JOINT LEGISLATIVE AUDIT AND REVIEW COMMISSION
OF THE VIRGINIA GENERAL ASSEMBLY

Special Report: Department of Education Student Assessment Program Procurement

In Brief…

Special Report: Department of Education Student Assessment Program Procurement

In September 2005, the Department of Education (DOE) awarded a six-year, nearly $140 million contract for student testing services, DOE's largest single procurement. The Chairman of the House Appropriations Committee subsequently requested that JLARC staff conduct a review of this procurement. The JLARC Chairman authorized the staff review and appointed a special subcommittee of members (identified at right).

The JLARC staff review found several strengths of the student assessment procurement process, including that it was conducted by an evaluation committee of professional staff without outside pressure, the steps in the procurement process were completed, and the selected proposal was below budget. However, a weakness of the process was that the evaluation committee was not required to document or provide the rationale for its scoring by the procurement act, the Superintendent, or the Board of Education. Evaluation committee members had some difficulty after-the-fact clearly articulating the basis for their scoring, but it appears that the committee chose the vendor it thought had a more clear and detailed proposal; a fully-developed, proven online testing system; and better staffing and management—which DOE believed meant less risk of future problems with the student assessment program.

JLARC on the Web:
http://jlarc.state.va.us

Members of the Joint Legislative Audit and Review Commission

Chairman
Delegate Lacey E. Putney

Vice-Chairman
Senator Thomas K. Norment, Jr.*

Delegate Vincent F. Callahan, Jr.
Senator John H. Chichester
Senator Charles J. Colgan
Delegate M. Kirkland Cox*
Delegate H. Morgan Griffith
Delegate Frank D. Hargrove, Sr.
Delegate Johnny S. Joannou*
Delegate Dwight C. Jones
Delegate Harry J. Parrish
Senator Walter A. Stosch
Delegate Leo C. Wardrup, Jr.*
Senator Martin E. Williams

Mr. Walter J. Kucharski, Auditor of Public Accounts

Director
Philip A. Leone

*Members of the Student Assessment Program Procurement Subcommittee

JLARC Staff for This Report

Bob Rotz, Senior Division Chief
Justin Brown, Project Leader
Greg Rest, Chief Methodologist
Christine Wolfe

COMMONWEALTH of VIRGINIA

Joint Legislative Audit and Review Commission

Philip A. Leone
Director

Suite 1100, General Assembly Building, Capitol Square
Richmond, Virginia 23219

(804) 786-1258

February 1, 2006

The Honorable Lacey E. Putney
Chairman
Joint Legislative Audit and Review Commission
General Assembly Building
Richmond, Virginia 23219

Dear Delegate Putney:

On October 25, 2005, you directed the Joint Legislative Audit and Review Commission staff to review the procurement process used by the Department of Education to award a contract for its student assessment program. The findings of the staff review were presented to the Commission on December 12, 2005. This special report includes the findings of that review, and completes the staff's work on this issue.

Sincerely,

Philip A. Leone
Director

Special Report: Department of Education Student Assessment Program Procurement

In Summary...

In September 2005, the Department of Education (DOE) awarded a six-year, nearly $140 million contract for student testing services, DOE's largest single procurement. Strengths of this procurement process are that it was conducted by a committee of professional staff without outside pressure, the steps in the procurement process were completed, and the selected proposal was below budget. However, a weakness of the process is that the selection committee was not required to document or provide the rationale for its scoring by the procurement act, the Superintendent of Public Instruction, or the Board of Education. Committee members had some difficulty after-the-fact clearly articulating the basis for their scoring, but it appears that the committee chose the vendor it thought had a more clear and detailed proposal; a fully-developed, proven online testing system; and better staffing and management—which DOE believed meant less risk of future program problems for the State.

This special report was conducted at the request of the Chairman of the House Appropriations Committee. On October 21, 2005, the Chairman sent a letter to the Chairman of the Joint Legislative Audit and Review Commission (JLARC). The letter requested that JLARC staff review the Department of Education's (DOE) most recent student assessment program procurement. A copy of this letter is provided in Appendix A of this report. During this review, JLARC staff examined relevant procurement laws and procedures, reviewed file documentation related to the procurement and previous procurements, and conducted 19 interviews with relevant individuals. A list of these interviews is provided in Appendix B of this report.

BACKGROUND ON RECENT STUDENT ASSESSMENT PROGRAM PROCUREMENT

In mid 2004, DOE began the procurement process for its student assessment program. This procurement, according to DOE, had two primary purposes:

1. Secure a vendor to develop, administer, score, and report pencil/paper and online Standards of Learning (SOL) tests for all school divisions and to provide services for the alter-

nate assessment programs for students with disabilities; and

2. Consolidate the statewide student assessment program—previously involving three vendors—into a single contract.

Two companies competed for the contract: Pearson Educational Measurement and Harcourt Assessment, Inc. The procurement process resulted in DOE entering into a contract with Pearson on October 3, 2005. A timeline of important events related to the student assessment program procurement and subsequent JLARC staff review is provided in Figure 1.

Figure 1
Timeline of Key Events in 2004 2005 SOL Procurement Process

Source: JLARC staff analysis.

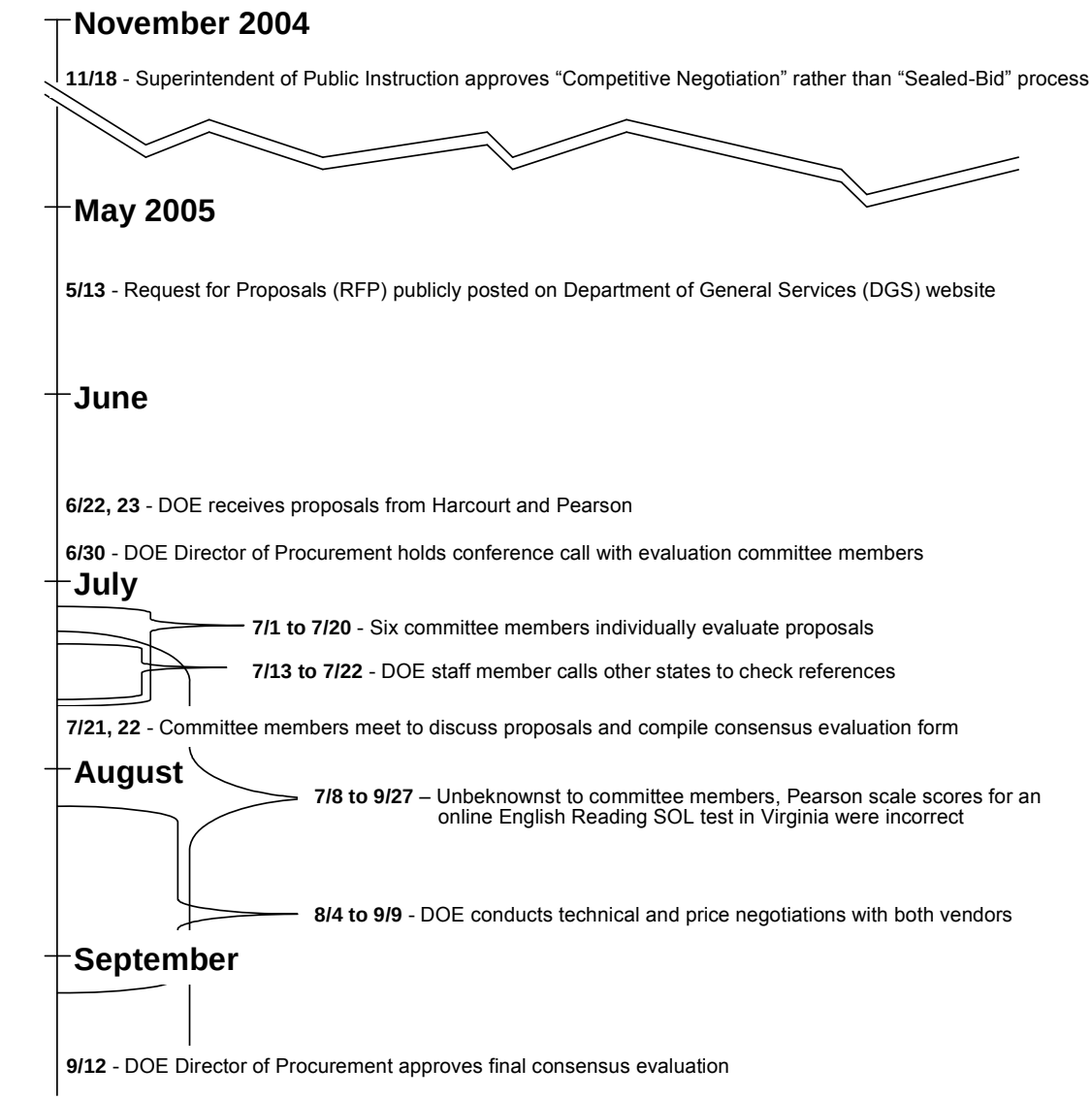


Figure 1 (continued)
Timeline of Key Events in 2004 2005 SOL Procurement Process

Source: JLARC staff analysis.

September 2005 (continued)

9/16
- Memo from DOE Director of Procurement to Superintendent recommends notification of intent to award contract to Pearson
- DOE places intent to award to Pearson on DGS e-procurement system
- Harcourt informed through e-mail that DOE intends to award to Pearson

9/21 - DOE receives letter from Harcourt, DOE treats as protest letter

9/22 - DOE sends letter to Harcourt denying protest

9/27
- Roanoke City asks Pearson to investigate report of scoring error
- Pearson formally notifies DOE of scoring errors affecting 60 student scores
- Pearson begins recalculating scores of affected students and offers $5,000 scholarships to five students not receiving diplomas because of error

October

10/3 - DOE receives signed contract from Pearson

10/17 - DOE officials discuss SOL contract award at House Appropriations Committee meeting

10/18 - *Richmond Times-Dispatch* prints article on SOL scoring mistakes made by Pearson

10/21 - House Appropriations Committee Chairman letter requesting JLARC staff review

10/25
- JLARC Chairman letter authorizing JLARC staff review and appointing special subcommittee (Delegates Wardrup, Cox, Joannou, and Senator Norment)
- JLARC Director sends letter to DOE notifying of JLARC staff review

November

11/14 - Interim JLARC staff briefing to special subcommittee

December

12/12 - Final JLARC staff briefing

Prior Student Assessment Program Procurements

Prior to this most recent procurement, there were two other major SOL procurements. The first of these was in 1996, when DOE initiated a procurement for SOL assessment development, administration, scoring, and reporting services, as well as the adoption of a nationally norm-referenced achieve-

ment test. This procurement was a competitive negotiation process, including evaluation criteria similar to those used during this most recent procurement. Harcourt Brace Educational Measurement was awarded the contract. Until this most recent procurement award, Harcourt continued as the prime vendor for Virginia's testing program, subcontracting some work to Pearson.

In 2000, DOE initiated a procurement for online SOL end-of-course testing. DOE evaluated 11 proposals as part of the online testing procurement and awarded three contracts to demonstrate online testing systems. Of the three vendors who were awarded contracts, DOE renewed only the Pearson contract to continue online testing implementation after the demonstration phase was completed and evaluated. This procurement was also a competitive negotiation process. In addition, DOE signed a contract with Questar Educational Systems in 2001 to provide administration, scoring, and reporting services for alternate assessments of students with disabilities.

JLARC REVIEW FOUND SOME POSITIVE ELEMENTS OF THE STUDENT ASSESSMENT PROGRAM PROCUREMENT

In May 2005, DOE issued a Request for Proposals (RFP) to unify its three existing testing contracts under one large contract. Virginia received two proposals in response to its RFP. These two proposals were from well-known, large vendors with experience in Virginia and other states: Harcourt and Pearson. This appeared to provide DOE with the option of choosing between two vendors that have conducted work for Virginia to create what is generally recognized as one of the more comprehensive testing programs in the nation.

Overall, JLARC staff found three positive aspects of the SOL procurement. First, DOE completed the steps required by Virginia's procurement process. Second, JLARC staff found no evidence of undue influence on the evaluation committee members or process by the Superintendent of Public Instruction or other individuals. Third, it appears that both cost proposals DOE received were below the initial budget estimate for the SOL program.

DOE Completed the Steps Required by Virginia's Procurement Process

In Virginia, the Department of General Services (DGS) "Agency Procurement and Surplus Property Manual" outlines the process that executive branch agencies are to use when procuring goods and services. Table 1 shows each step in the process and whether DOE completed it.

Table 1
DGS Competitive Negotiation Process Steps
Source: JLARC analysis of DGS "Agency Procurement and Surplus Property Manual."

Process Step	Required or Optional	Conducted by DOE During SOL Procurement Process?
Prepare written determination	Required	Yes
Prepare RFP	Required	Yes
Issue RFP	Required	Yes
Select evaluation committee	Required	Yes
Hold pre-proposal conference	Optional	No
Issue RFP addenda	Optional	Yes
Receive proposals	Required	Yes
Review proposals	Required	Yes
Give instructions to evaluation panel	Required	Yes
Hear oral presentations from vendors	Optional	Yes
Evaluate proposals	Required	Yes
Select offerors / Negotiations preparation	Required	Yes
Conduct negotiations, determine offeror to award	Required	Yes
Post notice of intent to award	Optional	Yes
Notify other offerors	Optional	Yes
Review insurance requirement	Required	Yes
Award contract	Required	Yes
Bonds	Required	N/A
Conduct post-award activities	Required	Yes

The remainder of this section briefly describes the major steps in this procurement process and whether DOE fulfilled the requirements.

Determination Was Consistent with Process and Previous Procurement. According to Section 4.22 of the "Agency Procurement and Surplus Property Manual," the work requested by DOE under this procurement is considered a "non-professional service." Consequently, DOE could have used a competitive sealed-bid or a competitive negotiation process. According to Section 7.0 of the procurement manual, "competitive negotiation may be the procurement method used for goods or non-professional services when it is not practicable or fiscally advantageous to use competitive sealed bidding." Further, Section 7.1 of the manual states that "competitive negotiation has the advantage of flexibility for describing in general terms what is being sought and the factors to be used in evaluating responses."

DOE staff indicate that, consistent with the provisions of the manual, they had good reasons for considering it to be not practicable and not fiscally advantageous to use a competitive

sealed-bid. As noted by the DOE Director of Procurement, who served as non-voting chair of the committee, this was a complex procurement, involving dual mode testing (paper and online) as well as the consolidation of three contracts. He noted there were too many variables, making it difficult for DOE to create the detailed specifications necessary for an invitation for bid. While a sealed bid can work for procuring many goods, the situation is different when complex services are to be purchased. DOE staff cited the flexibility offered by competitive negotiation – flexibility to describe the services required in general terms without being as prescriptive as would be required for a competitive sealed-bid process – as a primary reason for using a competitive negotiation. DOE staff also noted that this process allowed DOE to negotiate with vendors on both scope and price. The 2001 and 1996 procurements for student testing were also competitive negotiations.

Process to Develop and Release RFP Appears Appropriate. The State undertook a nearly nine-month process to develop what appears to be a comprehensive, detailed RFP. The RFP document was 146 pages long and included more than 90 separate requirements. DOE staff currently responsible for test development and administration developed and weighted the evaluation criteria. They were advised by the DOE procurement director, and the criteria and weightings are both within the bounds of procurement policy.

Through the DGS website, 106 individuals representing different organizations were notified of the RFP's pending release. DOE also ran an advertisement in the *Richmond Times-Dispatch* on May 8, 2005, publicizing that Virginia was going to release the RFP. Five vendors submitted a statement of intent to submit a proposal. Two of those vendors, Harcourt and Pearson Education Measurement, submitted proposals to DOE.

Evaluation Committee Was Comprised of State and Local Practitioners. According to Annex 7-B (Step 4) of the DGS procurement manual, the "evaluation panel should be kept small, three to five is recommended, including the buyer or some other person knowledgeable of the Virginia Public Procurement Act, the Vendors Manual, and the Agency Procurement and Surplus Property Manual." DOE chose an evaluation committee that included State and local perspectives, and that had background and expertise in test development and both paper/pencil and online aspects of test administration, scoring, and reporting.

The six voting committee members had the following job titles and qualifications:

- Assistant Superintendent, Assessment and Reporting, DOE;
- Team Lead – Test Program Administration, DOE Test Administration and Implementation Division;
- Team Lead – History, DOE Assessment and Development Division;
- Senior Testing Specialist, Virginia Beach Public Schools;
- Supervisor, Assessment and Staff Development, Gloucester Public Schools; and
- Coordinator, Testing and Adequate Yearly Progress, Shenandoah Public Schools.

The committee was chaired by the DOE procurement director, who provided advice about the procurement process and served as a non-voting member of the committee.

Multi-Step Process Was Used to Evaluate Vendor Proposals. Before the evaluations began, the DOE procurement director sent a memo with instructions to the six voting committee members. The memo stated that "all committee members must be present for all evaluation meetings (including oral presentations)." The memo provided a timeline and flow-chart of the process the committee would use. The memo also noted that the committee members "must independently review all proposals for proposal content using provided evaluation forms . . . ." There was also a conference call in which he provided instructions to the committee members. After receiving the instructions, each of the six committee members reported that they independently reviewed the two proposals and completed evaluation sheets provided to them by DOE. After roughly three weeks of individual evaluation, the committee members submitted their scoring sheets to the procurement director.

After this preliminary scoring, the evaluation committee members met in Richmond for two days to discuss their scores and identify questions or areas of clarification that would be sent to the vendors. Following that meeting, and after receiving responses to their questions and conducting negotiations, the evaluation committee members subsequently revised their scores based on the additional information. The preliminary and final scores of the evaluation committee for each criterion are shown in Table 2.

Table 2
Preliminary and Final Evaluation Scoring
Source: JLARC staff analysis of DOE documentation.

	Preliminary		*Final*	
	Harcourt	Pearson	Harcourt	Pearson
Committee Member Scoring				
Qualifications of Offeror	14.0	25.8	13.2	26.5
Test Development	7.3	14.7	17.7	17.0
Administration, Scoring, and Reporting	11.5	17.7	9.8	18.0
Project Management	9.8	13.2	9.5	13.3
Price Formula Scoring				
Price	15.0	10.65	15.0	11.17
TOTALS	**67.67**	**81.98**	**65.17**	**86.0**

Contract Allows State Flexibility and Penalizes Vendors for Mistakes. The period of performance for the contract with Pearson Educational Measurement began on October 1, 2005. The contract is for a period of six years, with four renewal periods of three years for a potential total of 18 years. While concern has been expressed about the length of this contract, it includes several provisions that appear to give the State the flexibility to end the contract. For example, the terms and conditions of the contract state:

- "The purchasing agency reserves the right to cancel and terminate any resulting contract, in part or in whole, without penalty, upon 60 days written notice to the contractor."
- "In case of failure to deliver goods or services in accordance with the contract terms and conditions, the Commonwealth, after due oral or written notice, may procure from other sources and hold the Contractor responsible for any resulting additional purchase and administrative costs. This remedy shall be in addition to any other remedies which the Commonwealth may have."

According to DOE staff and educational literature, mistakes, however unfortunate, do occur in the testing process. When asked about the role that these mistakes (as shown previously in Figure 1) played in their selection decision, committee members said that had the recent scaling mistakes made by Pearson been known during the evaluation process, it would not have changed the result of the process itself. Neither vendor's proposal disclosed previous mistakes, nor did the RFP request them to do so. Both proposals did include quality control plans as requested in the RFP.

When asked why the most recent scaling mistakes made by Pearson were not volunteered to the Appropriations Commit-

tee at its October 17, 2005, meeting, the Assistant Superintendent for Finance who was representing DOE at the meeting said that he was (1) focused on responding to the specific questions asked, and (2) was not aware of the errors that had occurred since they did not occur in his area of responsibility and they were not a topic on which he had been briefed about or had specific knowledge. However, the Assistant Superintendent for Assessment and Reporting who is responsible for the program was at the meeting, but did not address the committee.

To address future mistakes in testing, the new contract includes liquidated damages for poor performance of the vendor, including failure to accurately replicate the scaling and equating of the SOL tests that will result in payment of $5,000 per day for the first ten days of delay, $10,000 per day for days 11-30 of delay, and $15,000 per day beyond 30 days. The contract also stipulates damages for delays in delivering test materials or test scores and for security breaches resulting from contractor negligence.

No Evidence of Undue Influence by Superintendent or Other Individuals

It appears that throughout the process, neither the Superintendent of Public Instruction nor others within DOE or outside the agency influenced the evaluation committee's decision. When asked in separate interviews, all six committee members responded that they felt no pressure from anyone to score the proposals a certain way or reach a certain outcome. The file documentation indicates the superintendent was involved at only two points in the process: signing the request to use a competitive negotiation process and signing the authorization to notify of the intent to award a contract to Pearson.

Both Cost Proposals Were Below the Budget Estimate for the Program

According to an October 26, 2005 letter from DOE to the House Appropriations Committee Staff Director, $31,309,690 from the General Fund has been appropriated for SOL assessment in FY 2006. Based on the assumption that the same amount would be appropriated in the subsequent five years, DOE developed a six-year budget estimate for SOL assessment, which can serve as a benchmark for comparing the six-year cost estimates in the Pearson and Harcourt proposals. The six-year amount was $187.9 million.

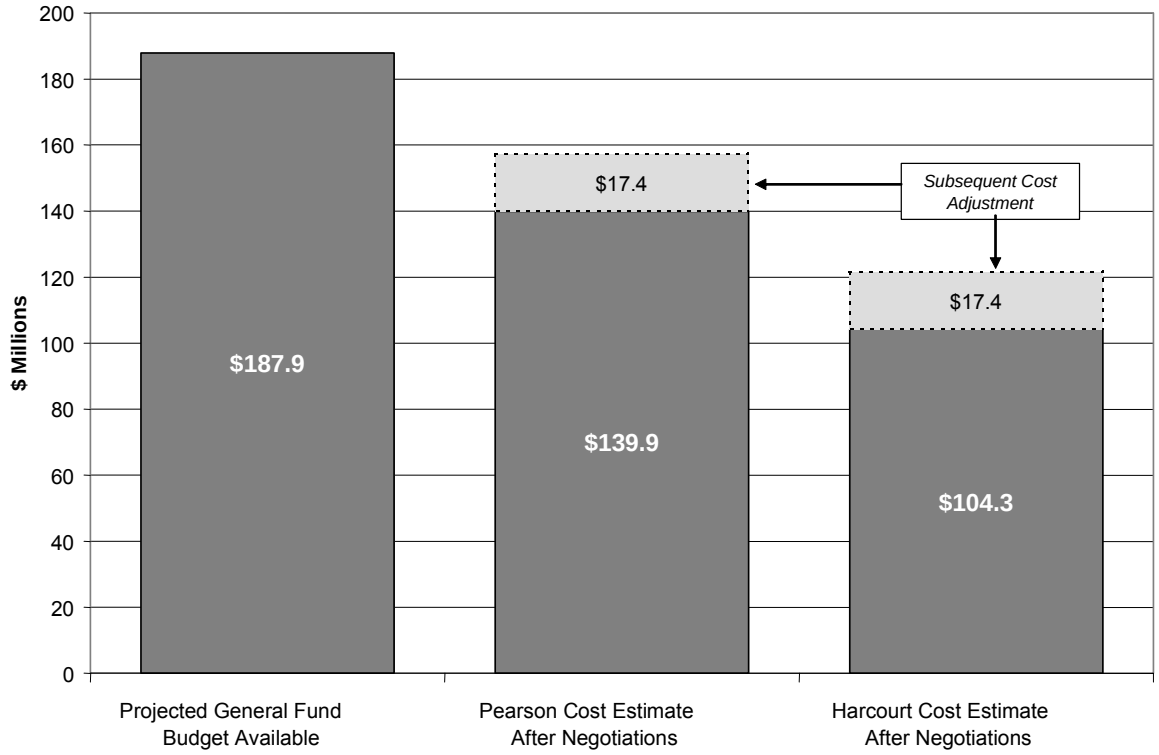
Both the Pearson and Harcourt estimates were substantially below the projected budget estimate. The cost estimate in the Pearson proposal for the six fiscal years totaled $139.9 million, or $48.0 million less than the benchmark General Fund budget estimate. Likewise, the cost estimate in the Harcourt proposal

was $104.3 million, or $83.6 million less than the benchmark General Fund budget estimate and $35.6 million less than the Pearson cost estimate. These comparisons are shown in Figure 2.

As shown in the figure, both the Pearson and Harcourt cost estimates were made at the same stage of negotiation. After the decision to award the contract to Pearson, adjustments to the estimated costs were made. These adjustments are based on projections about the variable portion of the cost proposals based on the number of tests that may be administered in the future, and would apply whether the decision was made to award the contract to Pearson or to Harcourt. These adjustments total $3.4 million for FY 2007 and may total as much as $3.5 million annually for FY 2008 through FY 2011.

Figure 2
Vendor Cost Estimates Compared to Estimated, Projected General Funds (Fiscal Years 2006 through 2011)

Source: JLARC staff analysis of October 26, 2005, DOE letter to Appropriations Committee and procurement files.



ELEMENTS OF THE EVALUATION PROCESS RAISE SOME QUESTIONS ABOUT DOCUMENTATION, THE BASIS FOR THE SCORES GIVEN, AND LEADERSHIP SCRUTINY

Despite the positive elements of the process identified above, JLARC staff identified several areas of concern. First, aside from the consensus scoring sheet, written documentation for

the rationale of the evaluation committee is lacking, and thus, the primary reasons explaining overall why the committee scored one vendor substantially higher than the other are not recorded. Second, for some of the individual evaluative criteria used to score the vendors, some of the verbal rationales offered by the committee after-the-fact are difficult to verify. Third, there appears to have been limited scrutiny and involvement of the Superintendent of Public Instruction and Board of Education concerning the committee's justification for scoring and overall rationale for the selection decision.

Written Documentation Explaining the Overarching Rationale(s) and Outcome of the Process Is Not Available

Annex 7-B (Step 9) of the DGS procurement manual states that "the evaluators are instructed to … prepare a tentative handwritten analysis describing the rationale leading to specific conclusions on the strengths and weaknesses of each proposal."

The 1996 procurement used evaluation sheets, but the file documentation for this procurement also includes handwritten analysis and comments on the scoring sheets.

DOE's Director of Procurement instructed the evaluation committee to electronically complete scoring sheets to document the points assigned to each vendor. These were retained as part of the procurement file. However, the director instructed each committee member to not write on the proposals, but rather make notations and write questions for clarification on Post-It® notes. Consequently, no handwritten analysis describing the strengths and weaknesses of each proposal was created or included in the procurement file. The 1996 procurement used evaluation sheets, but the file documentation for this procurement also includes handwritten analysis and comments on the scoring sheets.

Further, according to Annex 7-B (Step 10) of the procurement manual, "Individual committee member's written comments and points assigned should be included in the procurement file once the consensus evaluation form is prepared to support their consensus." Some evaluation committee members could not recall whether they had written comments, but all mentioned they used the Post-It® notes. Neither written comments nor the Post-It® notes are included in the procurement file.

Recommendation 1. In future procurements, DOE should follow procurement policies and procedures by instructing evaluators to prepare a tentative analysis describing the rationale leading to specific conclusions. DOE should also collect and keep any written comments made by evaluation committee members as part of the file documentation.

At the conclusion of the negotiations and after the final evaluations had been completed, the DOE Director of Procurement sent a memo to the Superintendent of Public Instruction seeking approval to award the contract to Pearson. A copy of this memo is provided in Appendix D. The memo was three paragraphs long, and briefly summarized the evaluation committee's process and noted that the committee was unanimous in its recommendation that the contract be awarded to Pearson.

However, the memo does not provide an overall, summary justification supporting the decision to award the contract to Pearson. The only other documents in the file that provide justification for selecting Pearson are the evaluation scoring sheets. When asked, DOE staff contended that this documentation, along with the entire procurement file, represented the facts supporting the decision to select and negotiate with Pearson. In contrast, the file documentation for the 1996 procurement includes a document, signed by all committee members, summarizing the rationale for the selection of a vendor.

The file documentation for the 1996 procurement includes a document, signed by all committee members, summarizing the rationale for the selection of a vendor.

The lack of file documentation regarding the justification for the award decision meant that JLARC staff had to ascertain much of the DOE rationale for selecting Pearson through interviews with the six committee members. Although scores were assigned by the evaluation committee to each proposal, the file documentation included no supporting information to justify why the scores were provided. To avoid such issues in the future, DGS should provide more specific direction to agencies for these types of procurements about justifying the award decision.

Recommendation 2. The Department of General Services should add language to the "Agency Procurement and Surplus Property Manual" clarifying to agencies that a separate memo detailing the justification for choosing a vendor is a required part of procurement file documentation. When two or more vendors are involved, the memo should specifically focus on the distinguishing characteristics between the two proposals that led to the contract award.

A memo such as the one recommended above would have provided a consolidated document for the superintendent, Board of Education, General Assembly, other vendors, and the public to understand the committee's rationale for awarding a contract to a specific vendor. Such a memo would seem to be especially important in competitive negotiation decisions where many factors need to be considered in making the decision, some of them less quantifiable than others. Without having an understanding of these less quantifiable factors through discussions with the evaluation committee members, it was difficult for JLARC staff to determine from the documentation

alone what additional value the State was receiving by selecting a more expensive vendor.

Two other characteristics of the scoring process, although not currently addressed by procurement policy, make it additionally difficult to independently reconstruct the evaluation committee's scoring rationale for some criteria:

- The sub-criteria within each criterion were not weighted; and
- Committee members made their own judgments about what scores would mean, and assigned scores accordingly. In the absence of notes about their own scoring, they generally did not appear to be able to reconstruct the basis for their scores from memory.

For Some Categories Scored by the Committee, After-the-Fact Rationales for the Scores Assigned Are Difficult to Verify

In light of the lack of written documentation, JLARC staff interviewed evaluation committee members during this review to ascertain the rationale for the scores that committee members assigned to the criteria used to evaluate the RFP submissions. Two months after the completion of the evaluation process, committee members generally had difficulty clearly articulating the basis for some of their scores. According to some members of the committee, the committee chose the vendor that it believed had a more clear and detailed proposal, better project management, and a fully-developed, proven online testing product, which they believed would mean less risk for the State. However, it is difficult to verify some of the after-the-fact rationales offered by committee members for some of the scores. The following is a discussion of the findings from this part of the JLARC staff review.

Rationale for Scoring Pearson Twice as Qualified as Harcourt Is Difficult to Verify. The gap between Harcourt and Pearson on the Qualifications of Offeror criterion was relatively constant between the preliminary and final scoring, with Pearson having between a 12- and 13-point advantage. This was the largest scoring gap of all the evaluation criteria, and was the criterion weighted the most points (30). The general description of this criterion is quite broad; vendors were evaluated on their "qualifications for providing test development, administration, scoring, and reporting services for large-scale high stakes assessment systems in both paper/pencil and online delivery modes." In addition to this general description of the criterion, the RFP provides four sub-criteria that the offerors were to address:

- Related company experience;
- Offeror's ability to provide quality services/deliverables and meet desired schedules in prior contracts of similar scope;
- Related experience and training of assigned personnel; and
- References from other clients for whom the offeror has provided similar services.

Evaluation committee members, when asked about the rationale for the scoring difference, generally were either not able to remember the specific rationale for their scores, or they cited Pearson's experience in other states, particularly their online testing experience. One committee member said that Pearson had a "mature online system" and that Harcourt's system was still in development. Another committee member said that Pearson is a "proven vendor of online testing and high volume paper/pencil testing." One committee member stated that Harcourt's past experience in Virginia had a negative effect on the score that this committee member gave to Harcourt.

According to Annex 7-B (Step 10) of the DGS procurement manual, "Although the scoring of points for other than cost is a matter of subjectivity, the committee member's judgment must be based on fact as presented in the proposal, question/answer conference, etc." However, the committee members' stated rationale for scoring Pearson twice as high as Harcourt on the Qualifications of Offeror criterion was difficult to verify. For example, committee members stated that one of the major reasons for this scoring difference was Pearson's online experience in other states, but it is unclear what information available to the evaluation committee during its deliberations was used as the basis for this scoring. Specifically:

- Pearson's executive summary indicated that in addition to Virginia, it had done online testing in Georgia, Maryland, South Carolina, and Florida. The proposal also noted that it had online pilot programs in Tennessee, Minnesota, Michigan, and Texas. However, neither the summary nor the proposal indicated the number of students tested online in any of the states listed. The summary to the proposal did note that "Virginia continues to be our model state."
- Harcourt's proposal indicated it had done online testing in New York City, Oklahoma, and Mississippi. The proposal indicated that the number of students tested online in these states was between 448 and 47,132.
- Pearson's complete reference list that was provided to DOE during the question and answer phase does not provide any indication of whether the programs referenced were online assessments, and although the number of

programs provided by Pearson is higher than Harcourt, there is no indication of the size of each program.

- The results of the nine reference phone calls made by DOE were provided to committee members after their preliminary scores were developed, and the committee members had already scored Pearson 11.8 points higher than Harcourt at this point. This means that these references could not have been a major reason for the difference in scores.
- Further, the documentation on these reference phone calls does not provide more substantial evidence of Pearson's online testing experience. The reference phone call documentation shows that Pearson had conducted online tests for two of the references, but shows Harcourt had conducted online tests for two references as well. The documentation also provides a mixed picture in general of the qualifications of the two vendors and the other states' satisfaction with the two vendors, and therefore does not provide a rational basis for the major differences in the final scores.

It appears that some evaluation committee members also used this criterion as a "catch-all" category to capture other factors also covered under other evaluation criteria.

It appears that some evaluation committee members also used this criterion as a "catch-all" category to capture other factors also covered under other evaluation criteria. Specifically, when asked to describe the rationale for their scores, committee members often mentioned Pearson's online testing qualifications as the rationale for awarding Pearson more points for Qualifications of Offeror. However, online testing was also given points in the Administration, Scoring, and Reporting category. Appendix C provides more information about the scoring for the Qualifications of Offeror criterion.

Rationale for Test Development Score Appears Reasonable. After the preliminary evaluation, Harcourt scored 2.6 points higher than Pearson on the Test Development criterion. However, committee members raised the Pearson score after getting the answers to questions and receiving a presentation from another company, Educational Testing Service (ETS), which would do the test development under a subcontract relationship with Pearson. This increase brought the two scores nearly even at 17.7 for Harcourt and 17 for Pearson.

Efforts to Assign Scores for Administration, Scoring, and Reporting Were Not Comprehensive. After the preliminary evaluation, Pearson scored 6.5 points higher than Harcourt in this criterion. Evaluation committee members, when asked about the rationale for the scoring difference, cited the fact that Pearson had an existing online system. The DOE committee member with the most technical expertise indicated that certain aspects of the Harcourt system raised "red flags" that led to giving a lower score. These included several security con-

cerns, questions about whether the system would require more bandwidth than would be available at the school divisions, and the use of a third-party browser application.

After the demonstrations of the two online systems, the evaluation committee's final scoring dropped Harcourt nearly two points and slightly raised Pearson's. When asked about the rationale for these changes, the DOE committee member with the most technical expertise indicated that there were several additional security concerns identified during the demonstration that led to Harcourt's score being reduced.

Presentations and demonstrations from the vendors was an optional step of the procurement process. DOE also had the option to choose to only negotiate and have system demonstrations from a single vendor. However, the online component of this evaluation was important enough that DOE chose to have both vendors provide a demonstration of their online systems. The file documentation indicates that during the question and answer process, numerous technical questions were asked about each system, including those related to security, application and infrastructure architecture, and end-user functionality. Unlike the 2000 procurement for online SOL testing, there was no field testing of the proposed systems. According to DOE, this was because in 2000, DOE was unsure of the types of technologies available for conducting online high stakes testing, while this recent procurement was to continue an already established online testing program. DOE believed that given Virginia's experience with online testing, field testing was not required. However, during this recent procurement process, the Harcourt system had not been used in Virginia and was not piloted or field tested in Virginia. This raises questions about the comprehensiveness of the approach used to understand and test Harcourt's system.

Two other aspects of the evaluation of the vendors' online capabilities raise some concern as well. First, the six evaluation committee members gave preliminary scores in the Administration, Scoring, and Reporting criterion before any demonstrations were given of the two online systems—which would be the primary way to administer, score, and report tests. While both proposals did include descriptions of how the online systems would be used for administration, scoring, and reporting, several committee members indicated they had questions about exactly how Harcourt's system would work in Virginia. One evaluation committee member indicated that they would have liked to have learned more about Harcourt's system but did not. This committee member was not invited to the vendor demonstrations.

While evaluation committee members revised their scores after the two demonstrations on a final score sheet, none of the local evaluation committee members actually attended the demonstrations.

While evaluation committee members revised their scores after the two demonstrations on a final score sheet, none of the local evaluation committee members actually attended the demonstrations. Rather, they were advised on a conference call by the DOE staff that did attend the demonstrations about the results. Second, the memo sent by the DOE Director of Procurement to the evaluation committee members providing them instructions about their participation on the committee indicated that "all committee members must be present for all evaluation meetings (including oral presentations)." As stated above, however, the three local members of the committee were not present for either demonstration.

Past Performance of Vendors in Virginia Appears to Have Played a Role in Scores for Project Management. The gap between Harcourt and Pearson on the Project Management criterion was relatively constant between the preliminary and final scoring, with Pearson having roughly a 3.5-point advantage. Evaluation committee members, when asked about the rationale for the scoring difference, cited the fact that Pearson provided more detailed project planning documentation, addressed the interaction among the different work activities, and had staff with more experience doing online testing.

None of the evaluation criteria specifically addressed the past performance of the vendor in Virginia. Further, neither vendor provided Virginia as one of their proposal references. However, it appears that past performance of the incumbent vendor in Virginia may have played a role, although evaluation committee members indicated that, whenever possible, they tried to not consider it. For example, several committee members mentioned that there had been turnover in Harcourt management which had caused difficulties for Virginia in the past. Several committee members also said that they doubted certain aspects of the Harcourt proposal because of previous difficulties Harcourt had in meeting certain deadlines.

Price Score Assigned According to Formula, but Consideration of Cost Beyond Scoring Was Limited. DOE assigned points for the Price criterion according to the formula prescribed in Annex 7-B (Step 10) of the DGS procurement manual. The scoring was done by the DOE Director of Procurement. The price proposals were not provided to the three local committee members, and generally not considered by the full committee in the context of the technical differences between each vendor. After the negotiations process in which Pearson lowered its cost proposal by approximately $10 million and Harcourt lowered its proposal by approximately $2 million, Harcourt had a nearly four-point advantage.

Although the evaluation committee members were told initially not to focus on the cost proposals, when asked, several DOE staff noted several concerns about Harcourt's price proposal. DOE staff noted an apparent inconsistency in Harcourt's price for FY 2007 through FY 2011 when compared to their current contract costs for FY 2006. Under that current contract for FY 2006, Harcourt provided the paper/pencil component of SOL assessment for $26.5 million. However, in their proposal for a new contract for FY 2006 (the contract transition year) through 2011, Harcourt proposed doing this component and more (that is, online and special education testing) in FY 2007 through 2011 for an average of $20.8 million per year. Harcourt's proposal also included minimal transition costs to assume the work being done by Pearson. DOE staff indicated that the $5.7 million difference might be due to either the FY 2007 through 2011 price being underestimated, or the FY 2006 price being higher than necessary. When asked whether they underestimated their costs for the future years, Harcourt staff said they were willing to stand by their cost estimates.

The DOE staff on the committee and the Assistant Superintendent for Finance did discuss the fact that the Pearson proposal was below the budgeted amount for the program. However, the scoring approach prescribed by the procurement manual does not provide a way to consider the true value of the marginal difference between cost proposals. Table 3 demonstrates this by providing hypothetical proposed costs, evaluation points, and the consequent dollar per point difference. As the table shows, the greater the disparity between cost proposals, the higher the dollar value assigned to each point.

Table 3
Hypothetical Illustration of the Difference in Dollars Per Price Point
Source: JLARC analysis.

	Pearson Costs		Difference Between Pearson and Harcourt Proposals		
	Proposed Amount	Points	Points	Dollars	Dollars Per Price Point Difference
Actual	$139,947,359	11.17	3.83	$35,695,610	$9,329,824
	209,921,039	7.45	7.55	105,669,290	13,994,736
Hypothetical	279,894,718	5.59	9.41	175,642,969	18,659,648
	419,842,077	3.72	11.28	315,590,328	27,989,472

As stated above, DOE did score the Price criterion according to the procurement manual. However, this was a mission-critical, high visibility, and high-dollar procurement, and there did not appear to be any attempt by DOE to consider whether the State was willing to pay $9.3 million for each difference in the point score. This raises questions about the approach that Virginia uses to score cost proposals, because without additional consideration of cost beyond the scoring process, there may not be sufficient analysis of the marginal benefit or value the State is receiving for additional cost. However, the way cost is addressed in Virginia's overall procurement process was not in the scope of this review, so JLARC staff are not making recommendations in this area.

Limited Scrutiny and Involvement of the Superintendent of Public Instruction and Board of Education Concerning the Committee's Justification for Scoring and Overall Rationale for the Selection Decision

As noted previously, a positive aspect of this procurement is that a committee of professional staff assessed the proposals, and there is no evidence of any inappropriate outside pressures upon the committee to influence their work. However, as the procurement process concluded, it transitioned from a contract award to a significant program decision. Consequently, obtaining a clear justification from the committee for the benefit of agency leadership, as well as potentially outside oversight and the public, appears to be important.

As such, there is a concern with DOE's limited effort to justify the committee's decision. While the committee's selection was based on the process and criteria set at the time of the RFP, the superintendent, and later the Board of Education, could have been more proactive in ascertaining the underlying basis for the scoring (and hence, the selection itself). This would have provided several benefits: (1) it would have required the committee to state the summary basis for its scoring at a time when their recollections of the process were still fresh; (2) it would have provided some additional assurance that the committee could successfully defend the rationale for its decision to an outside audience; and (3) it would have placed the superintendent and the board in a better position to publicly explain the underlying basis for an important departmental decision.

The Superintendent's Role. The Superintendent of Public Instruction delegated much of the responsibility for this decision to the Assistant Superintendent for Assessment and Reporting and the DOE Director of Procurement. During the interviews with DOE staff for this review, the procurement director recalled that the superintendent did inquire before the conclusion of the process about which vendors submitted pro-

posals. However, he said that he declined to tell her because it could have compromised the process. He said that the superintendent did not press the issue further.

According to the Assistant Superintendent for Finance, in mid-September 2005 there was a meeting at which several DOE staff provided information to the superintendent about the concluded procurement process. According to DOE, this meeting lasted for 45 minutes and was used to brief her on the process and its outcome. However, the discussion remained at a broad level, and did not address the underlying rationale(s) for the committee scoring.

The superintendent indicated that at this meeting, she asked three questions about the decision:

1. Was there a discernible difference between the scores?
2. Are there any "land mines" from the process?
3. Are the proposals realistic relative to the funding?

The superintendent indicated that she received positive answers to her questions, and then she approved the posting of the award notification by signing the memo requesting authorization to award the contract to Pearson (Appendix D). However, these questions were also not sufficient to elicit the underlying rationale(s) for the committee's scoring.

The Board of Education's Role. In this most recent procurement, the current board president recalls being advised early in the process that DOE was going to issue an RFP. The board president also recalls being advised at the conclusion of the process that DOE had made an award decision, but that Harcourt had submitted a letter that was being treated as a protest. The full board never met, however, to discuss the procurement process or the justification for the committee's selection. Furthermore, when asked in general terms, the board president was not able to provide information about the committee's overall rationale or justification for scoring the proposals.

Previous board meeting minutes, DOE documentation, and interviews suggest that the board was more involved in the 1996 procurement process. It appears this was in part because the creation of a testing program was a significant change in State policy, and therefore the board was involved in its role as a policy-making board.

Generally, though, it appears that policy-making boards in Virginia are removed from the procurement process. However, there are examples from other states in which the boards of education appear to be provided more detailed information

about the justification for significant procurement decisions at the conclusion of the process. A recent example is a November 2, 2005, memo sent by the California Department of Education to the California State Board of Education providing background and rationale for its selection of ETS to be the contractor for their Standardized Testing and Reporting program. The memo was 13 pages long and provided a summary of the strengths and weaknesses of each vendor proposal. It also had attachments summarizing evaluation scores and proposed costs, as well as a letter from the State Superintendent of Public Instruction recommending ETS. A similar memo and the one described in this report's Recommendation 2 would serve as a potential vehicle for summarily informing the board, school divisions, oversight organizations, and the public about the justification for important decisions such as this recent procurement.

Appendixes

Appendix A

Study Mandate

This appendix includes the letter requesting that JLARC staff conduct its review, which also includes as an attachment a letter sent to the Superintendent of Public Instruction.

COMMONWEALTH OF VIRGINIA
HOUSE OF DELEGATES
RICHMOND

APPROPRIATIONS COMMITTEE
9TH FLOOR, GENERAL ASSEMBLY BUILDING
CAPITOL SQUARE
POST OFFICE BOX 406
RICHMOND, VIRGINIA 23218
804-698-1590

VINCENT F. CALLAHAN, JR., CHAIRMAN

October 21, 2005

The Honorable Lacey E. Putney
Chairman
Joint Legislative Audit and Review Commission
Suite 1100, General Assembly Building
Richmond, Virginia 23219

Dear Lacey:

Earlier this week I wrote to Dr. Jo Lynne DeMary, Superintendent of Public Instruction, regarding the recently awarded contract for on-line testing services. You will recall that during the Appropriations Committee meeting on Monday, October 17th, concerns were raised regarding the difference in the contract cost between the two vendors and the ability of the successful vendor to accurately score the test.

In my letter to Dr. DeMary, which is enclosed, I shared with her that the testimony given by DOE staff failed to allay concerns raised by the Committee on the "premium" that will be paid for the services provided by NCS Pearson, Inc. within the scope of the contract.

Likewise, I am equally troubled by recent settlements made by Pearson, both in Virginia and the state of Minnesota, for scoring errors that erroneously failed students, resulting in several not graduating.

In light of these issues, I believe there is a compelling need for a complete review of the procurement process undertaken by the Department of Education. As such, I am respectfully requesting that you appoint a special subcommittee to review the process in which this award was made.

This review should include how the Request for Proposal (RFP) was developed, including how the criteria for evaluation and the assignment of point values, specifically for price and test administration and scoring were determined; the process of evaluating the submissions; whether the vendors fully disclosed to the evaluation panel any previous problems in administering or scoring on-line tests; and, whether references were properly contacted and the appropriate due diligence undertaken.

Finally, as a matter of policy, should the State Board of Education have played an active role in developing, evaluating and approving the contract? How does the role of the Board of Education in contract participation differ from other policy making Boards?

In my letter to Dr. DeMary, I requested complete copies of the RFP, any working papers that relate to the RFP, submissions from bidders, scoring and comments, and the final contract language and annual contract costs. This information is to be compiled and forwarded to Robert Vaughn, Director of the Appropriations Committee Staff. Once the materials have been received, Robert will in turn forward the documents to Phil Leone for his staff to begin their review.

In closing, I want to underscore that this review in no way reflects a lack of commitment to the Standards of Learning or the testing needed to ensure proper accountability to the Standards. However, it is imperative that we ensure the integrity of the testing services being provided to our school divisions are above reproach. Thank you in advance for your assistance.

Sincerely,

Vincent F. Callahan, Jr.
Chairman

Cc: The Honorable William J. Howell, Speaker of the House
Members, Appropriations Committee
Dr. Jo Lynne DeMary, State Superintendent for Public Instruction
Mr. Phil Leone, Director, Joint Legislative Audit and Review Commission
Mr. Robert P. Vaughn, Staff Director

Enclosure

COMMONWEALTH OF VIRGINIA

HOUSE OF DELEGATES

RICHMOND

APPROPRIATIONS COMMITTEE
9TH FLOOR, GENERAL ASSEMBLY BUILDING
CAPITOL SQUARE
POST OFFICE BOX 406
RICHMOND, VIRGINIA 23218
804-698-1590

VINCENT F. CALLAHAN, JR., CHAIRMAN

October 19, 2005

Dr. Jo Lynne DeMary
Superintendent of Public Instruction
P.O. Box 2120
Richmond, VA 23218-2120

Dear Jo Lynne:

I am sorry that your schedule did not allow for you to attend Monday's Appropriation's meeting. However, thank you for sending Lan Neugent and Dan Timberlake to present to the Committee the current status of the development of the Department of Education's (DOE) on-line testing initiative.

As you may now know, most of Committee questions and lengthy discussion centered on the recent awarding of the on-line testing contract to NCS Pearson, Inc. that totaled $139.5 million over a 6 year period. Specifically, two key issues were raised: the difference in the contract cost between the two vendors and the ability of the successful vendor to accurately score the test.

Testimony given by DOE staff failed to allay concerns raised by the Committee on the "premium" that will be paid for the services provided by Pearson within the scope of the contract. Quite frankly, there was neither sufficient information readily provided nor a persuasive rationale clearly articulated that would have justified the additional $35 million cost to be borne by Virginia's taxpayers.

Of equal concern was the absence of full disclosure by the DOE staff on the extent of testing and scoring errors encountered by Pearson. In response to questions raised regarding prior test scoring errors, DOE staff only acknowledged that Pearson had erroneously failed 8,000 students in Minnesota, resulting in 50 seniors not graduating.

What was not disclosed was the fact that similar errors have occurred here in Virginia with the SOL test administered by Pearson in the spring of 2005. I was dismayed to learn from a news article in the October 18 *Richmond Times-Dispatch* that approximately 60 Virginia students were told that they too had erroneously failed their SOL tests, resulting in five students not graduating on time. Clearly, given the line of questioning about the ability of Pearson to correctly score the tests, the Department should have been forthcoming and disclosed to the Committee the recent problems in Virginia.

It is also my understanding that Pearson has offered to settle with the students by providing $5,000 scholarships. While it was not clear in the article what the total cost of the settlement would be in Virginia, I understand that the settlement in Minnesota was $7.0 million.

The Committee is concerned that the settlement costs incurred by Pearson may contribute to the difference in the contract costs between the two vendors. The settlements and performance issues underscore the Committee's uneasiness. Likewise, unanswered questions about what, if anything, the Commonwealth's taxpayers would be getting for what strikes many on the Committee as $35 million in unnecessary or unwarranted spending. Clearly, the presentation did nothing to resolve these issues to the Committee's satisfaction.

In light of these several disconcerting issues, there is a compelling need for additional information to be provided to the Committee for closer scrutiny and legislative oversight of the additional contract costs.

I am requesting complete copies of the RFP, any working papers that relate to the RFP, submissions from bidders, process of the award, scoring and comments, the final contract language and annual contract costs along with DOE's plan to cover the additional $35 million contract expenses within existing resources.

If you have any questions regarding this request, please contact Robert Vaughn, Director of the Appropriations Committee Staff. Once the materials have been compiled, please forward them to Robert. Thank you in advance for your timely response to the Committee's concerns.

Sincerely,

Vincent F. Callahan, Jr.
Chairman

Cc: The Honorable William J. Howell, Speaker of the House
Members, Appropriations Committee
Mr. Robert P. Vaughn, Staff Director

Appendix B

Interviews Conducted by JLARC Staff During the Review

Contact	Date	Method: Meeting	Method: Phone
1. Jo Lynne DeMary, Superintendent of Public Instruction	November 3, 2005	x	
2. Harriet Dawson, Supervisor, Assessment and Staff Development, Gloucestor Division of Public Schools	November 4, 2005	x	
3. Barry Kavy, Senior Testing Specialist, Virginia Beach Division of Public Schools	November 7, 2005	x	
4. Deborah Swecker, Coordinator, Testing, Adequate Yearly Progress, Shenandoah Division of Public Schools	November 7, 2005	x	
5. Shelly Loving-Ryder, Assistant Superintendent, Assessment and Reporting, DOE	November 10, 2005	x	
6. Don Marsh, Team Lead – History, DOE Assessment and Development Division	November 10, 2005	x	
7. Sarah Susbury, Team Lead – Test Program Administration, DOE Test Administration and Implementation Division	November 10, 2005	x	
8. Representatives from Harcourt Assessment, Inc.	November 17, 2005	x	
9. Tracie Coleman, Internal Audit Director, DOE	November 17, 2005		x
10. Lolita Hall, Team Lead – Test Implementation, DOE Test Administration and Implementation Division	November 18, 2005	x	
11. Linda Ramthun, Director, Budgeting and Performance Management, Auditor of Public Accounts	November 21, 2005		x
12. Representatives from Pearson Educational Measurement	November 22, 2005	x	
13. Lan Neugent, Assistant Superintendent for Information Technology and Human Resources	November 22, 2005	x	
14. Sharron Hunt, Director of Testing, Georgia DOE	November 23, 2005		x
15. Deborah Love, Office of the Attorney General	November 23, 2005		x
16. Wiley Rowsey, Director of Procurement, DOE	November 28, 2005	x	
17. Dan Timberlake, Assistant Superintendent, Finance, DOE	November 29, 2005	x	
18. Tom Jackson, President, Virginia Board of Education	November 29, 2005	x	
19. Jennifer Stegman, Assistant Superintendent of Accountability and Assessment, Oklahoma DOE	November 29, 2005		x

Appendix C

Additional Information on Offerors' Experience in Other States and References

This appendix provides additional information on each offerors' relevant experience in other states and their references, both of which were part of the Qualifications of Offeror evaluation criterion.

INFORMATION PROVIDED IN PROPOSALS REGARDING EXPERIENCE IN OTHER STATES

The main section of the RFP that addresses the experience of the offerors is Section 7.7: Qualifications of Offeror. This section of the RFP asks for "evidence of experience in the development, administration, scoring and reporting of large-scale, high stakes assessment programs delivered in both online and paper/pencil modes" and "a list of all companies and/or agencies for which work similar to that required by this RFP has been completed, including detailed descriptions of the services performed, and numbers of students tested with paper/pencil and online tests." This section of the RFP also asks for evidence of involvement with the School Interoperability Framework Association (SIFA).

The two offerors responded to the requirements in Section 7.7 slightly differently. Harcourt provided a one-page overview of their experience with the dual administration of assessment programs (online and paper/pencil testing programs), and then provided descriptions of eight assessment programs that they have conducted in seven other states and one large city. For three of these clients, Harcourt administered tests online, and they provided the number of students tested online, as required.

Pearson's response to Section 7.7 was a table that listed four assessment programs conducted in three states. Specific descriptions of the work conducted in each of these three states were included in an appendix to Pearson's proposal. The descriptions provided did not clearly indicate whether the assessment programs were online programs, although one description mentioned an "online testing pilot." Pearson did not provide the number of students tested online for any of the references provided.

Each offeror also provided descriptions of work conducted in other states by their respective subcontractors.

Neither proposal specifically addressed scoring mistakes make by the offerors in Virginia or other states because this was not a requirement in the RFP. The RFP did require each offeror to propose quality control procedures for ensuring the accuracy of score reports, and both offerors included information related to this requirement in their proposals.

ADDITIONAL INFORMATION PROVIDED TO COMMITTEE MEMBERS AFTER THE PRELIMINARY EVALUATION

After the proposals were evaluated and preliminarily scored by the committee members, DOE asked each offeror to provide DOE with a complete list of references for their large-scale testing programs (defined by DOE as programs for which more than 20,000 students were tested). In response to this request, Harcourt provided references for 22 programs in 20 states (one of these programs is in Virginia), four programs in four large cities or counties, and three programs for other clients, for a total of 29 programs. Pearson provided references for 36 programs in 17 states (two of these programs are in Virginia), and seven references for projects where they were the subcontractor ("publisher"), for a total of 43 programs. However, neither Harcourt's nor Pearson's lists provide an indication of the number of students tested or number and type of tests provided, nor whether the programs included an online testing component.

REFERENCE PHONE CALLS MADE BY DOE

As part of the proposal evaluation process, DOE called other states to obtain references for each offeror. Based on the complete client list described above and the references provided in the initial proposals, the Assistant Superintendent for Assessment and Reporting selected four states to call for each offeror, and one state for Pearson's main subcontractor, ETS. In selecting a sample of the references to be contacted, DOE used the following guidelines: 1) states/districts that had large-scale, criterion-referenced testing programs such as Virginia's, and 2) programs whose contracts included handscoring of responses to writing prompts and online delivery of tests. The state of Minnesota, where Pearson had an issue with scoring mistakes in 2000, was not called as a reference for Pearson.

A DOE staff person called each of these states to obtain information on the types of work each offeror performed, whether this work included web-based (online) assessments, how well deadlines were met, and the general level of satisfac-

tion with each vendor (in terms of deliverables, quality control, and types of problems encountered). DOE asked each state a specific question on their level of satisfaction with the quality control procedures used by the contractor, which appears to be an attempt to determine whether the contractor made scoring mistakes in other states. In response to this question, most of the states made general comments on their level of satisfaction with the contractors' quality control procedures. However, two states made more specific comments. One of Pearson's references said there have been "a few incidents that the state was not pleased…" and one of Harcourt's references said that Harcourt "sent a data file for AYP [adequate yearly progress] that was not correct, the state had to completely redo the file."

The DOE staff person who made the phone calls provided handwritten notes on the results of the calls to the committee members during the two-day evaluation meeting in Richmond. This meeting took place after the committee members had developed their preliminary scores, and the committee members had already scored Pearson 11.8 points higher than Harcourt at this point.

The comments from other states were mixed, with each offeror receiving both positive and negative comments from other states. Representative examples of positive comments, taken directly from the DOE staff person's notes, include the following:

<u>Pearson</u>

- "Very responsive; project manager very good with keeping up with deadlines."
- [Web-based delivery] is a "new area for the state, but Pearson has been very helpful and knowledgeable about online testing because of experience in Virginia."
- "The state has been very satisfied with the deliverables from Pearson."

<u>Harcourt</u>

- "Web-based testing has worked very well; very few issues."
- "Very satisfied" with quality control procedures. "The contractor seems to use a high level Q.C. process; very rarely do materials need adjustments when they are reviewed by state staff."
- "Very well satisfied with program manager, staff assigned and national consultant."

Representative examples of negative comments, taken directly from the DOE staff person's notes, include the following:

Pearson

- "Sub-contract arrangement has been difficult; many of the problems seem to be around the subcontractor. Pearson has a very top-down management structure; staff assigned to work with [the] state typically not able to make decisions, [they are] not allowed to work directly with [the] contractor (ETS) which creates lots of problems—miscommunications; procedures not followed as specified in contract, etc."
- "On the issue of online assessment, it has been very challenging. Many of the problems have been in their QC process."
- "Have had a lot of glitches, but Pearson has work[ed] through the glitches."

Harcourt

- "The last two years had problems in the area of program management."
- "Late delivery of score reports in 2003; the state worked with the contractor to improve procedures which did work better."
- Web-based assessment was "successful, but painful; not satisfied with security procedures."

In an interview with JLARC staff, the DOE staff person who made the calls stated that that there were some "positive and negative comments about both." She said comparing the two was difficult because the other states were using the vendors for different things, and had different programs. She said because of this that one "can't make a fair judgment just based on these conversations."

Appendix D

DOE Memo Approving Intent to Award

This appendix includes the memo signed by the Superintendent of Public Instruction approving the posting of notification of intent to award the contract.

COMMONWEALTH of VIRGINIA

DEPARTMENT OF EDUCATION

RICHMOND, 23216

September 16, 2005

MEMORANDUM

TO: Jo Lynne DeMary

FROM: Wiley C. Rowsey

SUBJECT: Recommendation of Contract Award

RE: RFP # DAR2005, dated May 13, 2005, Virginia Statewide Student Assessment Program

Two proposals were received in response to the referenced RFP. The two offerors were Harcourt and Pearson. The evaluation committee, chaired by me as a non-voting member, consisted of three Department of Education staff members: Shelley Loving-Ryder, Sarah Susbury, Don Marsh and three directors of testing from local school divisions: Debbie Swecker, Barry Kavy and Harriet Dawson. Each of the committee members conducted an independent review and scoring of the proposals. A consensus meeting was held and all members were in agreement on the ranking: Pearson as number 1 and Harcourt as number 2.

Although the initial scoring placed the two offerors a number of points apart, the committee decided to negotiate with both. Several discussions were held with each offeror, including demonstrations of the proposed on-line test delivery system for each. After the negotiations were completed, each committee member was furnished with a copy of the results, and they independently re-scored the proposals. A final consensus meeting was held and the ranking remained the same. The committee was unanimous in its recommendation that the contract award be made to Pearson.

Based on the recommendation of the committee and after finalizing some items with Pearson, I respectfully request your approval to proceed with posting a notice of intent to award for a 10 day period and that contract performance commence as of October 1, 2005.

I approve the posting of the notice of intent to award to Pearson and the subsequent contract award.

Jo Lynne DeMary
Dr. Jo Lynne DeMary

Sept. 16, 2005
Date

Appendix E

Agency Response

As part of the validation process, State agencies are given the opportunity to comment on an exposure draft of the report. Appropriate technical corrections resulting from comments provided by DOE have been made in this version of the report. This appendix includes DOE's written response.

DEC - 7 2005

COMMONWEALTH of VIRGINIA

DEPARTMENT OF EDUCATION
P.O. Box 2120
Richmond, Virginia 23218-2120

JO LYNNE DEMARY, Ed.D.
Superintendent of Public Instruction

Office: (804) 225-2023
Fax: (804) 371-2099

December 7, 2005

Mr. Philip A. Leone
Director
Joint Legislative Audit and Review Commission
Suite 1100, General Assembly Building
Richmond, Virginia 23219

Dear Mr. Leone:

The Joint Legislative Audit and Review Commission (JLARC) report on the Department of Education's (DOE) student assessment program procurement confirms that the procedure followed by the department in selecting a vendor for this critical program was consistent with Virginia's procurement process and was not subject to undue influence by any party.

This finding is a credit to the leadership during the process provided by the assistant superintendent for assessment and reporting and the agency's procurement director. These senior DOE officials brought many years of expertise and knowledge to this procurement, and their wisdom and diligence, coupled with the expertise of the local testing directors who served on the evaluation panel, has produced a contract that meets the assessment requirements of Virginia's accountability program for schools and students.

The objectives of the 2005 Request for Proposals (RFP) were to continue the commonwealth's assessment program, consolidate the program under a single contract, and meet the General Assembly's goal of expanding online testing to all tested grades. As your staff documented, Harcourt Assessment, Inc. and Pearson Educational Measurement both submitted proposals in response to the RFP.

The Pearson proposal provided an established Web-based delivery system that incorporated the requirements for the Virginia assessment program. This system also had been implemented successfully in several states, including Virginia. Harcourt's online delivery system, on the other hand, would have required additional development to meet Virginia's needs. This was a crucial distinction, in my opinion, given the maturity of the commonwealth's assessment program.

The six members of the evaluation committee all rated Pearson's response to the RFP as the best overall proposal for providing a high-quality and effective assessment program for the commonwealth's public schools, and achieving the General Assembly's objective of expanding online testing. The members of this committee, which included DOE staff and directors of testing from three school divisions, brought a wealth of experience, insight, and integrity to the evaluation of the competing proposals. They performed a valuable service to the commonwealth in recommending a proposal that will keep Virginia in the forefront of what is now a national effort to raise student achievement by holding schools accountable for student learning and effective instruction.

It is important to note that the evaluation committee members reached their conclusions after independently reviewing the proposals, and then spending two full days in discussion and deliberation. A good deal of time, effort, and consideration was invested by these professionals to reach what they believed was a sound decision for the commonwealth. Given the 20-point gap between the overall scores for the two proposals, the thoroughness of the multi-step procurement process, and my confidence in the expertise of the six members of the committee, I was pleased to accept their recommendation.

Legislative scrutiny of an award of this size is appropriate, and I appreciate the professional manner in which JLARC conducted this review. I thank the members of my staff and the members of the evaluation committee for their cooperation.

Sincerely,

Jo Lynne DeMary

JLD/jt

JLARC Staff

Executive Staff

Philip A. Leone, Director
Glen S. Tittermary, Deputy Director

Division Chiefs

Robert B. Rotz, Senior Division Chief
Harold E. Greer III, Division Chief

Section Managers

Patricia S. Bishop, Fiscal & Administrative Services
Gregory J. Rest, Research Methods
Walter L. Smiley, Fiscal Analysis

Project Leaders

Aris W. Bearse
Ashley S. Colvin
Justin C. Brown
Martha L. Erwin
Eric H. Messick
Nathalie Molliet-Ribet
Kimberly A. Sarte

Project Staff

Janice G. Baab
Jamie S. Bitz
Jennifer Breidenbaugh
M. Angela Coleman
Eileen T. Fleck
Paula C. Lambert
Brad B. Marsh
Ellen J. Miller
Jason W. Powell
Tracey R. Smith
Elisabeth M. Thomson
Christine D. Wolfe
Kent S. Wyatt

Administrative and Research Support Staff

Joan M. Irby
Betsy M. Jackson

Recent JLARC Reports

Special Report: Higher Education, November 2002
Special Report: Medical Supplies and Pharmaceuticals, December 2002
VRS Semi-Annual Investment Report No. 19, December 2002
The Future of the Chesapeake Bay Bridge-Tunnel, January 2003
Review of Information Technology Systems Development, January 2003
Review of the Virginia Birth-Related Neurological Injury Compensation Program, January 2003
Review of Workforce Training in Virginia, January 2003
Review of the Charitable Gaming Commission, January 2003
Implementation of the Chesapeake Bay Preservation Act, January 2003
Special Report: State Spending on Regional Health Planning Agencies, June 2003
2003 Report to the General Assembly, September 2003
Technical Report: State Funding Formula for Educational Technology, September 2003
Review of State Spending: December 2003 Update
Implementation Review: Virginia Information Technologies Agency, December 2003 Status Report
Review of Virginia's Activity in Maximizing Federal Grant Funding, December 2003
Semi-Annual VRS Investment Report No. 21, December 2003
Best Practices for the Support Services of School Divisions, January 2004
Acclimation of Virginia's Foreign-Born Population, January 2004
Review of the State's Passenger Vehicle Fleet, January 2004
Review of Factors and Practices Associated with School Performance in Virginia, January 2004
Benchmarks: Virginia Compared to the Other States, July 2004
Tenure and Post-Tenure Review Policies at Virginia's Public Colleges and Universities, August 2004
Special Report: Impact of Proposed Child Day Care Center Regulations in Virginia, September 2004
Replacing Income Tax Revenues with Sales and Use Tax Revenues, November 2004
Interim Status Report: Impact of Virginia's Aging Population on State Agency Services, November 2004
Review of Emergency Medical Services in Virginia, November 2004
The Use and Financing of Trauma Centers in Virginia, December 2004
Review of State Spending: December 2004 Update
VRS Biennial Status and Semi-Annual Investment Report, No. 23, December 2004
Special Report: State Spending on Standards of Quality Costs, December 2004
Review of Nutrient Management Planning in Virginia, January 2005
Review of Child Protective Services in Virginia, January 2005
Semi-Annual VRS Investment Report # 24, July 2005
Special Report: Certain Personnel Issues at VRS, July 2005
2005 Biennial Report to the General Assembly, September 2005
Review of Homeland Security Funding and Preparedness, September 2005
Operation and Performance of Virginia's Social Services System, October 2005
Assessment of Reimbursement Rates for Medicaid Home and Community-Based Services, October 2005
Review of Land Application of Biosolids in Virginia, November 2005
Interim Report: Impact of Assisted Living Facility Regulations, November 2005
Virginia Compared to the Other States: A Compendium of State Statistics, December 2005
Impact of an Aging Population on State Agencies, January 2006
Impact of an Aging State Workforce, January 2006
Review of State Spending: December 2005 Update, January 2006
Special Report: Department of Education Student Assessment Program Procurement, February 2006

JLARC Home Page: http://jlarc.state.va.us

JLARC
Suite 1100
General Assembly Building
Capitol Square
Richmond, Virginia 23219
(804) 786-1258 Fax: 371-0101
http://jlarc.state.va.us