service properly advised IBM that it could not modify its machine to solve the problem before the retest, but IBM could have better prepared its machine by performing the modifications it made during the retest<sup>14</sup> before the retest. IBM failed to prepare its machine adequately, and is now blaming the Postal Service or Rapistan for its failure.

The contracting officer submitted comments on IBM’s submission as follows:

— IBM’s implication that the NGFSM Program team is prejudiced in favor of Alcatel is insulting, unjustified, and completely without any merit. The recommendation for noncompetitive purchase of Alcatel’s machine in September/October 1997 was reasonable, since at that time the IBM and Siemens machine had failed the prequalification tests.

— IBM is incorrect when it asserts that Alcatel had the most time to prepare for the retest. After the four-week initial test was concluded, Alcatel personnel left the NJ facility and the Alcatel machine was returned to processing mail under the Postal Service’s maintenance. After notification of the retest, Alcatel also had to send its employees to refurbish and repair its machine.

— IBM contends that the Postal Service should have evaluated IBM’s missort rate based on some unspecified combination of other missort statistics more favorable to IBM. It is well established that contracting officers must evaluate proposals according to the evaluation criteria stated in the solicitation. PM 4.2.5.a;

---

<sup>13</sup> A statement by Alcatel’s project manager describes the “refurbishment” and “comprehensive tune up” in advance of the retest, which included the substitution of a new version of a component intended to correct a problem which had required the component’s frequent replacement. The statement asserts that “[t]he use of this [new component] does not effect [sic] the performance of the machine . . . [and the replacement] did not effect [sic] the form fit or function of the machine, nor did it change the machine’s design.”

<sup>14</sup> That is, after the first (unscored) run, and following run 10.

*Serv-O-Matic*, P.S. Protest No. 91-32, August 9, 1991. In this solicitation, the most important evaluation factor was the performance cost model for which the retest amendment specified that the results of 14 runs of the Postal Service 1,000 piece test decks would determine missort rates. The Postal Service could evaluate IBM's missort performance only by calculating the average of its scores on the 14 test deck runs. The use of any other statistics would be contrary to the evaluation plan stated in the solicitation, clear error, and manifestly unfair to the other offerors.

Run 15 was performed at all sites only to ensure that all feeders could operate simultaneously. The solicitation test plan clearly states that the Postal Service would use runs 1-14 to determine the missort rate for the Performance Cost Model. It is unreasonable for IBM to suggest that run 15 should be used to substitute for or discredit runs 1-14.

The test runs IBM subsequently performed using dead letter mail were not official tests and were not verified by Postal Service Engineering officials responsible for the NGFSM program. The dead mail used was not the Postal Service retest test deck, and the results reflect IBM's modification made to resolve the double pull problem. If IBM were allowed to change its machine and record new scores, all offerors would have to be allowed the same opportunity, requiring another round of testing and a delay of over four months.

— IBM attempts to escape responsibility for its own failure by shifting it to the Postal Service. The Postal Service permitted a four-week Pretest period before commencement of the Competitive Test and IBM was given the time period from June 24 to July 16 to prepare for the retest. During those periods, IBM should have run practice tests, noted its machine's performance problems and adjusted it to correct them.

IBM knew that it had achieved a high missort rate on the four-week initial test.<sup>15</sup> IBM also knew the importance of the Performance Cost Model in the evaluation

---

<sup>15</sup> IBM admits that it was aware of its high missort rate at least as early as week 3 of the initial test. Even if the local testing personnel told IBM it was doing well in missort performance, IBM should have ignored such statements. Test personnel did not know how missort rates would be used in the Cost Model or the Postal Service Solicitation requirements, and IBM knew that its missort rates on live mail and on test decks both averaged more than the Postal Service's 1% or less requirement. At discussions on June 24, the Postal Service pointed out to IBM its excessive missort rate.

Further, It was IBM's responsibility, not the Postal Service's, to determine the cause of its problems. It is unbelievable that IBM had so little understanding of its own machine that the Postal Service's gratuitous suggestion that the missort problems looked like a software problem could influence it. Although IBM's post-retest White Paper stated that its high missort rate was caused by the loading of older and less effective software parameters, now IBM appears to be repudiating its prior assessment and claim that the entire problem is a mechanical one.

and the value of missorts in the Cost Model. IBM should have prepared its machine to produce the lowest missort rate possible.<sup>16</sup> Instead, IBM's machine performed poorly during runs 1-10 of the retest. It performed better, but still poorly, on runs 11-14, when its average missort rate still exceeded the Postal Service's requirements. Except for repeating the shuffling arguments previously made, the only reason IBM presents for removing the first 10 runs from its score is that it performed better on runs 11-14, on live mail during the four-week initial test,<sup>17</sup> and in its own tests after adding a modification.<sup>18</sup>

— IBM continues to argue that Alcatel's test decks were not adequately shuffled. In addition, it argues that the retest was unfair because the Postal Service did not have test procedures that ensured that test deck pieces would be presented in precisely the same order to each offeror. Both arguments are without merit.

It was obvious from the excess hand counts and missorted pieces for IBM and Alcatel that the IBM test decks had not been shuffled adequately until after run 5 (as reflected in the large discrepancies noted on runs 2 and 5) and that Alcatel's decks had all been adequately shuffled. While IBM points out instances where excess hand counts exceeded missorts recorded in various Alcatel runs, these were smaller discrepancies of a sort which can occur even in a random deck. Similar instances of these smaller discrepancies occurred for IBM after the Postal Service began adequately shuffling its test decks.

---

<sup>16</sup> IBM's suggestion that the Postal Service dissuaded it from correcting its machine's high missort rate is an untrue and unfair attempt to avoid its own responsibility. Postal Service personnel told IBM that (1) it could not make adjustments to its machine to correct its missort rate during the initial testing period and (2) that IBM's poor performance looked like a software problem. The first statement simply informs IBM of what the test plan specifies. The second was free advice that IBM had no obligation to consider. IBM engineers who designed the machine should have known enough about their own machine to determine if this advice was correct. Moreover, IBM stated that it had the wrong software parameters loaded during the beginning of the retest.

<sup>17</sup> The contracting officer does not dispute IBM's achieved live mail missort rates, but those rates exceeded the Postal Service 1% or less missort rate requirement.

<sup>18</sup> Contrary to IBM's assertions, IBM was favored during the retest. The IBM machine performed normally (with excessive missorts) with its older software parameters. That performance did not require a halt to testing to permit IBM to change software, but the Postal Service permitted such a halt and allowed IBM to adjust its machine. Alcatel's minor repair to its machine during its test did not affect its test scores. It was a convenience to the Postal Service which wanted more pieces to be sorted rather than rejected. Further, although Alcatel had to make a repair during the initial test, the Postal Service refused to disregard data unfavorable to Rapistan recorded prior to the repair, treating Alcatel exactly the same way IBM was treated during the retest. Finally, IBM was favored by the use of artificially low missort rates on runs 2, 3 and 5 caused by inadequate shuffling.

— IBM also contends that the Postal Service test procedures for the recent Optical Character Reader/Video Coding System (OCR/VCS) implies that the only fair way to have run the NGFSM tests was to order all test decks in precisely the same order. The OCR/VCS test and the retest are totally different. The OCR test determines how well the OCR reads addresses and matches the addresses to the correct ZIP Code by printing a log of the ZIP Code it associates with the address and certain other information. The physical mail pieces are all sorted to the *same bin*. To score this test, the correct ZIP Code for each mailpiece is determined by a computer and that ZIP Code is compared to the ZIP Code each OCR obtained. The Postal Service tests each contractor's OCR with the same test deck in the same order for convenience in making the required comparisons.

IBM's argument that the retest was unfair because the Postal Service did not have procedures to precisely order test decks fails because it is untimely. The Postal Service specified its original four-week test procedures and its retest procedures in the solicitation and its amendments. Postal Service regulations require that protests based upon alleged improprieties in the solicitation be received before the time set for receipt of proposals. Consequently, IBM should have challenged the original test plan procedures before the original four-week test began and certainly before the retest began.

— IBM continues to argue that Alcatel unfairly modified its machine between the end of the four-week initial test and the retest in order to improve its missort rate performance. IBM presents nothing more than speculation based on some of Alcatel's results to support that allegation.<sup>19</sup> It is clear from Rapistan's Comments and the Declaration of its program manager that that Alcatel did not unfairly modify its machine.<sup>20</sup> The one change it made was within the "no modification" guidelines the Postal Service established and had no impact on either missort or misfaced rates, the only criteria measured in the retest. Consequently, the change could never affect the award decision or prejudice IBM and therefore,

---

<sup>19</sup> What is more significant than the differences IBM describes between Alcatel's initial test deck scores and its retest test deck scores are the similarities between Alcatel's initial performance on test deck A and its performance on the retest test deck, which demonstrate that Alcatel did not modify its machine between the initial test and the retest. Further, Alcatel had a missort improvement is similar in size to the improvement IBM achieved between runs 1-10 and 11-14 of the retest, the result of permissible adjustments and repairs.

<sup>20</sup> IBM's counsel also implies, without foundation, that the Postal Service may have changed Alcatel's machine to improve its performance after the initial test ended and before Alcatel arrived to prepare its machine for the retest. The Postal Service did not perform anything but routine maintenance upon the Alcatel machine and would have assumed a risk of legal actions from Alcatel if it had changed any performance characteristics of the Alcatel machine prior to the retest.



can never be more than harmless error. See *Cohlma Airline, Inc.*, P. S. Protest No. 87-118, April 13, 1987.

IBM's description of the Postal Service's "no changes" rule is overly simplistic and inaccurate. The Postal Service explained at Discussions with each offeror that it could not "change its machine" prior to the retest to ensure that the performance characteristics of the machines recorded in the four-week initial test remained the same. The Postal Service was particularly concerned that offerors would slow down machines to lower missort rates. The Postal Service specifically qualified the "no change" rule to permit routine maintenance as well as adjustments and tuning of the machines. If all changes had been prohibited, the Postal Service would not have permitted IBM to make the changes it did during the retest. Rapistan's understanding of the "no changes" rule as prohibiting design changes is accurate, and its substitution of components did not violate the Postal Service's "no change" rule.<sup>21</sup>

— IBM continues to argue that the Postal Service permitted Alcatel an unfair advantage in the retest by permitting the existing operators at New Jersey, to continue at their jobs and serve as loaders and unloaders during the Alcatel retest. IBM concedes that it had ample time to prepare and train Postal Service employees in the simple job of loading mail onto feeders and removing full trays of mail. Consequently, IBM now only argues that because the Alcatel employees had been performing the job longer, they did a better job, improving Alcatel's missort rate.

As previously explained, the Postal Service experience is that employees quickly learn to perform the simple job of loading mail into automatic feeders and unloading full mail trays. Once employees learn these jobs, they do not get significantly better at them. As a result, the Postal Service does not believe that Alcatel had any advantage during retest.

Assuming, for the sake of argument, that an advantage existed, it is no more than the natural advantage of an incumbent which has long been accepted by the General Counsel as not unfair. *Pitney Bowes, Inc.*, Postal Service Protest No. 89-22, July 7, 1989 (There is "no requirement for equalizing competition by taking into consideration advantages [gained via incumbency or offeror's particular circumstances], nor do we know of any possible way in which such

---

<sup>21</sup> On a related issue, local test personnel correctly told IBM it could make no changes to its machine when it inquired during week 3 of the 4-week test. The test plan prohibited any changes or adjustments to the machine after testing had begun. The Postal Service Program Manager could have approved such a change, had IBM addressed its request to the Program Manager.

equalization could be effected.”) citing *Aerospace Engineering Services Corp.*, Comp. Gen. Dec. B-184850, March 9, 1976, 76-1 CPD ¶ 164.

In addition, IBM’s challenge is untimely. IBM should have known from the test plan that the Postal Service intended to use its own employees to act as loaders and unloaders for offeror machines. The test plan did not make any special provision for removing the employees currently processing mail with the Postal Service owned Alcatel machine. Siemens knew this fact and timely raised the issue to the Postal Service in April, 1998, before the four-week initial test began. IBM could have raised the issue then as well. If IBM wished to challenge this aspect of the four-week initial test and the retest, it should have raised the issue before the four-week initial test and with regard to the retest before the retest began. *Grand Rapids Label Company, Inc.*, P.S. Protest No. 96-22, January 31, 1997.

Rapistan submitted additional comments on IBM’s submission which generally paralleled the contracting officer’s arguments with citations to additional relevant law.

IBM submitted a final letter waiving its right to a conference, withdrawing the cost model calculation issue, and clarifying that “although it is alleging certain unequal or unreasonable treatment by [the Postal Service], it is not alleging that any individual [postal] employees acted in bad faith.”

## **DISCUSSION**

Much of the information submitted by the contracting officer, and some of the information submitted by the parties, was submitted subject to claims of privilege, as PM 3.6.7.d of the protest provision allows. As a result, substantial amounts of information about Alcatel’s test results was not available to IBM, and *vice versa*. While that information has been reviewed *in camera* in the course of preparing that decision, and informs it (*see, e.g., American Bank Note Company*, P.S. Protest No. 94-02, May 11, 1994), that information is not recited herein, except in general terms.

Basic to the protest process is the principle that the protester has the burden of establishing its case affirmatively, and that in doing so it must overcome the “presumption of correctness” which accompanies the statements of the contracting officer. *Timeplex Federal Systems, Inc.; Sprint Communications Company*, P.S. Protest Nos. 93-22; 93-24, February 2, 1994. Where bias, unfairness, or bad faith are alleged, they must be established by specific proof, not merely assumptions and suppositions. *Enpro Corporation*, P.S. Protest No. 91-48, October 9, 1991, *citing Thermico, Inc.*, P.S. Protest No. 90-71, December 21, 1990.

That principle is sufficient to resolve IBM's assertions that Alcatel's machine must have been improperly modified between the initial test and the retest,<sup>22</sup> and that its results on the retest must have been affected by the Postal Service's failure to shuffle the test decks adequately between runs. IBM's assertions are unsupported, they are refuted by the contracting officer, and nothing in the record which contradicts that refutation.

We reach a similar conclusion with respect to the issue of Alcatel's operators. The contracting officer has provided an adequate explanation of the limited role which the NGFSM operators play in the functioning of the machines, and IBM's contentions to the contrary are not persuasive.

The contentions that IBM was unfairly prohibited from modifying its machine between the initial test and the retest and that errors occurred in the course of its retest when the test decks were not adequately shuffled are not for our consideration. The tests were conducted pursuant a testing contract which contained the standard Claims and Disputes clause which implements the Contract Disputes Act of 1978. It provides, in part, that "[a]ll disputes arising under or relating to this contract [are to] be resolved under this clause." That clause provided IBM's remedy for the problems it identified in the course of the contract, for which the protest remedy is not available. *I.C. Inc.*, P.S. Protest No. 86-05, April 25, 1986.

The assertion that it was irrational not to consider information about IBM's performance outside the framework of the test is incorrect.

It is well settled that the evaluation of a proposal must be based on factors outlined in the solicitation. The contracting officer has broad discretion in the selection and weighting of evaluation criteria to determine which offers will best meet the Postal Service's actual needs. However, once offerors are informed of evaluation criteria, the procuring agency must adhere to those criteria . . . .

*TRW Financial Systems, Inc.*, P.S. Protest No. 91-19 May 29, 1991 (citations and internal quotations omitted). The contracting officer could not have considered IBM's other performance without modifying the evaluation scheme, and acted well within her discretion in declining to do so.

The protest is dismissed in part and denied in part.

---

<sup>22</sup> The repairs which Alcatel made in the course of the retest and the component substitution which it identified in its comments were consistent with the test instructions and the Postal Service's admonition, and, in an event, had no impact on the elements measured in the retest.

William J. Jones  
Senior Counsel  
Contract Protests and Policies