

**TRANSCRIPT OF PROCEEDINGS  
TESTING INTEGRITY SYMPOSIUM  
02.28.2012**

**EUNICE GREER:** GOOD MORNING. MY NAME IS EUNICE GREER. ON BEHALF OF THE DEPARTMENT OF EDUCATION, I HAVE THE PRIVILEGE OF WELCOMING JOHN EASTON, THE DIRECTOR OF IES, AND INVITING HIM TO KICK US OFF THIS MORNING. JOHN?

**JOHN EASTON:** THANKS VERY MUCH, EUNICE. AS EUNICE JUST SAID, I AM JOHN EASTON, AND I'M THE DIRECTOR OF THE INSTITUTE OF EDUCATION SCIENCES, AND I WANT TO WELCOME YOU ALL AND ESPECIALLY THANK YOU FOR COMING TO PARTICIPATE IN TODAY'S DISCUSSION. I KNOW MANY OF YOU HAVE TRAVELED TO BE WITH US, AND WE APPRECIATE THAT.

THE NATIONAL CENTER FOR EDUCATION STATISTICS IS SPONSORING TODAY'S SERIES OF PANEL DISCUSSIONS THAT FOCUS ON MAINTAINING THE INTEGRITY OF ASSESSING STUDENT ACHIEVEMENT. WE AT THE UNITED STATES DEPARTMENT OF EDUCATION ARE SEEKING TO COLLECT AND SHARE INFORMATION ABOUT BEST PRACTICES THAT HAVE BEEN USED TO PREVENT, DETECT, AND RESPOND TO IRREGULARITIES IN ACADEMIC TESTING.

EDUCATORS, PARENTS, AND THE PUBLIC IN GENERAL RELY ON ACCURATE, RELIABLE, AND TIMELY INFORMATION ON STUDENT ACADEMIC PERFORMANCE TO IMPROVE INSTRUCTION AND HELP ALL STUDENTS REACH AND MAINTAIN HIGH LEVELS OF ACHIEVEMENT. INDEED, THE AVAILABILITY OF VALID, RELIABLE, AND TIMELY DATA ON STUDENT PERFORMANCE IS ESSENTIAL FOR INFORMING INSTRUCTION, IDENTIFYING PROFESSIONAL DEVELOPMENT NEEDS,

HELPING ENSURE MEANINGFUL ACCOUNTABILITY, AND IMPLEMENTING EFFECTIVE EDUCATION REFORMS.

AS A FEDERAL STATISTICAL AGENCY, THE NATIONAL CENTER FOR EDUCATION STATISTICS IS ESPECIALLY CONCERNED WITH ISSUES OF DATA QUALITY AND ACCURACY, AS IS THE ENTIRE INSTITUTE OF EDUCATION SCIENCES AND THE DEPARTMENT OF EDUCATION. SO THIS SYMPOSIUM IS ONE OF SEVERAL STEPS THAT'S BEING TAKEN BY THE DEPARTMENT TO COLLECT INFORMATION AND GATHER SUGGESTIONS TO ASSIST STATE EDUCATION AGENCIES, LOCAL EDUCATION AGENCIES, AND TESTING INTEGRITY FOCUS ORGANIZATIONS THAT SERVICE THEM. THOUGH THE DEPARTMENT IS NOT DEVELOPING TESTING INTEGRITY REGULATIONS, IT ANTICIPATES MAKING USE OF THIS INFORMATION TO FACILITATE FURTHER DIALOGUE AND TO HELP SEAs AND LEAs IDENTIFY, SHARE, AND IMPLEMENT BEST PRACTICES FOR PREVENTING, DETECTING, AND INVESTIGATING IRREGULARITIES IN ACADEMIC TESTING.

MANY STATES AND ORGANIZATIONS, SUCH AS THE COUNCIL OF CHIEF STATE SCHOOL OFFICERS, HAVE DONE WORK TO IDENTIFY AND SHARE BEST PRACTICES AMONG STATES. THIS SYMPOSIUM IS DESIGNED TO BRING TOGETHER EXPERTS WHO CAN ALSO SHARE INFORMATION ABOUT THE LATEST THINKING ON ASSURING THE INTEGRITY OF OUR TESTING PROGRAMS. LAST MONTH, THE DEPARTMENT ISSUED A REQUEST FOR INFORMATION TO COLLECT INPUT FROM THE GENERAL PUBLIC ABOUT THE INTEGRITY OF ACADEMIC TESTING.

THE DEPARTMENT POSED A SERIES OF QUESTIONS, TO WHICH WE INVITED INTERESTED MEMBERS OF THE PUBLIC TO RESPOND. THIS SYMPOSIUM IS DESIGNED TO SHOWCASE EXTERNAL EXPERTS WHO CAN ENGAGE IN FURTHER DISCUSSION AND PROBE THESE ISSUES IN GREATER DEPTH. FOLLOWING THE SYMPOSIUM, THE DEPARTMENT WILL PUBLISH A DOCUMENT SUMMARIZING THE RECOMMENDATIONS THAT WERE DEVELOPED AS A RESULT OF THE REQUEST FOR INFORMATION AND FROM THE SYMPOSIUM, AS WELL AS OTHER RESOURCES IDENTIFIED BY PANELISTS TODAY. I'D LIKE TO NOW TURN THIS OVER TO JACK BUCKLEY.

JACK IS THE COMMISSIONER OF THE NATIONAL CENTER FOR EDUCATION STATISTICS. JACK IS ON LEAVE FROM HIS POSITION AS PROFESSOR OF APPLIED STATISTICS AT NEW YORK UNIVERSITY. HE'S WELL KNOWN FOR HIS RESEARCH ON SCHOOL CHOICE AND ON STATISTICAL METHODS FOR PUBLIC POLICY.

THANK YOU, AND WELCOME AGAIN.

**JACK BUCKLEY:** THANK YOU VERY MUCH, JOHN, AND THANK YOU ALL FOR COMING.

WHETHER YOU'RE HERE IN THE ROOM OR PARTICIPATING VIA WEBCAST, WE APPRECIATE YOUR ATTENTION TODAY, AND WE APPRECIATE THAT YOU'RE

JOINING US TO DISCUSS TESTING INTEGRITY AND TESTING  
IRREGULARITIES.

I WANT TO START BY DEFINING WHAT I MEAN BY THAT, ALTHOUGH I HAVE A  
FEELING THAT THE DEFINITION WILL BE A MOVING TARGET AS THE PANELS  
SPEAK AND AS THE DISCUSSION COMMENCES.

FOR STARTERS, A TESTING IRREGULARITY INCLUDES ANY OCCURRENCE THAT  
MAY INAPPROPRIATELY INFLUENCE A STUDENT'S PERFORMANCE ON AN  
ACADEMIC ASSESSMENT, PROVIDE THE APPEARANCE OF IMPROPRIETY, OR  
OTHERWISE CONSTITUTE A BREACH IN TEST SECURITY OR AN IMPROPER  
ADMINISTRATION OF ACADEMIC TESTING.

TO HELP FOCUS THE DISCUSSION TODAY, WE'VE TRIED TO ORGANIZE THE  
PANELS INTO 4 TOPICS--PREVENTING AND REDUCING TESTING  
IRREGULARITIES IN ACADEMIC ASSESSMENTS, DETECTING AND ANALYZING  
TESTING IRREGULARITIES, REVIEWING AND INVESTIGATING ALLEGED  
TESTING IRREGULARITIES, AND, LOOKING TO THE FUTURE, HOW THESE 3  
ISSUES MIGHT CHANGE FOR ASSESSMENTS DELIVERED ONLINE AND BY  
COMPUTER. FROM OUR OUTSTANDING SET OF PANELISTS TODAY, YOU CAN  
EXPECT TO HEAR ABOUT BEST PRACTICES, SOLUTIONS, ADVICE, TECHNICAL  
INFORMATION, LEGAL, REGULATORY, AND POLICY APPROACHES TO THE  
PROBLEM, AND OTHER LESSONS LEARNED RELATED TO THE PREVENTION,

DETECTION, AND INVESTIGATION OF ALLEGED OR ACTUAL TESTING  
IRREGULARITIES.

AS WE ALL KNOW, STATES EMPLOY A VARIETY OF SAFEGUARD PROCESSES AND  
PROCEDURES TO ENSURE THAT ASSESSMENTS ARE ADMINISTERED UNIFORMLY  
AND SECURELY. THERE ARE SOME IMPORTANT COMMONALITIES THAT WE'VE  
SEEN.

ACROSS STATES, MANUALS USED BY TESTING ADMINISTRATORS ARE  
DEVELOPED FOR THE ADMINISTRATION OF STATE ASSESSMENTS, AND STATES  
TRAIN ANNUALLY ON THEIR TEST SECURITY REQUIREMENTS, COVERING A  
RANGE OF TOPICS, INCLUDING SECURING TEST MATERIALS, HANDLING TEST  
MATERIALS THROUGHOUT THE ADMINISTRATION, AND SECURING MATERIALS  
POST ADMINISTRATION. ANOTHER MAJOR COMPONENT OF THESE TRAINING  
SESSIONS INCLUDES HOW TO HANDLE ANY DEVIATION FROM SPECIFIED  
SECURITY PROCEDURES.

WE FOUND A MAJORITY OF STATES INCLUDE DETAILED INSTRUCTIONS TO  
ADMINISTRATORS ON HOW TO REPORT ANY CONCERNS THAT MAY IMPACT THE  
VALIDITY OF STUDENT PERFORMANCE. IN A NUMBER OF STATES, STAFF  
INVOLVED IN THE ASSESSMENT ARE REQUIRED TO SIGN AN AFFIRMATION OF  
SECURITY, ENSURING THEIR UNDERSTANDING OF PROCESSES AND  
PROCEDURES.

AT THE CONCLUSION OF THE ASSESSMENT, THERE ARE ALSO COMMON SAFEGUARDS EMPLOYED ACROSS THE STATES. THE MAJORITY OF THESE, AS MENTIONED ABOVE, ARE DIRECTED TOWARDS MAINTAINING SECURITY OF THE TESTING ITEMS AND ENSURING THAT TEST SECURITY HAS NOT BEEN COMPROMISED. INCLUDED IN MANY OF THE TRAINING SESSIONS AND MANUALS ARE PROCEDURES FOR SECURING AND SHIPPING MATERIALS TO TEST-SCORING ORGANIZATIONS THAT CONTRACT TO STATE DEPARTMENTS OF EDUCATION.

AS JOHN SAID, EDUCATORS, PARENTS, AND THE PUBLIC IN GENERAL RELY ON ACCURATE, RELIABLE, AND TIMELY INFORMATION ON STUDENT ACADEMIC PERFORMANCE TO IMPROVE INSTRUCTION AND HELP ALL STUDENTS MAINTAIN HIGH LEVELS OF ACHIEVEMENT. STATE EDUCATION AGENCIES AND LOCAL EDUCATION AGENCIES WORK TO ENSURE THE INTEGRITY OF THE DATA THAT THEY USE TO MEASURE STUDENT ACHIEVEMENT AND ENSURE MEANINGFUL EDUCATIONAL ACCOUNTABILITY.

IN ORDER TO OBTAIN THESE DATA, STATES MUST ESTABLISH AND MAINTAIN ASSESSMENT SYSTEMS THAT ARE VALID, RELIABLE, AND CONSISTENT WITH NATIONALLY RECOGNIZED PROFESSIONAL AND TECHNICAL STANDARDS. EVEN THE SLIGHTEST APPEARANCE OF IMPROPRIETY IN THE TEST ADMINISTRATION PROCESS CAN UNDERMINE STATE ACCOUNTABILITY SYSTEMS.

AS WE ALL KNOW, TESTING IRREGULARITIES HAVE OCCURRED THROUGHOUT THE UNITED STATES, IN URBAN, SUBURBAN, AND RURAL SETTINGS, LOW-

INCOME AND AFFLUENT AREAS, AND IN CHARTER AND NON-CHARTER SCHOOLS. TESTING IRREGULARITIES ARE NOT NEW PHENOMENA AND CAN BE TRACED TO THE BEGINNING OF STATE ASSESSMENT PROGRAMS IN THE MID 1800s.

SOMETIMES, TESTING IRREGULARITIES HAVE INVOLVED LARGE-SCALE AND COORDINATED EFFORTS AFFECTING MULTIPLE SCHOOLS AND/OR DISTRICTS WITHIN A STATE, BUT OTHER TIMES, IRREGULARITIES HAVE BEEN ISOLATED INCIDENTS IN SINGLE SCHOOLS. SOME OF THE ALLEGED TESTING IMPROPRIETIES COULD BE DESCRIBED AS TOP-DOWN, INITIATED OR CONDONED BY A SCHOOL OR DISTRICT. IN OTHER CONTEXTS, CHEATING WAS CLEARLY CONFINED TO CASES INVOLVING INDIVIDUAL TEACHERS ACTING INDEPENDENTLY.

IT'S IMPORTANT TO REMEMBER THAT USING ACCURATE METHODS TO DETECT AND VERIFY IRREGULARITIES IS CRITICAL TO PROTECTING THE INTEGRITY OF ASSESSMENTS. AS JOHN MENTIONED, THE PURPOSE OF THIS SYMPOSIUM IS TO SHARE BEST PRACTICES AND PROCEDURES, AS RECOMMENDED BY TESTING EXPERTS AND EDUCATION PRACTITIONERS, TO IMPROVE THE ABILITY OF STATE AND LOCAL EDUCATION AGENCIES TO PREVENT, DETECT, AND RESPOND TO TESTING IRREGULARITIES.

WE HAVE ASSEMBLED 4 PANELS TODAY, MADE UP OF SOME OF THE LEADING EXPERTS IN TESTING INTEGRITY, TO DISCUSS THESE ISSUES. EACH PANEL WILL BEGIN WITH PRESENTATIONS FROM THE PANELISTS, FOLLOWED BY A

MODERATED DISCUSSION AMONG PANELISTS, AND WILL CONCLUDE WITH QUESTION AND ANSWER FROM OTHER PANELISTS, AS WELL AS THE AUDIENCE AND WEBCAST VIEWERS.

AND BEFORE ANY OF OUR DISTINGUISHED PANELISTS SPEAK, I SHOULD SAY THAT THE COMMENTS TODAY REFLECT THE EXPERTISE AND OPINIONS OF THESE PANELISTS, AND NOT THE OFFICIAL POSITION OF THE U.S. DEPARTMENT OF EDUCATION.

THE FIRST 2 PANELS WILL BE FOLLOWED BY A ONE-HOUR BREAK FOR LUNCH. THERE WILL ALSO BE 2 PANELS AFTER LUNCH, WITH A 15-MINUTE BREAK BETWEEN THEM.

NOW I'D LIKE TO INTRODUCE OUR FIRST SPEAKER OF THE DAY.

KATHI KING IS THE MATHEMATICS DEPARTMENT CHAIR AND 12<sup>TH</sup> GRADE MATHEMATICS TEACHER AT MESSALONSKEE HIGH SCHOOL IN OAKLAND, MAINE. SHE HAS OVER 34 YEARS' TEACHING EXPERIENCE IN CALIFORNIA, NEW YORK, AND MAINE.

MS. KING'S TEACHING RESPONSIBILITIES INCLUDE ADVANCED PLACEMENT CALCULUS AT THE AB AND BC LEVEL, AND AS DEPARTMENT CHAIR, SHE'S FOSTERED THE CREATION OF COURSES IN ROBOTICS AND PRE-ENGINEERING. SHE ALSO TEACHES STATISTICS AT THE COLLEGE LEVEL.

MS. KING IS A FORMER MEMBER OF THE NATIONAL ASSESSMENT GOVERNING BOARD, WHICH SETS POLICY FOR THE NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS, OR NAEP. SHE RECEIVED A MASTER'S DEGREE IN COMPUTER TECHNOLOGY EDUCATION FROM THOMAS COLLEGE AND A MASTER'S DEGREE AND BACHELOR'S DEGREE IN MATHEMATICS FROM CALIFORNIA STATE UNIVERSITY, NORTHRIDGE.

MS. KING.

**KATHI KING:** GOOD MORNING. THE LENS I WOULD LIKE TO USE THIS MORNING FOR MY COMMENTS IS THROUGH THE EYES OF A CLASSROOM TEACHER. I'VE BEEN TEACHING FOR CLOSE TO 40 YEARS, IN SCHOOLS OF EVERY TYPE IMAGINABLE, INCLUDING JUNGLE SCHOOLS IN THE U.S. PEACE CORPS, SMALL PRIVATE SCHOOLS, LARGE URBAN SCHOOLS IN THE STATES OF CALIFORNIA, NEW YORK, AND MAINE.

OVER THESE 4 DECADES, I'VE WATCHED TESTING INTEGRITY TRANSFORM FROM HEAVY RELIANCE ON PERSONAL HONOR CODES TO STAYING ONE STEP AHEAD OF STUDENTS WHO WANT TO COMPROMISE THE PROCESS FOR PERSONAL GAIN.

CHEATING IS NOTHING NEW. IT'S ALWAYS BEEN WITH US. IN ORDER TO MAINTAIN TESTING INTEGRITY, I BELIEVE FOCUS SHOULD BE DIRECTED TO 3

MAIN AREAS: DESIGN, VIGILANCE, AND CONSEQUENCES. THESE 3 AREAS OF FOCUS APPLY TO STUDENTS, EDUCATORS, ADMINISTRATORS, AND EDUCATIONAL LEADERSHIP. THE VERY FACT THAT THIS SYMPOSIUM HAS BEEN CALLED IS A REFLECTION OF A RESPONSIBLE COMMUNITY EAGER TO REEVALUATE AND TAKE APPROPRIATE ACTION TO IMPROVE THE TESTING ENVIRONMENT. DESIGN.

YEARS AGO, AS I WAS PASSING BACK THE FIRST TEST IN AN AP CALCULUS CLASS, ONE OF THE STUDENTS RAISED HIS HAND AND HE COMPLAINED THAT HE HAD THE EXACT SAME RESPONSES AS HIS NEIGHBOR, AND SHE GOT A 90 AND HE GOT A 20. I THINK WE'VE ALL MAYBE BEEN THERE. IT SEEMS HE'D NEVER BEEN IN A MATH CLASS WHICH USED MULTIPLE VERSIONS OF A TEST.

GOOD DESIGN DOESN'T COME CHEAP, WHETHER IT'S THE CLASSROOM, THE STATE, OR THE NATION. THE TIME TO CREATE MULTIPLE VERSIONS AND THE TIME TAKEN IN CLASS TO ADDRESS TESTING INTEGRITY ARE VALUABLE MINUTES TAKEN AWAY FROM TEACHING AND LEARNING, BUT THEY'RE ABSOLUTELY NECESSARY. A STUDENT WHO PLANS ON CHEATING LEARNS VERY DIFFERENTLY THAN A STUDENT WHO PLANS ON TAKING OWNERSHIP FOR HIS OR HER LEARNING.

EACH YEAR IN MY SCHOOL, IN MATH AND ENGLISH, WE PRE- AND POST-TEST WITH THE NORTHWEST EVALUATION ASSOCIATION ASSESSMENTS, OR AS WE LIKE TO CALL IT, NWEA. STUDENTS UNDERSTAND THE FACT THAT THEY WILL

BE ANSWERING QUESTIONS TAILORED TO THEIR INDIVIDUAL LEVELS OF ACHIEVEMENT. THERE IS A CERTAIN RESOLVE WHICH EXISTS IN THIS CULTURE OF COMPUTER-ADAPTIVE TESTING WHICH RESULTS IN ABSOLUTELY NO ATTEMPT IN ANY WAY TO CHEAT. ONE ADDITIONAL AND DESIRABLE FEATURE IS THAT SCORES ARE IMMEDIATELY AND ELECTRONICALLY REPORTED AND RECORDED, AND ONLY IF THE STUDENT HAS MET WHAT I CALL THE "HONEST ATTEMPT CRITERIA." SO DESIGN IS KEY.

MY NEXT FOCUS IS VIGILANCE. A SIMILAR STUDENT RESOLVE EXISTS TOWARD CHEATING ON S.A.T. TESTS AT MY SCHOOL. I'VE PROCTORED S.A.T. TESTS OVER THE DECADES, AND STUDENTS ARE WELL AWARE OF THE FACT THAT THERE ARE MULTIPLE VERSIONS. WHEN I PROCTOR S.A.T.s, IT'S IN MY SCHOOL, OFTEN IN MY OWN CLASSROOM, WITH STUDENTS WHO ATTEND MY HIGH SCHOOL. ANONYMITY ISN'T SOMETHING THAT CAN BE USED TO BEAT THE TEST.

THE PROBLEM, AS WE'VE SEEN RECENTLY IN NEW YORK, IS NOT A DESIGN PROBLEM, BUT A VIGILANCE PROBLEM. STUDENTS PAYING AS MUCH AS \$3,500 RELIED ON ANONYMITY TO HAVE SOMEONE ELSE TAKE THE TEST FOR THEM. HIGH STAKES AND THE IMMENSE PRESSURE TO GET INTO SOME OF THE MORE PRESTIGIOUS UNIVERSITIES AND COLLEGES WAS ONE OF THE REPORTED MOTIVATIONS FOR THIS BREACH. AND JUST A QUICK AND GREAT EXAMPLE OF VIGILANCE ARE THE WONDERFUL WEBSITES AVAILABLE TO COMBAT

PLAGIARISM. MY SCHOOL LIBRARIAN IS LIKE A REGULAR SHERLOCK HOLMES WHEN IT COMES TO THIS, AND WE REALLY RELY ON HER.

RECENTLY, VIGILANCE HAS HIT AN EVEN GREATER CHALLENGE, HOWEVER, THAN MONITORING INDIVIDUAL STUDENTS. WITH RELIANCE AT DISTRICT AND STATE LEVELS ON HIGH-STAKES TESTING, EDUCATORS ARE FEELING INTENSE PRESSURES THAT HAVE NEVER EXISTED BEFORE. LET'S NOT FOOL OURSELVES. THERE ARE PRIZES TO BE HAD, PRIZES FOR WHICH TEACHERS, PRINCIPALS, AND EVEN UNIVERSITY ADMINISTRATORS ARE WILLING TO PUT THEIR INTEGRITY ON THE LINE FOR. HOW DID IT EVER GET TO THIS? HOW DID TESTING EVER GAIN PRIORITY OVER TEACHING AND LEARNING?

I THINK ABOUT AN OLYMPIC ATHLETE TRYING TO SHAVE SECONDS OFF HIS TIME. HE MIGHT USE TESTING, EVALUATION, AND ASSESSMENT TO GET A TRUE PICTURE IN AN EFFORT TO IMPROVE. THE DIFFERENCE IS THE DESIRE TO USE THE INFORMATION TO GET BETTER. WHY CAN'T ACADEMIC TESTING BE LIKE THAT? WHY CAN'T AN ACADEMIC TEST BE A TOOL WHICH WE CAN USE TO IMPROVE?

WE NEED TO PUT LEARNING BACK INTO THE TESTING PROCESS BECAUSE LEARNING IS THE BUSINESS OF EDUCATION, NOT TESTING. YEARS AGO, I RECALL READING "FREAKONOMICS" BY THE ECONOMIST STEVEN LEVITT, AND MAYBE I'M JUST NAIVE, BUT READING THE DETAILS OF TEACHERS ERASING AND REPLACING ANSWERS WAS BEYOND BELIEF.

I FIND GREAT CONSOLATION IN THE FACT THAT SOFTWARE EXISTS TODAY TO IDENTIFY THIS TYPE OF CHEATING. WE DON'T ALWAYS GET LUCKY TO CATCH OFFENDERS RED-HANDED, AND SO SAFEGUARDS MUST BE BUILT INTO THE TESTING PROCESS. AS A FORMER MEMBER OF THE NATIONAL ASSESSMENT GOVERNING BOARD AND CHAIR OF THE ASSESSMENT DEVELOPMENT COMMITTEE, I BECAME FAMILIAR WITH NAEP TESTING FOR THE NATION'S REPORT CARD. THE STRATEGY OF BRINGING IN INDEPENDENT TEST ADMINISTRATORS IS ANOTHER WAY TO ELIMINATE THE TEMPTATION TO ALTER TEST RESULTS AND THUS PROMOTE TESTING INTEGRITY. VIGILANCE IS KEY.

MY FINAL COMMENTS ARE ON CONSEQUENCES. EACH YEAR WHEN I BRIEF MY AP STUDENTS BEFORE THE TEST, I EXPLAIN TO THEM THAT IF THEY CHEAT, THEY PUT THE SCHOOL AT RISK FOR NOT BEING ABLE TO OFFER AP EXAMS. THE CONSEQUENCES ARE STRINGENT ENOUGH TO BE A DETERRENT, AND STUDENTS PAY ATTENTION. IF THE STAKES ARE HIGH, THEN THE PUNISHMENTS MUST BE COMMENSURATE.

PEOPLE MAKE MISTAKES, AND WE ALL LEARN AND GROW FROM THEM, BUT WE'RE TALKING ABOUT COMPROMISING TESTING INTEGRITY FOR PERSONAL GAIN. WE NEED TO WORK HARD TO GET AN ACCURATE PICTURE OF JUST WHAT WE'RE DEALING WITH AND THEN DEVISE STRATEGIES TO ADDRESS IT.

THE GOOD NEWS IS THAT INCIDENCES ARE COMING TO LIGHT AND BEING DEALT WITH. I WOULD LIKE TO COMMEND THE CITY OF ATLANTA FOR CONFRONTING THE TEST FRAUD WHICH OCCURRED THERE. IT WAS A DIFFICULT AND PAINFUL PROCESS WHICH MADE IT CLEAR THAT THIS TYPE OF BEHAVIOR CANNOT BE TOLERATED. DEPENDING ON THE CONSEQUENCES, THIS INCIDENT CAN SERVE AS A DETERRENT AND CONTRIBUTE TO A CHANGE OF ATTITUDE. I'M NOT HERE TO POINT FINGERS, BUT ACCORDING TO MOST OF THE SURVEY DATA I'VE SEEN, THE CASUAL ATTITUDE TOWARD CHEATING IS PERVASIVE AMONGST STUDENTS, EDUCATORS, PARENTS, AND COMMUNITY MEMBERS.

I WOULD LIKE TO CONCLUDE BY SAYING THAT THE TIME SPENT TODAY HERE IS ABSOLUTELY ESSENTIAL TO MY JOB AS A CLASSROOM TEACHER. GOOD STUDENTS MAKE GOOD CHOICES EVERY DAY, AND THEY ARE DEPENDING ON ALL OF US TO SAFEGUARD TESTING INTEGRITY SO THAT THE REWARDS FOR CHEATING AND THE REWARDS FOR WORKING HARD CAN NEVER BE THE SAME. THE PRESENCE OF SMARTER BALANCE AND PARCC HERE TODAY GIVE ME HOPE THAT THE NEW ERA OF TESTING WHICH IS JUST AROUND THE CORNER WILL PROVIDE CLASSROOM TEACHERS WITH TOOLS WHICH WILL HELP US GET BACK TO THE BUSINESS OF TEACHING AND LEARNING.

THANK YOU SO MUCH, AND I HOPE MY COMMENTS WILL BE HELPFUL IN OUR WORK TODAY.

[APPLAUSE]

**JACK BUCKLEY:** THANK YOU VERY MUCH, KATHI.

NOW IT'S MY DISTINCT PLEASURE TO WELCOME KAYA HENDERSON, CURRENTLY THE CHANCELLOR OF THE DISTRICT OF COLUMBIA PUBLIC SCHOOLS. HER EDUCATION CAREER BEGAN AS A MIDDLE SCHOOL SPANISH TEACHER IN THE SOUTH BRONX, AND SHE SPENT HER SUMMERS OVERSEEING THE PROFESSIONAL DEVELOPMENT OF NEW TEACHERS AT SUMMER INSTITUTES WITH TEACH FOR AMERICA. SHE ALSO SERVED TEACH FOR AMERICA AS A RECRUITER AND NATIONAL DIRECTOR OF ADMISSIONS. IN 1997, SHE BECAME THE EXECUTIVE DIRECTOR OF TFA-DC.

SHE CAME TO DCPS AS DEPUTY CHANCELLOR IN 2007, WHERE SHE LED THE DISTRICT'S HUMAN CAPITAL EFFORTS AND SERVED AS CHIEF NEGOTIATOR FOR THE 2010 CONTRACT BETWEEN DCPS AND THE WASHINGTON TEACHERS' UNION.

HENDERSON'S TEAM ALSO LED THE DEVELOPMENT OF IMPACT, THE DISTRICT'S NEW TEACHER ASSESSMENT SYSTEM.

SHE RECEIVED HER BACHELOR'S DEGREE IN INTERNATIONAL RELATIONS FROM GEORGETOWN SCHOOL OF FOREIGN SERVICE AND HER MASTER OF ARTS IN LEADERSHIP ALSO FROM GEORGETOWN UNIVERSITY. I'LL TURN IT OVER TO KAYA.

**KAYA HENDERSON:** GOOD MORNING.

IN MARCH OF 2011, LESS THAN 6 MONTHS AFTER I WAS NAMED INTERIM CHANCELLOR FOR DC PUBLIC SCHOOLS, "USA TODAY" PRINTED A STORY WHICH ALLEGED THAT THERE WERE SIGNIFICANT TESTING IMPROPRIETIES AT DC PUBLIC SCHOOLS.

UPON LEARNING OF THESE ALLEGATIONS, MY SOLE INTEREST WAS IN DETERMINING WHAT ACTUALLY HAPPENED AT OUR SCHOOLS AND HOW COULD WE ENSURE THAT PARENTS, TEACHERS, STUDENTS, AND POLICYMAKERS WOULD HAVE CONFIDENCE IN OUR RESULTS? THIS SEEMINGLY SIMPLE BUT CRITICAL DESIRE TURNED OUT TO BE MUCH MORE DIFFICULT THAN I ANTICIPATED.

OVER THE PAST DECADE, WITH THE PASSAGE OF NO CHILD LEFT BEHIND, TEST SCORES HAVE PLAYED A CRITICAL ROLE FOR BOTH SUPPORTERS AND CRITICS OF OUTCOME-BASED EDUCATION REFORM. WE USE TESTS TO DETERMINE OUR STUDENTS' PROGRESS, WE USE TESTS TO CHART IMPROVEMENTS IN SCHOOLS AND DISTRICTS, WE USE TESTS AS ONE ELEMENT OF OUR EVALUATION OF TEACHER EFFECTIVENESS, AND PARENTS USE TEST SCORES TO MAKE DECISIONS ABOUT WHERE THEY WANT TO SEND THEIR STUDENTS TO SCHOOL.

BECAUSE EDUCATION QUALITY IS MULTIFACETED AND DIFFICULT TO EXPLAIN, WE FREQUENTLY--PROBABLY TOO FREQUENTLY--USE TEST SCORES AS OUR ONE PROXY FOR DISCUSSING REAL EDUCATION QUALITY. OUR USE OF TEST SCORES AND DATA HAVE HELPED US ADVANCE A MEANINGFUL CONVERSATION ABOUT WHAT IT MEANS FOR SCHOOLS TO DO WELL. WE NOW FOCUS ON ISSUES RELATED TO THE BLACK/WHITE ACHIEVEMENT GAP WITH GREATER FOCUS, WE CAN ZERO IN ON THE PERFORMANCE OF STUDENTS WITH SPECIAL NEEDS AS WE WORK TO ENSURE THAT ALL OF OUR STUDENTS RECEIVE A QUALITY EDUCATION, AND WE CAN TARGET DIFFERENTIATED INTERVENTIONS TO STUDENTS WHO ARE WELL BEHIND IN THEIR PROGRESS WHEN COMPARED TO THEIR PEERS.

WE'VE CREATED AN EDUCATION ENVIRONMENT IN WHICH IT IS NOW THE NORM TO TALK ABOUT STUDENT ACHIEVEMENT DATA. EVEN DC CITY COUNCILMEMBERS WHO ONCE SPOKE ONLY OF PAROCHIAL ISSUES RELATED TO THEIR SCHOOLS NOW QUIZ ME REPEATEDLY ON SCHOOL PERFORMANCE DATA. BUT WE FAIL DEEPLY IN EXPLAINING TO PARENTS AND THE PUBLIC WHAT TEST SCORES MEAN, HOW THEY SHOULD BE USED, AND HOW THEY COULD BE INTERPRETED.

PRINCIPALS OFTEN VIEW TEST SCORES AS THE MEASURE--NOT ONE MEASURE-- IN HOW THEIR SCHOOL IS DOING. WE FAIL TO EXPLAIN WHAT PROGRESS, AS OPPOSED TO ABSOLUTE SCORES, MEANS FOR A SCHOOL, AND WE'VE CREATED A CULTURE IN WHICH STUDENTS, TEACHERS, AND PRINCIPALS FEAR

THEIR END-OF-YEAR TESTS, RATHER THAN LOOKING AT THEM AS AN OPPORTUNITY TO DEMONSTRATE THEIR PROGRESS AND IDENTIFY AREAS FOR IMPROVEMENT.

WHILE I WAS FAMILIAR WITH THESE CHALLENGES FROM MY YEARS OF WORKING IN URBAN EDUCATION, I DIDN'T REALIZE THAT WE HAD ALSO FAILED TO ESTABLISH CLEAR MEANS OF DETERMINING WHAT RELIABLE TEST SCORES ARE, WE FAILED TO CREATED ANY NATIONALLY ACCEPTED STATISTICAL ANALYSIS TO DETERMINE WHEN SCORES ARE SUSPECT, AND WE FAILED TO PROPOSE A RELIABLE INVESTIGATIVE PLAN WHEN RESULTS ARE CALLED INTO QUESTION. WE'VE DONE THE HARD WORK TO ENSURE THAT WE COULD HAVE A DISCUSSION OF EDUCATION OUTCOMES THAT WOULD BE INFORMED BY TEST SCORES, BUT WE'VE NOT APPLIED EQUAL EFFORTS TO HELPING POLICYMAKERS OR OUR COMMUNITY UNDERSTAND WHAT TEST SCORES ARE FOR, WHAT THEY MEAN, AND HOW WE CAN MAKE SURE THAT THEY'RE RELIABLE.

UPON LEARNING OF THE IMPROPRIETIES ALLEGED BY "USA TODAY," I QUICKLY DOVE INTO THE WORLD OF TEST INTEGRITY. I FOUND A DIZZYING WORLD OF DATA ANALYSIS, NON-STANDARDIZED APPROACHES, AND UNRELIABLE OUTCOMES, NONE OF WHICH SEEMED DIRECTED AT HELPING DISTRICTS OR COMMUNITIES GAIN CONFIDENCE IN THEIR OUTCOMES--MOST OF WHICH, IN FACT, PUT DISTRICTS IN A DEFENSIVE POSTURE.

BY ANY REASONABLE MEASURE, DCPS HAD DONE EVERYTHING THAT WE WERE SUPPOSED TO DO TO ADDRESS ALLEGED TESTING IMPROPRIETY.

WE HAD A RIGOROUS OVERSIGHT SYSTEM FOR ADMINISTERING TESTS THAT ENSURED THAT CENTRAL OFFICE STAFF MEMBERS AND OTHER THIRD PARTIES WERE PRESENT TO OBSERVE TESTING.

WE HAD A CONFIDENTIAL TIP LINE WHICH ALLOWED CONCERNED PARTIES TO REPORT WRONGDOING.

WE INVESTIGATED EVERY SINGLE ALLEGATION OF IMPROPRIETY THAT WE RECEIVED, INCLUDING BOTH THOSE REPORTED BY INDEPENDENT SOURCES IN SCHOOLS AND IN THE COMMUNITY AND THOSE IDENTIFIED FOR US BY OUR OFFICE OF THE STATE SUPERINTENDENT OF EDUCATION.

WE USED AN OUTSIDE CONTRACTOR TO CONDUCT INVESTIGATIONS IN CLASSROOMS WHERE THE STATE IDENTIFIED CONCERNS, AND WHEN WE IDENTIFIED THE FEW STAFF WHO HAD COMPROMISED TEST INTEGRITY, WE ACTED SWIFTLY TO PUNISH AND, WHEN APPROPRIATE, TO TERMINATE THE STAFF WHO WERE INVOLVED.

AT THE SAME TIME, IT WAS EASY SPORT FOR THE PRESS TO PLAY THE "WHAT MORE COULD BE DONE" GAME. COULDN'T WE HAVE LOOKED AT RIGHT-TO-WRONG ERASURES IN MORE SCHOOLS? WASN'T THERE MORE THAT OUR VENDOR COULD

HAVE DONE TO IDENTIFY INCONSISTENCIES? COULDN'T INVESTIGATIONS BE MORE THOROUGH? CAN'T WE RELEASE THE INFORMATION TO THE PRESS SO THAT THE PRESS AND THE PUBLIC COULD DECIDE FOR THEMSELVES?

AND THE ANSWER TO ALL OF THOSE QUESTIONS, OF COURSE, IS YES. THERE ARE INNUMERABLE ADDITIONAL THINGS THAT WE COULD HAVE DONE, BUT THERE WAS NO REASON TO BELIEVE THAT ANY OF THESE ACTIONS WOULD HAVE YIELDED MORE RELIABLE RESULTS OR MORE ACCURATE RESULTS. IT WAS CLEAR TO ME THAT IT WOULD BE VERY EASY FOR A DISTRICT LIKE OURS TO FALL DOWN A RABBIT HOLE OF TESTING INVESTIGATIONS ONLY TO FIND OUT THAT BECAUSE THERE ARE NO WIDELY ACCEPTED STANDARDS, THERE'S NO AGREED-UPON RESULT THAT WOULD HAVE SATISFIED THE PRESS OR THE PUBLIC.

WE HAVE A HUGE NUMBER OF TOOLS THAT WE CAN USE TO DETERMINE IF WE THINK THERE HAS BEEN IMPROPRIETY IN A TEST ADMINISTRATION. WE CAN LOOK AT RIGHT-TO-WRONG ERASURES, THE MOST POPULAR AND EASIEST TO UNDERSTAND. WE LOOK AT ITEM PATTERN ANALYSIS, STUDENT GROWTH ANALYSIS, CHANGE FROM THE PREDICTED OUTCOMES OF THE INTERIM ASSESSMENTS, STUDENT PERFORMANCE COMPARED TO RESULTS FROM OUR OBSERVATIONS OF TEACHERS, AND AN EVALUATION OF THE STUDENTS WHO CROSS SIGNIFICANT THRESHOLDS-- FOR EXAMPLE, FROM BELOW BASIC TO BASIC OR FROM BASIC TO PROFICIENT.

WHILE THESE TOOLS ARE ALL USEFUL AND MAKE INTUITIVE SENSE, THERE'S NO CLEAR STANDARD FOR USING THESE MEASURES. SHOULD WE INVESTIGATE EVERY CLASSROOM WITH AN ABOVE-AVERAGE LEVEL OF ERASURES, KNOWING THAT THIS WOULD FORCE US TO INVESTIGATE HALF OF ALL OF OUR CLASSROOMS AND WOULD ONLY CATCH ONE FORM OF CHEATING?

SHOULD WE INVESTIGATE CLASSROOMS WHERE THERE'S A HIGH LEVEL OF GROWTH, KNOWING THAT THIS WOULD PLACE A HUGE BURDEN ON OUR ABSOLUTELY HIGHEST PERFORMING TEACHERS, THE VERY ONES WE WANT TO REWARD AND KEEP IN OUR CLASSROOMS?

IF WE'RE GOING TO USE ANY OF THESE MEASURES, HOW MANY STANDARD DEVIATIONS ABOVE THE MEAN DOES ONE NEED TO BE TO WARRANT AN INVESTIGATION?

AND SHOULD WE USE SOME COMBINATION OF FACTORS, KNOWING THAT THIS WILL INVOLVE DOZENS, IF NOT HUNDREDS, OF DECISIONS, EACH OF WHICH MAY LIMIT THE INVESTIGATION AND EACH OF WHICH MAKES EASY FODDER FOR THOSE WHO WOULD LIKE TO TEAR DOWN OUTCOME-BASED EDUCATION REFORM?

AND WHAT SHOULD INVESTIGATIONS LOOK LIKE? SHOULD WE INTERVIEW STUDENTS? AT WHAT AGE? HOW MANY? GIVEN THAT TEST RESULTS ARE NOT AVAILABLE UNTIL MONTHS AFTER ADMINISTRATION AND THAT INFORMATION REGARDING POTENTIAL IMPROPRIETY MIGHT NOT BE AVAILABLE FOR MANY

MORE MONTHS, CAN WE RELY ONLY ON EYEWITNESSES? CAN WE INVESTIGATE WITHOUT ASKING LEADING QUESTIONS?

BECAUSE THERE IS NO STANDARD, EITHER FOR IDENTIFYING POTENTIAL WRONGDOING OR FOR INVESTIGATING ONCE CHEATING IS ALLEGED, WE ARE LEFT WITH A FUZZY PICTURE OF WHAT RELIABLE OUTCOMES ARE. AS A RESULT, WE DO 3 THINGS THAT ARE VERY BAD FOR CHILDREN.

WE TURN TESTS, WHICH ARE IN PLACE SO THAT WE CAN CELEBRATE STUDENT SUCCESSES AND ADDRESS OUR AREAS OF CONCERN, INTO GAMES FOR ADULTS. WE UNDERMINE OUR VERY BEST TEACHERS BY CALLING INTO QUESTION THEIR GOOD WORK AND GIVING THEM NO OPPORTUNITY TO DEFEND THEMSELVES AND CLEAR THEIR NAME. AND WE ALLOW CRITICS OF OUTCOME-BASED REFORM TO UNDERMINE OUR EFFORTS TO ENSURE THAT WE IDENTIFY THE NEEDS OF EACH AND EVERY STUDENT AND TO WORK TO ENSURE THAT EACH STUDENT ACHIEVES PROFICIENCY NOT BECAUSE IT'S A GOAL, BUT BECAUSE IT'S A BASELINE.

TOGETHER, WE MUST ESTABLISH A CLEAR SET OF FACTORS THAT CAN BE APPLIED NATIONALLY TO DETERMINE WHEN TESTING IMPROPRIETIES OCCUR, HOW TO CONDUCT CONSISTENT STATISTICAL ANALYSIS TO HIGHLIGHT POTENTIAL WRONGDOING, AND HOW TO INVESTIGATE CLASSROOMS WHERE ANALYSIS RAISES QUESTIONS.

IN YESTERDAY'S "NEW YORK TIMES," MICHAEL WINERIP SAID HE WAS DISAPPOINTED THAT I WAS INVITED TO SPEAK AT THIS CONFERENCE. MORE GENERALLY, THE PRESS HAS FRAMED THE CHALLENGE OF TEST INTEGRITY AS THEIR STRUGGLE TO OUTHWIT DECEITFUL SCHOOL DISTRICTS. THIS APPROACH IS EXACTLY WRONG.

SCHOOL DISTRICTS LIKE OURS STRUGGLE WITH THESE ISSUES CONSTANTLY. WE DON'T STRUGGLE TO HIDE CHALLENGES OR TO CONCEAL WRONGDOING. WE STRUGGLE TO ENSURE THAT OUR TEST RESULTS ARE RELIABLE AND TRUSTWORTHY. WE STRUGGLE TO MAKE SURE WE KNOW HOW OUR STUDENTS ARE DOING SO WE CAN INFORM PARENTS, PROVIDE INTERVENTIONS, AND MAKE CORRECTIONS.

TOGETHER, DCPS AND OUR STATE SUPERINTENDENT HAVE DEVELOPED A STRONG APPROACH TO IDENTIFYING POTENTIAL INCONSISTENCIES AND INVESTIGATING THESE CASES, AND WE ARE HAPPY TO USE THIS AS A STARTING POINT FOR OUR NATIONAL CONVERSATION.

HOWEVER, WE KNOW THAT ABSENT A NATIONAL STANDARD FOR THIS WORK, OUR ANALYSIS WILL ALWAYS BE SUBJECT TO SECOND-GUESSING. NO SCHOOL DISTRICT, NO TESTING COMPANY HAS BEEN ABLE TO PROVIDE A PROTOCOL FOR ENSURING THAT THERE IS INTEGRITY TO OUR TEST RESULTS, AND SO I CHOSE TO STEP FORWARD AND ASK THE U.S. DEPARTMENT OF EDUCATION FOR GUIDANCE.

THE DEPARTMENT HAS BROUGHT YOU TOGETHER, A GROUP OF EXPERTS, TO HELP TACKLE THIS CHALLENGING ISSUE. ALONE, MY SCHOOL DISTRICT CANNOT HIRE A TEAM OF STATISTICIANS TO DEVELOP A RIGOROUS, NATIONALLY ACCEPTED INVESTIGATIVE STRATEGY. BY ITSELF, OUR STATE CANNOT DEVELOP A PROTOCOL FOR INVESTIGATIONS THAT WILL WITHSTAND SCRUTINY.

ALTHOUGH WE STRUGGLE WITH THE CHALLENGE OF VALIDATING OUR TEST RESULTS, NO ONE SCHOOL DISTRICT OR STATE CAN ARRIVE AT A VIABLE SOLUTION BY ITSELF. AND SO TOGETHER, WE FACE A CHALLENGE AND AN OPPORTUNITY. OUR CHALLENGE IS THAT WE MUST COME TOGETHER TO RECOGNIZE THAT THIS IS NOT A BATTLE OF DISTRICTS VERSUS THE PRESS. IT SHOULD BE OUR SHARED CHALLENGE TO VALIDATE TEST RESULTS SO WE CAN FOCUS ON THE REAL ISSUE OF IMPROVING EDUCATIONAL OPPORTUNITIES.

WE ALSO HAVE AN OPPORTUNITY FOR YOU TO PROVIDE A UNIVERSALLY ACCEPTABLE NATIONAL STANDARD FOR TEST INTEGRITY. TOGETHER, WE HAVE TO ESTABLISH A NATIONAL CONSORTIUM THAT WILL BRING TOGETHER THE EXPERTISE OF STATISTICIANS, THE KNOWLEDGE OF TESTING EXPERTS, AND THE REAL-WORLD EXPERIENCES OF DISTRICT AND STATE LEADERS.

IF WE ARE GOING TO MOVE FORWARD WITH EDUCATION REFORM THAT PRIORITIZES STUDENT OUTCOMES, THAT PRIZES CLEAR, UNDERSTANDABLE, AND RELIABLE RESULTS, AND THAT INFORMS PARENTS SO THEY CAN MAKE WISE DECISIONS ABOUT THEIR STUDENTS' EDUCATION, WE MUST MAKE SURE THAT OUR TESTING OUTCOMES ARE TRUSTED, BY PARENTS, BY POLICYMAKERS, AND BY THE PRESS.

AND SO TODAY I CALL ON YOU TO RECOGNIZE THE IMPORTANT ROLE THAT YOU PLAY AND TO ACCEPT THAT YOUR SUCCESS OR FAILURE WILL NOT BE MEASURED BY STATISTICIANS OR YOUR COLLEAGUES IN THIS ROOM. WE WILL ONLY BE SUCCESSFUL WHEN WE CAN MEANINGFULLY EXPLAIN TO THE PARENTS OF OUR STUDENTS THAT THE TEST SCORES THEY RECEIVE PROVIDE CRITICAL AND RELIABLE INFORMATION ABOUT THEIR STUDENTS' ACHIEVEMENT AND THAT THESE RESULTS ARE ONE IMPORTANT ELEMENT IN GAUGING STUDENT SUCCESS. THANK YOU.

[APPLAUSE]

**JACK BUCKLEY:** THANK YOU VERY MUCH, KAYA.

NOW, IN THE FIRST OF WHAT'S GOING TO BECOME A FAMILIAR DRILL, WE'RE GOING TO TAKE A COUPLE MINUTES WHILE WE SHUFFLE NAME PLACARDS UP TO THE FRONT FOR OUR FIRST PANEL. SO IF OUR PANEL CAN JOIN US AT THE FRONT OF THE ROOM, WE'LL MOVE YOUR NAMETAGS.

[BREAK]

PANEL I: PREVENTION OF IRREGULARITIES IN ACADEMIC TESTING

**JACK BUCKLEY:** AS I MENTIONED IN THE INTRODUCTION, OUR FIRST PANEL ADDRESSES THE QUESTION OF HOW TO PREVENT TESTING IRREGULARITIES, WITH A FOCUS ON WHAT SEAs AND LEAs CURRENTLY DO, BARRIERS TO IMPLEMENTING BEST PRACTICES, AND LESSONS LEARNED FROM THE FIELD.

OUR FIRST PANELIST, AUDREY AMREIN-BEARDSLEY, IS CURRENTLY ASSOCIATE PROFESSOR IN THE MARY LOU FULTON TEACHERS COLLEGE AT ARIZONA STATE UNIVERSITY. HER RESEARCH INTERESTS INCLUDE EDUCATIONAL POLICY, EDUCATIONAL MEASUREMENT, AND RESEARCH METHODS, AND SHE HAS PUBLISHED EXTENSIVELY ON ACADEMIC TESTING.

GREGORY J. CIZEK IS PROFESSOR OF EDUCATIONAL MEASUREMENT AND EVALUATION AT THE UNIVERSITY OF NORTH CAROLINA-CHAPEL HILL, WHERE HE TEACHES COURSES IN PSYCHOMETRICS, ASSESSMENT, STATISTICS, RESEARCH METHODS, AND PROGRAM EVALUATION. HIS INTERESTS INCLUDE STANDARD SETTING, VALIDITY, TEST SECURITY, AND TESTING POLICY, AND HE'S THE AUTHOR OF "FILLING IN THE BLANKS" IN 1999, "CHEATING ON TESTS: HOW TO DO IT, DETECT IT, AND PREVENT IT" IN 1999, "DETECTING AND PREVENTING CLASSROOM CHEATING," 2003, AMONG OTHER PUBLICATIONS.

SCOTT NORTON IS THE ASSISTANT SUPERINTENDENT OF THE OFFICE OF STANDARDS, ASSESSMENTS, AND ACCOUNTABILITY AT THE LOUISIANA DEPARTMENT OF EDUCATION. HE HOLDS A PH.D. IN EDUCATIONAL ADMINISTRATION AND SUPERVISION FROM LOUISIANA STATE UNIVERSITY. HIS OFFICE IS RESPONSIBLE FOR IMPLEMENTATION OF CONTENT STANDARDS AND DEVELOPMENT OF THE LOUISIANA COMPREHENSIVE CURRICULUM, ALL PARTS OF THE STATEWIDE ASSESSMENT PROGRAM, AND IMPLEMENTATION OF THE STATE ACCOUNTABILITY SYSTEM. HE'S ALSO THE STATE'S K-12 LEAD IN THE PARTNERSHIP FOR ASSESSMENT OF READINESS FOR COLLEGE AND CAREERS, OR PARCC, ONE OF THE TWO RACE TO THE TOPASSESSMENT CONSORTIA.

FINALLY, JAMES S. LIEBMAN IS THE SIMON H. RIFKIND PROFESSOR OF LAW AT COLUMBIA LAW SCHOOL, WHERE HE HAS BEEN A MEMBER OF THE FACULTY SINCE 1985. FROM 2006 TO 2009, HE TOOK A PARTIAL LEAVE FROM LAW SCHOOL TO SERVE AS CHIEF ACCOUNTABILITY OFFICER OF THE NEW YORK CITY DEPARTMENT OF EDUCATION. WHILE THERE, HE CREATED AND LED THE DIVISION OF ACCOUNTABILITY AND ACHIEVEMENT RESOURCES, DIRECTED THE CITY'S ASSESSMENT PROGRAMS, LED THE DESIGN AND IMPLEMENTATION OF THE CITY'S COMPREHENSIVE NEW ACCOUNTABILITY SYSTEM, FACILITATED THE ESTABLISHMENT IN EACH OF THE CITY'S 1,500 SCHOOLS OF INQUIRY TEAMS FOR USING DATA TO IMPROVE STUDENT LEARNING, AND DESIGNED AND DEPLOYED AN AWARD-WINNING CITY-WIDE INSTRUCTIONAL DATA SYSTEM.

WE'LL BEGIN WITH AUDREY.

**AUDREY AMREIN-BEARDSLEY:** GOOD MORNING. THANK YOU FOR HAVING US TODAY.

I'M GOING TO START WITH DEGREES OF CHEATING AND THE PREVENTION OF TESTING IRREGULARITIES. BACKGROUND -- WE GOT SOME BACKGROUND FROM BEFORE, THE PRESENTERS BEFORE. FROM THE RESEARCH, HERE ARE SOME THINGS THAT I PULLED FROM THE LAST TWO DECADES OF RESEARCH.

FOR STARTERS, PHELPS, 2005, WROTE, "IF A TEACHER'S PERFORMANCE IS JUDGED, IN WHOLE OR IN PART, ON THE BASIS OF THEIR STUDENTS' TEST RESULTS, CERTAINLY THEY ARE GIVEN AN INCENTIVE TO CHEAT." NICHOLS AND BERLINER, 2007, WROTE, "HIGH-STAKES TESTING ALMOST ALWAYS CORRUPTS THE INDICATORS USED AND THE EDUCATORS JUDGED BY SUCH TESTS." CAMPBELL'S LAW, 1976--"THE MORE ANY QUANTITATIVE SOCIAL INDICATOR IS USED FOR SOCIAL DECISION-MAKING, THE MORE SUBJECT IT WILL BE TO CORRUPTION PRESSURES AND THE MORE APT IT WILL BE TO DISTORT AND CORRUPT THE SOCIAL PROCESSES IT IS INTENDED TO MONITOR."

BUT BACK IN 1999, BEFORE NO CHILD LEFT BEHIND WAS PASSED, SACKS QUESTIONED THE PREVALENCE OF CHEATING, ARGUING THAT IT IS "MORE LIKELY RARE THAN COMMON."

LORRIE SHEPHARD, PROFESSOR AT BOULDER, COLORADO, 1990, ARGUED A DECADE EARLIER THAT CHEATING "IS GENERALLY BELIEVED TO OCCUR IN A VERY TINY PERCENTAGE OF SCHOOLS," APPROXIMATELY 1.3%.

AGAIN, THIS IS PRE-NO CHILD LEFT BEHIND. WHAT WE DO KNOW, IT'S VERY DIFFICULT TO QUANTIFY THE LEVEL AT WHICH CHEATING OCCURS, AND IT'S ALSO DIFFICULT TO QUANTIFY HOW IT OCCURS, AND THAT'S WHAT WE'VE ENGAGED IN THIS RESEARCH STUDY THAT I'M GOING TO TALK ABOUT NEXT.

RECENTLY, AS WE ALL ARE AWARE AND ARE HERE FOR, ANALYSTS FOUND THAT ONE IN 5 ELEMENTARY AND MIDDLE SCHOOLS ACROSS THE STATE OF GEORGIA SUBMITTED "HIGHLY ABNORMAL" ANSWER SHEETS, WITH ALMOST 90% OF ONE SCHOOL'S SCORES SUSPECT. AND NEWSPAPER ARTICLES CONTINUE TO BE PUBLISHED ON PRETTY MUCH A DAILY BASIS NOW ABOUT TEACHERS AND ADMINISTRATORS ACCUSED OF CHEATING ON HIGH-STAKES TESTS, YET STILL AN ACCURATE NUMBER OF INCIDENCES STILL ELUDES US.

THEORETICAL ASSERTIONS. WE CAN ASSUME WITH CONFIDENCE THAT THERE ARE MANY MORE INCIDENTS OF CHEATING ON HIGH-STAKES TESTS THAN ARE REPORTED IN THE MEDIA. SO THE PROBLEM IS PROBABLY WORSE THAN WE EVEN EXPECT IT TO BE GIVEN WHAT WE'VE READ. WE CAN ALSO ASSUME WITH CONFIDENCE THAT THIS HAS BECOME INCREASINGLY APPARENT POST-NO CHILD LEFT BEHIND, WITH THE HIGHER-STAKES CONSEQUENCES ATTACHED TO TEST OUTCOMES.

BUT WHAT IS CHEATING? HOW DO WE DEFINE CHEATING? AND MIGHT WE UNDERSTAND CHEATING BETTER TO PREVENT SUCH INCIDENCES FROM, FOR EXAMPLE, CAUSING DRAMATIC IRREGULARITIES AND DISTORTING VALID INTERPRETATIONS? WE HAVE A RELIABILITY ISSUE IN TERMS OF CONSISTENCY OVER TIME ON TEST SCORES, BUT I THINK MORE IMPORTANTLY, WE HAVE AN ISSUE WITH THE ABILITY TO WHICH WE CAN MAKE VALID INTERPRETATIONS ABOUT STUDENT LEARNING, STUDENT ACHIEVEMENT, TEACHER QUALITY, TEACHER EFFECTIVENESS, SCHOOL QUALITY, AND SO FORTH. IT'S THE VALID INTERPRETATIONS WITH WHICH I THINK WE'RE MOST CONCERNED.

THE EMPIRICAL APPROACH THAT WE TOOK--THERE'S A GROUP OF COLLEAGUES OF MINE AT ARIZONA STATE UNIVERSITY. WE INVESTIGATED THE TYPES OF AND DEGREES TO WHICH TEACHERS IN ARIZONA--THE SAMPLE WAS ABOUT 3,000 FROM THE POPULATION--ENGAGED IN TEST-RELATED CHEATING PRACTICES ON STATE-MANDATED, HIGH-STAKES, LARGE-SCALE TESTS. THIS IS THE NO CHILD LEFT BEHIND TEST IN ARIZONA, WHICH IS CALLED THE AIMS.

KNOWING THAT THESE BEHAVIORS DO CREATE IRREGULARITIES IN ACADEMIC TESTING, PARTICULARLY WHEN CONSEQUENCES ARE ATTACHED, THE GOAL HERE WAS TO UNDERSTAND WHAT CHEATING LOOKS LIKE IN THE CLASSROOM AT THE APPLIED LEVEL, AND BY UNDERSTANDING, TO INFORM POLICIES--STATE

LEVEL, DISTRICT LEVEL, CLASSROOM LEVEL, EVEN--POLICIES TO PREVENT FURTHER CHEATING.

ONCE WE STARTED GATHERING THE DATA, WE DETERMINED THAT WE COULD NOT CAPTURE--ONE DEGREE OF CHEATING WAS DIFFERENT THAN ANOTHER INCIDENCE OF CHEATING, SO WE KNEW THAT WE HAD TO COME UP WITH SOME KIND OF DEFINITION. WE CAME UP WITH A TAXONOMY OF CHEATING.

WE BASED THIS ON THE FIRST-, SECOND-, AND THIRD-DEGREE OFFENSES IN THE FIELD OF LAW, SO WE BASICALLY CALLED THIS THE FIRST DEGREE OF CHEATING, SECOND DEGREE OF CHEATING, AND THIRD DEGREE OF CHEATING TO TALK ABOUT THIS MORE ELOQUENTLY, WE HOPE, AND TO ALSO UNDERSTAND IT TO PREVENT IT.

WE HOPED THE TAXONOMY WAS USEFUL FOR UNDERSTANDING AND PREVENTING FURTHER IRREGULARITIES. THE TAXONOMY CAME OUT TO BE LIKE THIS.

THIS EMERGED AFTER THE DATA WERE COLLECTED FROM THESE 3,000 TEACHERS IN THE STATE OF ARIZONA. CHEATING IN THE FIRST DEGREE, JUST AS IN THE FIELD OF LAW--WILLFUL AND PREMEDITATED, THE MOST SERIOUS AND THE MOST WORTHY OF SANCTIONS.

FOR EXAMPLE, THINGS THAT CAME OUT THROUGH THE DATA--ERASING AND CHANGING STUDENTS' TEST ANSWERS, FILLING IN BUBBLES LEFT BLANK,

OVERTLY AND COVERTLY PROVIDING CORRECT ANSWERS ON TESTS,  
FALSIFYING STUDENT I.D. NUMBERS, BECAUSE IF THE STUDENT I.D.  
NUMBERS ARE FALSE, THEN THE STUDENT'S TESTS USUALLY GET REJECTED  
FROM THE SYSTEM, EXCLUDING/SUSPENDING STUDENTS WITH POOR ACADEMIC  
PERFORMANCE--THE ONES THAT ARE LESS DESIRABLE WHEN TESTING TIME  
COMES.

CHEATING IN THE SECOND DEGREE WE CLASSIFIED AS OFTEN MORE SUBTLE,  
DEFINED MORE CASUALLY, NOT NECESSARILY PREMEDITATED OR WITH  
MALINTENT. FOR EXAMPLE, ENCOURAGING STUDENTS TO REDO PROBLEMS OR  
DOUBLE-CHECK THEIR WORK, ACCOMMODATING FOR "STUPID MISTAKES"--WE  
TALKED TO A LOT OF TEACHERS. A LOT OF TEACHERS REPORTED THEY'D WALK  
BY AND THEY KNEW THAT THEIR STUDENTS KNEW THE RIGHT ANSWER, AND  
THEY'D GIVE THEM A NUDGE AND SAY, YOU KNOW, "DOUBLE-CHECK THAT  
ONE." SO IT WAS SOMETHING THAT WASN'T FIRST DEGREE--IT WASN'T  
OUTRIGHT PREMEDITATED--BUT IT WAS SOMETHING THAT A LOT OF TIMES,  
THE TEACHERS THOUGHT THAT THEY WERE HELPING THEIR STUDENTS, WHEN IT  
COULD BE DEFINED AND CLASSIFIED AS CHEATING. DISTRIBUTING "CHEAT  
SHEETS," SOMETIMES THAT INCLUDE VOCABULARY WORDS. THE VOCABULARY  
MIGHT BE INCLUDED IN THE TEST STEMS OR THE QUESTIONS. TALKING  
STUDENTS THROUGH PROCESSES AND DEFINITIONS, GIVING STUDENTS EXTRA  
TIME ON TESTS OR TIME DURING LUNCH, RECESS, AND BEFORE OR AFTER  
SCHOOL.

CHEATING IN THE THIRD DEGREE WE DEFINED AS CAUSED BY INDIFFERENCE, RECKLESSNESS, OR NEGLIGENCE, AND ALSO REFERRED TO AS INVOLUNTARY CHEATING. IN MANY OF THESE INCIDENCES, TEACHERS DIDN'T THINK WHAT THEY WERE DOING WAS CHEATING. A LOT OF TIMES, THEY THOUGHT IT WAS REALLY GOOD TEST PREPARATION PRACTICES. THINGS LIKE, WHICH WE'VE ALL HEARD BEFORE, "TEACHING TO THE TEST," WHICH IS VERY CONTROVERSIAL BECAUSE THERE'S A DIFFERENCE BETWEEN TEACHING TO THE TEST, WHICH IS ACTUALLY THE TEST ITEMS, AND THEN TEACHING TO THE STANDARDS. SO THERE'S A LOT OF CONTROVERSY. JAMES POPHAM FROM UCLA TALKS ABOUT THIS QUITE A BIT.

ACCESS TO TEST BLUEPRINTS, WHICH I BELIEVE IS A STATE POLICY. BLUEPRINTS ARE PUT ONLINE, BUT A LOT OF TIMES, TEACHERS TAKE THOSE AND TEACH TO THE BLUEPRINT, SO THERE'S ACTUALLY ANOTHER CAVEAT THERE, WHERE IT'S NOT JUST TEACHING TO THE TEST; IT COULD BE TEACHING TO THE BLUEPRINT, WHICH STILL MARGINALIZES SOME THINGS THAT ARE NOT TESTABLE.

NARROWING OF THE CURRICULUM. I THINK WE ALL UNDERSTAND WHAT THAT IS, BUT MARGINALIZING THINGS LIKE PHYSICAL EDUCATION, RECESS, EVEN SOCIAL STUDIES AND SCIENCE, ESPECIALLY RIGHT BEFORE THE TESTS ARE ADMINISTERED. HYPER-UTILIZING "CLONE ITEMS," WHERE YOU TAKE THE ITEMS IN THE WORD PROBLEMS AND YOU CHANGE THE NAMES OF THE PEOPLE

IN THE WORD PROBLEMS AND YOU CHANGE THE NUMBERS, BUT YOU BASICALLY USE THE SAME TYPES OF TESTS AND DEVELOP CLONE ITEMS FOR TEST PREPARATION. IS THAT PROFESSIONAL? IS IT ETHICAL? IS THAT APPROPRIATE TEST PREPARATION PRACTICE?

THE CHEATING IN THE THIRD DEGREE IS WHERE IT GETS REALLY INTO THE GRAY AREA, AND A LOT OF PEOPLE DIFFER ON WHETHER THEY AGREE THAT THAT'S CHEATING OR NOT. AND ALSO INORDINATELY FOCUSING ON TEST-TAKING STRATEGIES. DO WE HAVE TO CLICK TO GET THAT BACK?

WE LOST OUR...

BUCKLEY: TIME'S UP, TIME'S UP.

[LAUGHTER]

AMREIN-BEARDSLEY: OK. ALL RIGHT. I'LL JUST TURN AROUND. THERE WE GO.

OK. BUT I FAST-FORWARDED. NOW I GOT TO GO BACK.

OK. THAT WAS OUR TAXONOMY, AND THIS IS THE CHART THAT'S PUT TOGETHER IN THE ACTUAL PAPER. THE PAPER IS AVAILABLE. IT'S FREE ONLINE. IT'S CALLED "CHEATING IN THE FIRST, SECOND, AND THIRD DEGREE," IN CASE ANYONE IS INTERESTED IN READING FURTHER. IT'S

ABOUT 60 SINGLE-SPACE PAGES, SO IT'S QUITE THE READ, BUT THESE ARE THE MAIN PIECES THAT I PULLED OUT OF IT.

YOU'LL SEE IN THIS CHART HERE, WE'VE SEPARATED INTO FIRST-DEGREE CHEATING, SECOND-DEGREE CHEATING, AND THIRD-DEGREE CHEATING. IF YOU LOOK AT THE FAR LEFT, FOR EXAMPLE, IT'S CATEGORIZED BY, A-- CHEATING BEHAVIOR IS ERASED AND CHANGED TEST ANSWERS. THE WHITE COLUMN IS HEARSAY--"I'VE HEARD ABOUT OTHER TEACHERS DOING THIS, BUT I'VE NEVER DONE THIS MYSELF"--AND THEN THE DARKER COLUMN IS SELF-ADMITTED--"YEAH, I HAVE DONE THIS BEFORE IN THE PAST, AT SOME LEVEL."

THEN YOU CAN GO ACROSS, AND IT'S THE SAME AS THE HEARSAY AND THEN THE SELF-REPORTED ACROSS, AND YOU CAN SEE JUST BY LOOKING AT THIS THAT THE FIRST-DEGREE CHEATING, THE MOST OUTRIGHT PREMEDITATED, IS WITH THE LOWEST FREQUENCY, AND PARTICULARLY WITH THE SELF-REPORT, ALTHOUGH THERE'S ALSO A SOCIAL DESIRABILITY THREAT THERE, WHETHER TEACHERS ARE REALLY GOING TO ADMIT THEY ACTUALLY DID IT, BUT IT'S A LOOSE INDICATOR OF WHAT THIS MIGHT LOOK LIKE.

YOU CAN SEE AS YOU GET TO THE SECOND-DEGREE LEVELS AND THE THIRD-DEGREE LEVELS, THESE THINGS ARE MORE COMMON. FOR EXAMPLE, LEVEL "H," WHICH IS ENCOURAGED STUDENTS TO REDO PROBLEMS, WAS PRETTY COMMON AMONG THESE PARTICIPANTS. OK?

SO WHAT TO DO? WE FRAMED THESE RECOMMENDATIONS ALSO WITHIN THE SAME TAXONOMY. FIRST-DEGREE RECOMMENDATIONS--KEEP TESTS SECURE BEFORE AND AFTER ADMINISTRATION, WHICH SEEMS TO BE PRETTY COMMON PRACTICE THESE DAYS; PREVENT EXPOSURE, PHOTOCOPYING, WHICH WAS VERY COMMON, ITEM TRANSFORMATION AND CLONING, WHICH WAS VERY COMMON, AND THE YEAR-TO-YEAR USE OF OLD FORMS. A LOT OF TEACHERS REPORTED THAT IF THEY EVEN HAD ACCESS, THEY WOULD EITHER MEMORIZE THE ITEMS AND WRITE THEM DOWN, OR IF THEY HAD A CHANCE, THEY WOULD MAKE PHOTOCOPIES AND USE THEM AS THE TEST-PREPARATION MATERIALS FOR THE FOLLOWING YEAR 1, YEAR 2, YEAR 3. THEY'D USE THEM CONTINUOUSLY OVER TIME.

HAVE THE LEAST LIKELY TO DISTORT AND ARTIFICIALLY INFLATE IN CHARGE OF TESTING PROCEDURES. SO DECREASE THE LIKELIHOOD OF THINGS LIKE ORAL EMPHASIS OF CORRECT ANSWERS, REWORDING PROBLEMS, DEFINING KEY TERMS, EXTRA TIME ON TESTS, TEST MANIPULATION, WHAT WE CALL "CLEANING UP THE TEST."

ADMINISTER TESTS IN ARTIFICIAL ENVIRONMENTS. ADMINISTRATION SHOULD OCCUR WHERE ACCESS TO CURRICULAR MATERIALS, RESOURCES, AND VISUALS IS MOST LIMITED. ONE OF THE THINGS THAT WAS ASKED ON THE SURVEY IS WHETHER TEACHERS KEPT UP CLASSROOM RESOURCES AROUND THE CLASSROOM, INCLUDING POSTERS AND NUMBER CHARTS, AND A LOT OF THEM WERE

ACTUALLY VERY ANGRY BY THAT ITEM BECAUSE THEY THOUGHT THAT IT WAS NOT FAIR TO MAKE THE ENVIRONMENT ARTIFICIAL, THAT THE STUDENTS SHOULD HAVE ACCESS TO THOSE RESOURCES, AND SO THAT'S SOMETHING THAT WE'D HAVE TO DEBATE, AS WELL--WHETHER IT SHOULD BE AN ENTIRELY ARTIFICIAL ENVIRONMENT OR IT SHOULD JUST BE WHAT THEY'RE AROUND ON A DAILY BASIS ANYWAY.

PUT IN PLACE POLICIES TO ENSURE THAT THE MOST "UNDESIRABLE" OR LOWEST-SCORING STUDENTS CANNOT BE EXEMPTED OR EXCLUDED OR CONVINCED TO STAY HOME UNTIL THE TEST IS OVER. THAT WAS QUITE COMMON, ACTUALLY--CONVINCING SOME OF THESE STUDENTS THAT, "DON'T BOTHER TO COME TO SCHOOL TODAY," PARTICULARLY ENGLISH-LANGUAGE LEARNERS AND SPECIAL EDUCATION STUDENTS WHO ARE TYPICALLY SOME OF OUR LOWEST-SCORING STUDENTS ANYHOW. SECOND- AND THIRD-DEGREE RECOMMENDATIONS. THIS IS WHERE THE GRAY AREA COMES BACK IN. WHETHER WE DEFINE THESE ARE CHEATINGS OR APPROPRIATE TEST-PREPARATION PRACTICES HAS TO BE DETERMINED, PROBABLY MORE ON A LOCAL BASIS.

FIRST THING HERE IS DO NOT BECOME OVERLY DEPENDENT ON TECHNICAL SOLUTIONS. FOR EXAMPLE, ERASURE ANALYSES AND OTHER TECHNICAL APPROACHES CAN ONLY DO SO MUCH. IF I WAS A TEACHER, I THINK IT WOULD BE VERY LIKELY IF I KNEW THAT THE ERASURE ANALYSES WERE THE KEY INDICATORS THAT WERE GOING TO CATCH ME AS A CHEATER, AS A TEACHER CHEATER, I PROBABLY WOULD JUST CONVINCED THE STUDENTS TO

LEAVE THE ANSWERS BLANK AND JUST GO BACK IN AND FILL THEM IN MYSELF OR WHATEVER IT MIGHT BE.

WE CAN'T OVER-RELY ON TECHNICAL SOLUTIONS TO THIS. I THINK WE HAVE TO BE VERY HOLISTIC IN THE APPROACHES THAT WE TAKE AND UNDERSTAND THE COMPLEXITY AND THE DEGREES TO WHICH THESE THINGS EXIST AND OCCUR. A HEALTHY TESTING CULTURE IS MOST IMPORTANT. MANY DO NOT CONSIDER, AGAIN, WHAT THEY ARE DOING IS CHEATING, AND WHILE ONE EDUCATOR MIGHT CONSIDER SOME OF THESE PRACTICES SMART, OTHERS MIGHT CONSIDER THEM UNPROFESSIONAL OR UNETHICAL.

SOME IRREGULARITIES OCCUR AS PART OF DISTRICT OR SCHOOL TEST PREPARATION AND ADMINISTRATION POLICIES, TOO. THERE WERE A LOT OF TEACHERS THAT SAID, "I DON'T AGREE WITH WHAT MY PRINCIPAL IS SAYING THAT WE SHOULD DO WITH TEST PREPARATION OR USING THESE MATERIALS FOR TEST PREPARATION OR BRINGING IN TEST CONSULTANTS SAYING, YOU KNOW, 3 IN A ROW, NO, NO, NO, BECAUSE NO SELF-RESPECTING TEST AUTHOR WOULD EVER PUT THE ANSWERS C, C, C." REALLY MANIPULATIVE TYPES OF TEST PREPARATION EXERCISES THAT SOME OF THESE STUDENTS ARE GOING THROUGH, AND THE TEACHERS REJECTED THAT, BUT THEY WOULD SAY, "WELL, THE ADMINISTRATOR...THIS IS THE ADMINISTRATOR. IT'S ONE OF THOSE TOP-DOWN INITIATIVES, AND HOW DO WE COMBAT THIS? WE AGREE THAT WE'RE CHEATING, BUT IT'S ALSO THE ADMINISTRATOR SAYING THAT WE'RE GOING TO DO THIS FOR PROFESSIONAL DEVELOPMENT DAY."

PARENT/TEACHER/ADMINISTRATOR ANONYMOUS WHISTLE-BLOWING SYSTEM WAS A SUGGESTION, BUT MANY BELIEVED THAT WHEN CHEATING WAS EVEN REPORTED ANONYMOUSLY, IT WENT IGNORED OR CONSEQUENCES WERE NEGLIGIBLE--A SLAP-ON-THE-HAND TYPE OF CIRCUMSTANCE OR CONSEQUENCE.

OOH! I'M ALMOST OUT OF TIME. COME TO COLLECTIVE UNDERSTANDINGS AND LOCAL POLICIES ABOUT PROFESSIONAL AND ETHICAL TEST PREPARATION PRACTICES. I THINK THE BIG KEY HERE IS COLLECTIVELY DETERMINE WHETHER THINGS LIKE TEACHING TO THE TEST, WHETHER THESE THINGS ARE APPROPRIATE, AND I THINK THAT THE LOCAL OR THE DISTRICT LEVEL, YOU CAN COME TO DO THAT. IT'S A MORE CULTURALLY APPROPRIATE SOLUTION TO THIS.

AND KEEP DECISIONS ALWAYS, AND POLICIES, IN THE BEST EDUCATIONAL INTERESTS OF THE CHILDREN. AND I THINK I CAN SKIP THAT, SINCE I'M OUT OF TIME.

THANK YOU.

**JACK BUCKLEY:** THANK YOU VERY MUCH. GREG.

**GREG CIZEK:** THANK YOU, JACK, EUNICE.

IT'S AN HONOR, AND I REALLY APPRECIATE BEING INVITED TO SPEAK HERE. MY COMMENTS ARE ENTITLED "SHARED RESPONSIBILITIES." I WANT TO TALK ABOUT THE ROLE THAT STUDENTS, EDUCATORS, TEST DEVELOPERS, AND PROFESSIONAL ASSOCIATIONS HAVE IN PREVENTING.

FIRST, I'LL JUST QUICKLY GLOSS OVER SOME THINGS RELATED TO WHAT--I THINK STUDENTS PLAY A ROLE IN THIS, AS WELL. THAT IS--3 QUICK THINGS. STUDENTS NEED TO BE A PART OF AN ACADEMIC COMMUNITY THAT EMBRACES INTEGRITY. STUDENTS NEED TO ADOPT MORE LEARNING VERSUS PERFORMANCE GOALS, AND EDUCATORS HAVE A ROLE, OBVIOUSLY, IN FACILITATING THAT. AND STUDENTS HAVE SOME OBLIGATIONS, ALSO, TO REPORT THEIR CONCERNS. WE KNOW THAT HONOR CODES ARE EFFECTIVE. THERE'S REASONS WHY THEY'RE EFFECTIVE, AND IT'S NOT BECAUSE PEOPLE SIGN SOMETHING. IT'S BECAUSE THE CULTURE, THE SCHOOL, EVERYONE IN THE COMMUNITY SEES THEMSELVES AS ALL EMBRACING ACADEMIC INTEGRITY.

EDUCATORS HAVE A ROLE--TRAINING IN ASSESSMENT AND QUALIFICATION. THE FIRST THING--LET'S SEE. DID I SKIP ONE? YEAH. THE FIRST THING I SHOULD SAY ABOUT EDUCATORS IS THAT THEY CAN DO THE SAME KINDS OF THINGS THAT STUDENTS DO IN TERMS OF BEING RESPONSIBLE FOR IMPLEMENTING A CULTURE WHERE LEARNING IS THE GOAL, AS OPPOSED TO PERFORMANCE, REPORTING CONCERNS, AND EMBRACING THE CULTURE OF ACADEMIC INTEGRITY.

BUT BEYOND THAT, I LABELED THIS "EDUCATORS" BECAUSE I THINK TOO OFTEN WHEN WE SPEAK ABOUT THIS ISSUE, WE SAY "TEACHERS." IF WE KNOW ANYTHING NOW, IT'S SCHOOL COUNSELORS, PRINCIPALS, SUPERINTENDENTS, CHANCELLORS-- EVERYBODY HAS HAD THEIR FINGER IN THE PIE AT ONE TIME OR ANOTHER, SO IT'S UNFAIR, I THINK, TO PIN IT EXPRESSLY ON JUST TEACHERS.

THE FIRST THING I'LL TALK ABOUT IS TRAINING IN ASSESSMENT. IT'S WOEFULLY INADEQUATE. IN MANY STATES, YOU DON'T HAVE TO BE QUALIFIED IN ASSESSMENT TO BECOME A TEACHER, YOU DON'T HAVE TO BE QUALIFIED IN ASSESSMENT TO BECOME A PRINCIPAL. THAT'S A SHAME. STATE POLICIES SHOULD CHANGE IN THAT MATTER. WHAT HAPPENS IS IT LEADS TO A DISTRUST OF TESTING.

I HEAR THE MOST OUTLANDISH THINGS IN MY STATE. WE GIVE DIFFERENT FORMS OF A TEST--THE RED, THE BLUE, OR THE PURPLE, OR SOMETHING LIKE THAT--AND EVERYONE KNOWS THE RED ONE IS HARDER. THAT'S RIDICULOUS. THEY'RE PARALLEL FORMS. IF YOU KNOW ANYTHING ABOUT TESTING, YOU WOULD KNOW THAT IT WOULD BE IRRESPONSIBLE TO GIVE A HARDER TEST TO SOME KIDS AND EASIER TEST TO OTHERS. IT LEADS TO MISTRUST OF PEOPLE WHO TEST. IT LEADS TO OUTLANDISH THINGS LIKE PEOPLE SAYING, "WELL, THIS ISN'T RELIABLE. THE TEST THAT THE STATE GIVES, IT'S BIASED. IT'S UNRELIABLE. IT'S A SINGLE SNAPSHOT OF WHAT A KID CAN DO. WHO CAN COUNT ON ONE DAY'S PERFORMANCE?"

THAT'S ABSURD. IF YOU UNDERSTOOD HOW LARGE-SCALE TESTS ARE PUT TOGETHER--I'LL JUST SINGLE OUT WES BRUCE. THE SINGLE LARGE-SCALE TEST THAT HE GIVES TO STUDENTS IN THE STATE OF INDIANA EVERY YEAR IS BY FAR AND AWAY THE MOST OBJECTIVE, FAIREST, LEAST BIASED, MOST DEPENDABLE MEASURE THAT THAT KID WILL SEE THE ENTIRE YEAR, ESPECIALLY COMPARED TO THE STUFF THAT'S PRODUCED, LARGELY, BY CLASSROOM TEACHERS WHO DON'T HAVE TRAINING IN ASSESSMENT OR WHO USE WHAT HAPPENS TO BE PASTED IN THE BACK OF SOME INSTRUCTOR'S MANUAL, AND THEY RELY ON THAT KIND OF STUFF.

LARGE-SCALE TESTS ARE REALLY PRETTY GOOD, AND THERE'S JUST GROSS IGNORANCE BECAUSE--I'M NOT FINGERING TEACHERS OR PRINCIPALS BECAUSE OF THIS; I'M FINGERING STATES WHO DON'T REQUIRE EDUCATORS TO HAVE A FIRM GROUNDING IN ASSESSMENT.

I TALK ABOUT QUALIFICATION. YOU NEED TO BE QUALIFIED TO TEACH. YOU ALSO NEED TO BE QUALIFIED TO GIVE AND INTERPRET ASSESSMENTS, AND WE DON'T HIGHLIGHT THAT ENOUGH. PEOPLE NEED TO BE FAMILIAR WITH HOW TO GIVE A TEST AND FAMILIAR WITH THE DOs AND DON'Ts OF QUALITY DATA-INFORMATION GATHERING. THAT ALSO INCLUDES EXAMINATION PROCTORING OR TEST PROCTORING. IT'S NOT JUST A DAY OFF OR A PERIOD OFF TO WATCH OR READ A BOOK OR CATCH UP ON SOME GRADING. THERE'S A

CONSCIENTIOUS ROLE THAT EDUCATORS HAVE TO PLAY WHEN A TEST IS GOING ON, AND THERE'S ALSO A NEED TO ASSIST STUDENTS.

I'M TALKING ABOUT ASSISTING STUDENTS IN AN APPROPRIATE WAY.

IF A STUDENT DOESN'T KNOW HOW TO FILL IN A GRIDDED RESPONSE ITEM, YOU'RE GOING TO GET AN INVALID TEST SCORE FOR THAT KID, SO A TEACHER SHOULD BE THERE OR AN EDUCATOR SHOULD BE THERE TO HELP PEOPLE UNDERSTAND THE TESTING PROCESS. THERE SHOULD BE SPECIFIC RESPONSIBILITIES ASSIGNED FOR PEOPLE IN TERMS OF REVIEWING OR MONITORING TEST RESULTS.

IN TOO MANY JURISDICTIONS, THERE'S NOT SOMEBODY WHOSE SPECIFIC JOB IT IS, AS PART OF THEIR JOB RESPONSIBILITY, TO MONITOR TEST RESULTS. THERE SHOULD BE DISSEMINATION OF THE PROCEDURES THAT ARE IN PLACE TO PREVENT AND FOLLOW-UP DISSEMINATION OF THE CONSEQUENCES THAT HAVE TAKEN PLACE. WHEN YOU HEAR THAT YOUR NEIGHBOR GOT AUDITED, IT'S A STRONG DISINCENTIVE FOR YOU TO CHEAT ON YOUR TAXES, SO EVEN THE CONSEQUENCES SHOULD BE DISSEMINATED, AS WELL.

AND WE SHOULD CONSIDER PROCTORING ASSIGNMENTS FOR EDUCATORS. FOR LOWER GRADES, IT'S REALLY IMPORTANT TO HAVE THE CLASSROOM TEACHER THERE, BUT THERE IS ABSOLUTELY NO REASON WHERE AN END-OF-GRADE TEST

IN ALGEBRA AT THE HIGH-SCHOOL LEVEL HAS TO BE NURTURED BY THAT HIGH-SCHOOL KID'S CLASSROOM TEACHER. THERE CAN BE INDEPENDENT ASSIGNMENT OF PROCTORING.

TEST PROVIDERS, CONTRACTORS, TEST DEVELOPERS HAVE A ROLE, TOO. SOME BROADER ONES--I THINK WE NEED A CLEAR DEFINITION OF CHEATING, AND I'M DIFFERENTIATING CHEATING FROM AN IRREGULARITY. AN IRREGULARITY IS WHERE SOMEBODY JUST INNOCENTLY DOES SOMETHING INAPPROPRIATE. IT'S NOT AN INTENTIONAL ATTEMPT TO ALTER THE MEANING OF A TEST SCORE, BUT CHEATING IS, AND WE NEED A CLEAR DEFINITION OF WHAT THAT IS.

WE NEED CLEARER, I PROBABLY SHOULD HAVE SAID, EDUCATOR REFERENCE MATERIALS. I'VE SEEN A LOT OF THE TEST ADMINISTRATION MANUALS THAT ARE DISTRIBUTED TO FIELD PERSONNEL TO USE, AND THEY'RE JUST IN TYPICALLY TESTING JARGON AND NOT WELL SUITED TO WHAT EDUCATORS NEED TO UNDERSTAND HOW TO APPROPRIATELY GIVE A TEST. SO THINGS LIKE EXAMPLE/NON-EXAMPLE. IF YOU DO THIS WITH A STUDENT, THAT'S REALLY GOOD. THAT HELPED ENSURE A VALID TEST SCORE. IF YOU DO THIS WITH A STUDENT, THAT WAS INAPPROPRIATE. THAT WOULD BE CHEATING. THOSE KINDS OF EXAMPLES/NON-EXAMPLES NEED TO BE DISSEMINATED.

AND PERHAPS SOMETHING LIKE--I'M GOING OUT A BIT ON A LIMB HERE--BUT I SAID PEOPLE NEED TO BE QUALIFIED TO GIVE TESTS JUST LIKE THEY'RE

QUALIFIED TO TEACH. WE DON'T ACTUALLY HAVE PEOPLE QUALIFIED TO BE ABLE TO GIVE TESTS. I'M AT A UNIVERSITY WHERE IF I'M GOING TO DO ANY KIND OF RESEARCH, WHICH IS REQUIRED IN MY JOB, I HAVE TO QUALIFY TO PROTECT THE RIGHTS OF HUMAN PARTICIPANTS IN RESEARCH. OTHERWISE, I CAN'T EVEN DIRECT A DISSERTATION, MUCH LESS CONDUCT RESEARCH MYSELF.

SOME CONTRACTOR, I THINK, OUGHT TO TAKE THE LEAD IN DEVELOPING A WEB-BASED QUALIFICATION TOOL THAT PEOPLE COULD GO IN AND READ SCENARIOS ABOUT WHAT'S APPROPRIATE OR INAPPROPRIATE PRACTICE AND BE QUALIFIED TO GIVE ASSESSMENTS IN THEIR SCHOOLS.

WE CAN RELY ON LESS-CORRUPTIBLE FORMATS. MULTIPLE CHOICE IS PRETTY EASY TO GAME, BUT IT'S MUCH MORE DIFFICULT TO ERASE A STUDENT'S ESSAY AND WRITE A NEW ONE FOR THEM, SO THE MORE WE RELY ON CONSTRUCTED-RESPONSE FORMATS, I THINK...PEOPLE SAY THERE'S A COST THERE. YES, THERE IS A COST THERE, AND I'M SORRY, BUCK UP. IT'S A COST, OK?

WE SPEND A TINY AMOUNT—THIS IS ANOTHER FALLACY. WE SPEND A TINY AMOUNT IN EDUCATION ON ASSESSMENT OF STUDENT OUTCOMES. IT'S ALWAYS BLOWN OUT OF PROPORTION, ALL THIS TIME AND EXPENSE WE DO ON TESTING. IT'S TINY COMPARED TO THE EDUCATION BUDGET. THE AMOUNT OF MONEY WE SPEND TO ENSURE THE INTEGRITY OF THOSE DATA IS EVEN

TINIER, SO IT'S ABOUT TIME WE PONY UP SOME MONEY TO DO A BETTER JOB ON ASSESSMENT. LAST THING THERE SAID COMPUTER-BASED DELIVERY OF ASSESSMENTS IS LESS SUSCEPTIBLE TO CORRUPTION, AS WELL.

FINALLY, SOME NARROWER, MORE TECHNICAL THINGS, I THINK, TEST PROVIDERS OR CONTRACTORS CAN DO--SIMPLE STUFF, LIKE REQUIRING SEATING CHARTS. I RECENTLY LEARNED THIS REALLY SEXY TERM--"BATCH HEADER." APPARENTLY, BATCH HEADER IS WHAT IDENTIFIES WHO MONITORED THE ANSWER SHEETS THAT ARE ABOUT TO BE PROCESSED BY THE SCANNING AND SCORING SYSTEMS. IT'S IMPORTANT TO KNOW WHO WAS RESPONSIBLE FOR ADMINISTERING A TEST, SO A SIMPLE ADDITION TO A TEST DEVELOPER'S AND STATE'S PROCEDURES OF INCLUDING BATCH HEADERS IS REALLY AN IMPORTANT STEP.

CHAIN OF CUSTODY--KNOW WHO HAD TEST MATERIALS AT ALL TIMES. DELIVER MATERIALS JUST IN TIME AT THE DISTRICT AND GET THEM OUT OF THE SCHOOLS JUST IN TIME. REQUIRE NONDISCLOSURE AGREEMENTS AND SIGNED STATEMENTS THAT NOT ONLY SAY, "I'M WILLING TO UPHOLD THE RULES," BUT, "I UNDERSTAND WHAT THE PENALTIES ARE." TOO OFTEN, WE LEAVE OFF THAT THERE'S SOMETHING SERIOUS THAT CAN HAPPEN IF YOU DON'T DO THIS

RIGHT. TEST DEVELOPERS CAN ALSO DO MORE PURPOSEFUL TEST DESIGN IN TERMS OF WHERE THE PLACEMENT OF FIELD TEST BLOCKS ARE AND

SCRAMBLING OF FORMS AND THINGS LIKE THAT, AND TEST DEVELOPERS CAN ALSO DO A BETTER JOB OF CONSULTING WITH STATES AND DISTRICTS ON METHODS OF DETECTION.

WHEN I WAS INVOLVED IN A RECENT ISSUE, WE ACTUALLY WENT AND LOOKED AT ANSWER SHEETS, AND THE SCANNERS ARE SET PRETTY CONSERVATIVELY. WHEN YOU LOOK AT ANSWER SHEETS BY HUMAN EYE, YOU SEE A LOT MORE ERASURES. ON AVERAGE, WHAT WE SAW WAS ABOUT 10% MORE ERASURES THAN THE SCANNERS DETECTED, AND IT'S EASY TO SEE. SO THOSE RESPONSIBLE FOR TESTING NEED TO CONSULT WITH FOLKS ABOUT THEIR DETECTION METHODS.

FINALLY, I'LL WRAP UP, NEARLY, WITH POLICY DEVELOPMENT. PROFESSIONAL ASSOCIATIONS HAVE A ROLE HERE, AS WELL. I'M UNDER THE IMPRESSION--WAYNE, MAYBE YOU CAN CONFIRM THIS--BUT OUR PROFESSIONAL ASSOCIATION, THE NATIONAL COUNCIL ON MEASUREMENT IN EDUCATION, WORKED AS FAST AS A PROFESSIONAL ASSOCIATION COULD POSSIBLY WORK TO DEVELOP A DRAFT MODEL POLICY ON TEST INTEGRITY TO DELIVER TO THE U.S. DEPARTMENT OF EDUCATION.

YOU SHOULD HAVE THAT NOW. IF NOT, IT IS NOW READY FOR YOU, AND WE WILL DELIVER IT TO YOU. WE'RE PLEASED TO PROVIDE THAT, AND I THINK IT SHOULD, WE HOPE, BECOME SOMETHING THAT DISTRICTS AND STATES CAN RELY ON, AND THE U.S. DEPARTMENT OF ED, AS SOME POLICY GUIDANCE.

FINALLY, JUST SOME CONCLUDING THOUGHTS. SHOULD WE BE SURPRISED THAT CHEATING IS OCCURRING? NO. IT HAPPENS EVERYWHERE. EVERYBODY CHEATS ON EVERYTHING. SOME SMALL PERCENTAGE OF PEOPLE WILL CUT CORNERS WHEREVER THERE ARE INCENTIVES TO DO SO.

SHOULD WE BE SHOCKED? ABSOLUTELY. YOU KNOW, WHERE BARRY BONDS CAN SAY, "I TAKE STEROIDS, BUT YOU DON'T PAY ME TO BE A ROLE MODEL FOR KIDS. YOU PAY ME TO HIT HOME RUNS, FILL SEATS, AND MAKE A LOT OF MONEY FOR MY TEAM AND ENTERTAIN PEOPLE," HE'S ABSOLUTELY RIGHT. EDUCATORS, HOWEVER, CAN'T SAY THAT. WE, BY AND LARGE, ARE CHARGED WITH BEING ROLE MODELS, SO WE SHOULD BE SHOCKED AT THAT.

WHAT'S A WRONG REACTION? "DO AWAY WITH TESTS." TOO MANY PEOPLE ARE SAYING, "OH, TESTS HAVE CONSEQUENCES," AND "OH, THAT'S TERRIBLE. WE SHOULD DIMINISH THE CONSEQUENCES. WE SHOULD LESSEN OUR RELIANCE ON TESTING."

I THINK THAT'S ABSURD. I THINK THAT'S LIKE SAYING IF THERE'S VOTER FRAUD, WE SHOULD DO AWAY WITH FREE ELECTIONS. NOBODY WOULD CONSIDER THAT, AND DOING AWAY WITH THE SOUND INFORMATION WE GET FROM TESTS AND THE CONSEQUENCES ASSOCIATED WITH THEM IS PROBABLY THE MOST WRONG RESPONSE. THE BEST RESPONSE IS PREVENTION, AND THE GOOD NEWS

IS THERE IS PLENTY THAT CAN BE DONE. THERE ARE A LOT OF METHODS TO PREVENT CHEATING AND TEST IRREGULARITIES.

AND FINALLY, I'LL JUST END WITH, IT'S A SHARED RESPONSIBILITY AMONGST STUDENTS, EDUCATORS, PROFESSIONAL ASSOCIATIONS, AND ALL INVOLVED IN THE ENTERPRISE OF TESTING.

**JACK BUCKLEY:** ALL RIGHT. THANKS VERY MUCH.

SCOTT.

**SCOTT NORTON:** I THINK I'M UP. I ALSO WANT TO SAY THANK YOU TO THE DEPARTMENT FOR INVITING ME TO PARTICIPATE. I APPRECIATE THAT VERY MUCH.

I'M FROM LOUISIANA. I'M HERE REALLY, I THINK, TO REPRESENT THE SEA PERSPECTIVE. I THINK LOUISIANA IS PROBABLY PRETTY TYPICAL OF STATES THAT ARE FAIRLY INVOLVED IN TEST SECURITY, AND I'LL TAKE A STAND ON THAT AND SAY STATES SHOULD BE INVOLVED IN TEST SECURITY. IF WE'RE RESPONSIBLE FOR THE TESTS, WE'RE RESPONSIBLE TO MAKE SURE THAT THE RESULTS ARE VALID AND RELIABLE. WHAT THAT MEANS FOR US MOSTLY IS DEVELOPMENT OF STATE POLICY THROUGH THE STATE BOARD, AND I'LL TALK ABOUT THAT JUST A LITTLE BIT, AND THEN FOLLOW-UP ON THAT

POLICY, WORKING WITH THE DISTRICTS AND SCHOOLS. AS ALREADY STATED, IT IS A SHARED RESPONSIBILITY AMONG ALL THOSE PARTIES.

SO WHAT DOES THAT MEAN? WHAT DOES A POLICY LOOK LIKE? I'LL JUST GIVE YOU A FEW OF OUR HIGHLIGHTS.

WE DO REQUIRE THAT SCHOOL DISTRICTS DEVELOP THEIR OWN DISTRICT TEST SECURITY POLICY. IT HAS TO INCLUDE EVERYTHING THAT WE REQUIRE.

THERE ARE SOME LOCAL THINGS THAT WE ASK THEM TO PUT IN, AND THEY SUBMIT THOSE TO THE STATE. WE REQUIRE OATH OF SECURITY, AS WAS MENTIONED EARLIER. WE HAVE DONE THAT FOR SOME TIME FOR THE ADULTS INVOLVED IN TEST ADMINISTRATION. IN RECENT YEARS, WE'VE ALSO ASKED THE STUDENTS TO SIGN AN OATH OF SECURITY. WHETHER OR NOT THAT'S HELPFUL, WE'LL PROBABLY DISCUSS THAT MORE THROUGHOUT THE DAY.

WE REQUIRE CERTAIN PROCEDURES FOR THE HANDLING OF TEST MATERIALS. THAT'S ALSO BEEN DISCUSSED A LITTLE BIT ALREADY. OUR STATE POLICY ALLOWS THAT THE STATE SUPERINTENDENT CAN CANCEL OR VOID SCORES IF THERE'S INAPPROPRIATE FINDINGS, EVEN IF THE DISTRICT DISAGREES, SO THAT'S WRITTEN INTO OUR POLICY.

THE POLICY REQUIRES SEVERAL SPECIFIC TYPES OF ANALYSES. IT DOES NOT REQUIRE SOME OTHERS THAT I'LL MENTION IN JUST A MINUTE. WE REQUIRE ERASURE ANALYSIS AND PLAGIARISM, AND I'LL MENTION THOSE IN THE NEXT

SLIDE OR TWO. AND THEN THERE ARE PROCEDURES THAT TELL DISTRICTS WHAT TO DO--WHAT TO DO IF YOU FIND A STUDENT THAT MAY BE INVOLVED IN SOME TEST SECURITY BREACH, WHAT TO DO IF THERE'S AN ADULT INVOLVED, AND SO FORTH. SOME OF THAT IS SPELLED OUT IN THE POLICY.

REFERRED TO IN THE POLICY BUT A LITTLE MORE HANDS-ON, WE DO QUITE A BIT OF STATE TEST MONITORING. LITERALLY, THAT'S JUST WHAT IT SOUNDS LIKE. WE GET IN OUR CARS AND DRIVE AROUND ACROSS THE STATE. WE PROVIDE TRAINING TO THE STATE DEPARTMENT STAFF WHO DO THE MONITORING, GIVE THEM A LITTLE BACKGROUND ON WHAT THEY SHOULD BE LOOKING FOR. THERE'S A FORM THEY FILL OUT. THOSE ARE REVIEWED ONCE THEY RETURN AND SORTED INTO VARIOUS CATEGORIES OF OFFENSES, FROM MINOR TO MORE SEVERE.

WHO DO WE GO SEE? WELL, WE GO SEE ANYBODY WHO HAS A PRIOR PROBLEM. WE SINGLE THOSE OUT. YOU CAN EXPECT A VISIT FROM THE STATE IF YOU HAD AN ISSUE THE YEAR BEFORE. THE REST OF THE VISITS ARE RANDOM. WE TRY TO MAKE SURE WE ARE IN EVERY SCHOOL DISTRICT, JUST SO THEY KNOW WE'RE COMING. THE VISITS ARE ANNOUNCED IN THAT WE REMIND THEM WE'RE COMING, BUT UNANNOUNCED IN THAT WE DON'T SAY SPECIFICALLY WHICH SCHOOLS WE'RE COMING TO.

SO HOW MANY SCHOOLS CAN THE STATE DEPARTMENT COVER? IT'S A BIG STATE. ABOUT 1,400 PUBLIC SCHOOLS. WE GET TO ABOUT 250-400 SCHOOLS

EVERY YEAR. I THINK OUR LINEUP FOR THIS SPRING, WE'RE GOING TO BE IN ABOUT 300, 325 SCHOOLS. IT TAKES ABOUT 125 PEOPLE TO DO THAT, SO THAT IS QUITE AN EFFORT, BUT THAT HAS BEEN A GOOD THING FOR US. JUST A COUPLE OF WORDS ABOUT PLAGIARISM, AND THEN I'LL MENTION ERASURE ANALYSIS ALSO.

PLAGIARISM, WE DO ASK THE VENDOR TO LOOK FOR CERTAIN TYPES OF JUST WHAT YOU WOULD EXPECT--REPETITIVE INFORMATION ACROSS THE STUDENT RESPONSES. BACK IN THE DAYS OF PAPER, WHEN THEY WERE LOOKING AT THE ACTUAL PAPERS, WE HAD TO MAKE SURE THEY WERE BATCH HEADERS TOGETHER SO THAT IF THERE WAS A PATTERN, IT COULD BE SEEN. THAT'S DONE VIA SCANNER NOW, BUT IT'S THE SAME THING. WE NEED TO TRY TO MAKE SURE THAT THE REPOSSES ARE LOOKED AT BY SOME OF THE SAME INDIVIDUALS.

WE FOLLOW THAT UP AT THE STATE LEVEL WITH A COUPLE OF LAYERS OF REVIEW BEFORE SCORES ARE CANCELED, BUT AFTER THOSE REVIEWS, WE DO IN FACT CANCEL SCORES OR VOID SCORES BASED ON PLAGIARISM. I'LL PAUSE FROM THE SLIDE JUST A SECOND AND SAY, I THINK IF YOU ONLY READ THE ARTICLES ABOUT TEST SECURITY, YOU MIGHT BE LED TO BELIEVE THAT ERASURE ANALYSIS IS TEST SECURITY. AS HAS BEEN MENTIONED HERE A FEW TIMES ALREADY, IT IS JUST ONE TOOL. IT'S A GOOD ONE, AND WE USE IT IN LOUISIANA, BUT IT REALLY IS JUST ONE THING THAT WE CAN DO.

WE DO TYPICALLY LOOK FOR WRONG-TO-RIGHT ERASURES AND PROVIDE A THRESHOLD AND LOOK FOR CLASSROOMS WHERE THAT THRESHOLD IS EXCEEDED. THERE'S A COUPLE MORE WORDS HERE, BUT I THINK THE POINT I REALLY WANT TO MAKE IS THAT WHILE WE DO ERASURE ANALYSIS, AND OTHER STATES DO THE SAME, THE SPECIFIC WAY THAT WE DO IT IS VERY HOMEGROWN. IT'S LOUISIANA-SPECIFIC. WE CAME UP WITH THOSE RULES OURSELVES, AND SHOULD THERE BE MORE STANDARDIZATION THERE? SOMETHING I THINK WE SHOULD PROBABLY TALK ABOUT.

WE ALSO ASK THE DISTRICTS TO DO QUITE A BIT. THEY ARE THERE WHERE THE TESTS ARE GIVEN. THEY ARE CLOSEST TO THE SOURCE OF THE PROBLEM IF THERE IS ONE, AND WHAT THEY ARE ABLE TO DO IS VOID THE SCORES THEMSELVES. WHILE THE STATE IS IN CHARGE OF ERASURE ANALYSES AND PLAGIARISM, THE DISTRICTS CAN FIND MANY OR LOTS OF OTHER THINGS, AND IF THEY DO, THEY ARE ABLE TO REPORT UP TO THE STATE AND CANCEL OR VOID THOSE SCORES, AS WELL. I TALKED FAST SO I COULD LEAVE A COUPLE OF MINUTES AT THE END TO SAY, WHILE WE DO A LOT, I THINK LOUISIANA IS PROBABLY, AS I SAID, TYPICAL, AND MAYBE A LITTLE MORE INVOLVED THAN SOME STATES.

I THINK THERE'S A COUPLE OF AREAS THAT WE WOULD LIKE TO IMPROVE ON, AND SOME OF THESE HAVE ACTUALLY BEEN MENTIONED ALREADY. CHANCELLOR HENDERSON--I WROTE IT DOWN. SHE SAID SHE SPENT A LOT OF TIME DETERMINING WHAT ACTUALLY HAPPENED. WELL, TRUER WORDS, RIGHT? WE

SPEND A LOT OF TIME BETWEEN THE STATE AND DISTRICTS TRYING TO DETERMINE WHAT ACTUALLY HAPPENED, AND WE REALLY RELY ON THE DISTRICTS HEAVILY TO HELP US DETERMINE WHAT ACTUALLY HAPPENED. AND THAT'S A GOOD THING, BECAUSE THEY'RE THERE WHERE THE SOURCE OF THE PROBLEM IS, BUT IN CASES WHERE THE DISTRICT IS NOT AS INTERESTED OR NOT AS CAPABLE OF DOING THAT, I THINK, AT LEAST IN OUR STATE, OUR POLICY IS A LITTLE BIT LIGHT ON THEN WHAT THE STATE SHOULD DO TO INTERVENE AND PROVIDE MORE OVERSIGHT IN THAT INVESTIGATION, OR PERHAPS EVEN BE INVOLVED IN THOSE INVESTIGATIONS.

FOR VERY ESTABLISHED PROCEDURE, SUCH AS ERASURE ANALYSIS--AND I'VE ALLUDED TO THIS ALREADY--I THINK THERE PROBABLY COULD BE MORE CONSISTENCY IN HOW THAT'S DONE. LIKE I SAID, WE HAVE RULES. WE HAVE VERY SPECIFIC RULES BETWEEN OUR STATE AND THE VENDOR, BUT THOSE ARE REALLY HOMEGROWN RULES, AND THERE PROBABLY, PERHAPS, COULD BE A BETTER WAY OR MORE STANDARDIZATION.

THIRDLY, IT'S REALLY THE SAME POINT, JUST APPLIED TO THE OTHER TYPES OF ANALYSES, OTHER THAN ERASURE ANALYSIS. WE'VE WORKED WITH CAVEON, WE'VE TALKED TO OTHER STATES, WE'VE DONE LOTS OF THINGS, AND WHAT I'VE FOUND IS, IF YOU LOOK AT AREAS OUTSIDE OF ERASURE ANALYSIS--FOR EXAMPLE, UNUSUAL GAINS OR LOSSES--THERE MAY BE TOOLS TO DO THAT ANALYSIS, BUT EVEN ONCE YOU HAVE THE ANALYSIS, IT'S VERY DIFFICULT TO KNOW WHAT TO DO NEXT.

AND AGAIN, CHANCELLOR HENDERSON SAID IT VERY WELL--THERE ARE NO WIDELY ACCEPTED PRACTICES ABOUT THOSE KINDS OF THINGS, SO EVEN IF YOU DO THOSE THINGS, IT'S A LITTLE BIT UNCLEAR AT THE STATE LEVEL WHAT YOU SHOULD DO WITH THAT INFORMATION AND WHAT RISES TO THE LEVEL OF INTERVENTION. PERHAPS THE NCME STANDARDS WILL BE A GOOD STEP IN THIS DIRECTION, BUT I DO SUPPORT THE DEVELOPMENT OF SOME WIDELY ACCEPTED TEST SECURITY PRACTICES. THAT WOULD BE VERY HELPFUL TO THE STATE. THANK YOU.

**JACK BUCKLEY:** THANKS VERY MUCH. JIM.

**JIM LIEBMAN:** THANK YOU, JACK.

I'M GOING TO TALK FASTER THAN SCOTT, I PROMISE YOU, TO TRY TO GET WHAT I NEED TO SAY OR WANT TO SAY HERE.

FIRST OF ALL, I THINK IT'S REALLY IMPORTANT TO--LET'S SEE IF I CAN GET THIS ADVANCED HERE; THERE--TO AVOID A COUPLE OF FALLACIES THAT PEOPLE HAVE ALREADY TALKED ABOUT. ONE FALLACY IS THAT BECAUSE CHEATING CAN OCCUR, IT WILL OCCUR, AND IT NEUTRALIZES THE VALUE, THE MANY VALUES OF TESTING THAT PEOPLE HAVE TALKED ABOUT. THAT'S NOT TRUE, AND AS WE ARE HEARING, THERE ARE MANY THINGS THAT YOU CAN DO ABOUT IT.

BUT THERE'S THE REVERSE FALLACY, WHICH IS TO THINK THAT BECAUSE A LOT OF PEOPLE WHO RAISE THE CHEATING ISSUE REALLY DON'T CARE ABOUT CHEATING, THEY JUST DON'T LIKE TESTING OR THEY DON'T LIKE ACCOUNTABILITY MUCH, CAN LEAD YOU TO THINK THAT THERE REALLY ISN'T AN ISSUE THERE, AND THERE IS AN ISSUE THERE, AND YOU NEED TO TAKE IT SERIOUSLY. WHEN YOU PUT STAKES ON OUTCOMES, PEOPLE WILL REACT TO THOSE STAKES, AND ONE OF THOSE REACTIONS CAN BE CHEATING. I THINK THE TWO BASIC ANTIDOTES TO THOSE, BEYOND ALL OF THE SPECIFICS THAT WE'RE GETTING INTO, IS, INFORM YOUR EDUCATORS AND DEMONSTRATE THROUGH YOUR ACTIONS THAT YOU REALLY DO CARE ABOUT CHEATING, AND IF YOU HEAR ABOUT IT, YOU WILL RESPOND TO IT, AND YOU'RE CONSTANTLY TAKING STEPS TO AVOID IT, AND IT'S REALLY IMPORTANT TO SAY THE SAME THING TO THE PUBLIC SO THAT THE PUBLIC UNDERSTANDS.

SO NOW, JUST TO GET INTO A COUPLE OF THE SPECIFIC THINGS WE DO, I'VE USED, IN THE HEADERS HERE, PROFESSOR AMREIN-BEARDSLEY'S TAXONOMY, WHICH I THINK IS VERY HELPFUL, BUT TO GET AT THE REALLY SYSTEMATIC CHEATING, A COUPLE OF THINGS TO DO.

FIRST, TEST ADMINISTRATION. HERE, KEEP THE TEST WINDOW SHORT. IN NEW YORK STATE, THE TEST WINDOW IS LIKE TWO WEEKS. IN NEW YORK CITY, WE KEPT IT TO A DAY OR TWO TO MAKE SURE THAT PEOPLE AREN'T PASSING THE BOOKS AROUND OR TALKING TO EACH OTHER OR WHATEVER. AND

THEN JUST MAKE SURE THAT THE MATERIALS ARE IN THE SCHOOLS FOR THE LEAST AMOUNT OF TIME NECESSARY.

WE SHRINK-WRAP EVERYTHING FOR A SCHOOL INTO ONE PACKAGE. INSIDE THAT PACKAGE IS A SHRINK-WRAPPED PACKAGE FOR EACH CLASSROOM, AND IN THAT SHRINK-WRAPPED PACKAGE FOR EACH CLASSROOM, WE DO SOMETHING WITH THE ELEGANT NAME OF PRE-SLUGGED, WHICH MEANS THAT STUDENTS' NUMBERS ARE ALREADY PUT ONTO THE SHEET. THOSE BIG PACKAGES FOR THE SCHOOL ARE OPENED WITHIN AN HOUR OF THE START TIME, AND THE PACKAGES FOR EACH CLASSROOM ARE OPENED WITHIN 15 MINUTES OF START TIME, AND WE DO MONITORING FROM THE DISTRICT, IN ADDITION TO SOME MONITORING DONE BY THE STATE, TO BE THERE WHEN ALL OF THAT IS SUPPOSED TO HAPPEN AND MAKE SURE THAT IT DOES HAPPEN ON TIME-- AGAIN, UNANNOUNCED MONITORING. WE ONLY GET TO ABOUT 10% OF THE SCHOOLS, BUT MAKE IT VERY, VERY VISIBLE THAT THAT'S ALL HAPPENING, AND THEN OF COURSE GET THOSE MATERIALS OUT OF THE SCHOOL AS QUICKLY AS POSSIBLE. EVERYBODY WHO HANDLES THE MATERIAL HAS TO SIGN SOMETHING SAYING "I GOT XNUMBER, AND I RETURNED X NUMBER, AND IN BETWEEN I DID NOTHING ELSE WITH THEM. I DIDN'T COPY THEM. I DIDN'T DISTRIBUTE THEM IN ANY WAY." THAT'S ALL SIGNED AND CERTIFIED, AND IF THINGS ARRIVE LATE, THE PRINCIPALS HAVE TO EXPLAIN WHY THAT HAPPENED.

ON THE SCORING SIDE OF THIS, I THINK ALSO VERY IMPORTANT TO SCORE OFF-SITE. IF YOU DO USE JUST THE MULTIPLE CHOICE, THAT'S EASY. YOU CAN SCAN THEM OUT OF THE SCHOOL REALLY QUICKLY AND GET THEM SCORED ELECTRONICALLY ELSEWHERE. AS WAS SUGGESTED, WE IN NEW YORK USE A LOT OF CONSTRUCTED RESPONSE, ESSAYS, AND ALL OF THAT. ALL OF THOSE ARE GRADED BY TEACHERS. REALLY IMPORTANT TO GET THAT GRADING OUTSIDE OF THAT SCHOOL AND INTO ANOTHER SCHOOL, SOMETHING WE'VE BEEN DOING WITH THE ELEMENTARY AND MIDDLE BUT JUST HAVE STARTED DOING IN HIGH SCHOOL, AND IT'S REALLY WAY LATE FOR THAT TO HAVE HAPPENED, BUT FINALLY IT IS HAPPENING. EVERYTHING THAT IS READ IS READ BY 2 OR 3 TEACHERS AT LEAST, AND THERE'S A COMPARISON BY A TABLE LEADER OF THE EXTENT TO WHICH THERE'S CORRESPONDENCE BETWEEN WHAT THEY'RE GETTING.

WHERE THERE ISN'T CORRESPONDENCE, THERE'S STEPS THAT ARE TAKEN. THE CITY OF NEW YORK HAS JUST DISTRIBUTED A DISTRIBUTED SCORING RFP THAT YOU MIGHT WANT TO TAKE A LOOK AT, BECAUSE I THINK IT'S PRETTY MUCH STATE OF THE ART, WHAT'S BEING ASKED FOR. WE DON'T KNOW YET WHAT THEY'RE GOING TO GET BACK, BUT ESSENTIALLY NOT TO HAVE SOMEBODY DO THE SCORING, BUT TO HAVE SOMEBODY COME IN AND DO ALL OF THE DISTRIBUTION, PICKING UP, EVERYTHING SCANNED, INCLUDING THE CONSTRUCTED RESPONSES, SO THEY GO IMMEDIATELY INTO THE MACHINES AND CAN'T BE TAMPERED WITH AND CAN BE DISTRIBUTED WHEREVER YOU WANT THEM TO BE GRADED AND CAN HAVE INSTANT, AUTOMATIC, CITY-WIDE

COMPARISON OF HOW EACH SCORE COMPARES TO SCORES ELSEWHERE IN THE CITY SO THAT YOU CAN BE SURE YOU'RE GETTING INTER-RATER RELIABILITY AND THE LIKE.

FOR BOTH FIRST- AND SECOND-DEGREE CHEATING, WHERE IT'S GETTING A LITTLE BIT MORE HAPHAZARD AND MAYBE NOT EVEN INTENTIONAL, WE DO HAVE A TEST COORDINATOR ASSIGNED IN EACH SCHOOL. THAT'S SOMEBODY WHO'S REALLY RESPONSIBLE FOR THIS AND CAN SORT OF WORK WITH THE PRINCIPAL, BUT ALSO MAKE SURE THAT NOT TOO MUCH PRESSURE AND RELIANCE IS PUT ON THE PRINCIPAL.

WE HAVE A TERRIFIC TEST ADMINISTRATION HANDBOOK. I SEE, I THINK, HERE IN THE AUDIENCE THE ORIGINAL AUTHOR OF THAT, WHO'S NOW IN BALTIMORE. IT'S A REALLY VERY CLEAR BOOK THAT EVERYBODY IS SUPPOSED TO READ. ANYONE WHO'S GOING TO PROCTOR OR TEACH, BE INVOLVED IN THE TEST, HAS TO SIGN THAT THEY'VE READ IT, THEY UNDERSTAND IT, THEY KNOW WHAT THE PENALTIES ARE FOR VIOLATING.

ONE POINT ON HERE THAT'S REALLY IMPORTANT--TEACHERS ARE ALLIES. MY SENSE IS THAT TEACHERS REALLY DON'T LIKE CHEATING, FOR THE MOST PART, AND WHEN THEY SEE IT, THEY'RE OFFENDED BY IT. IF YOU MAKE A WAY TO ENABLE THEM TO PROVIDE INFORMATION ABOUT THAT, THEY WILL. CHANCELLOR WALCOTT RECENTLY JUST SAID EMAIL HIM. WE'VE GOT LOTS OF

WAYS YOU CAN DO THAT IN NEW YORK, BUT YOU CAN ALSO EMAIL THE CHANCELLOR.

SO I THINK YOU PROVIDE ANONYMITY WHERE YOU NEED TO. OFFER CONFIDENTIALITY WHERE THAT IS HELPFUL. PROCTORING IS A REALLY IMPORTANT POINT. AS WAS POINTED OUT ALREADY, WE HAVE A LOT OF MATERIALS ABOUT THAT, FAQs AND THINGS LIKE THAT. BASICALLY, THE TEACHER SHOULD TALK FROM THE FRONT OF THE ROOM ABOUT WHAT THE PROCEDURES ARE, SHOULD CIRCULATE, BUT WHEN CIRCULATING, SHOULD NOT HAVE PENS, PENCILS, ANY MATERIALS AT ALL IN HER HAND, SHOULD TRY TO KEEP THE CONVERSATION WITH STUDENTS TO A MINIMUM AND FROM THE FRONT OF THE ROOM, AS OPPOSED TO WITH INDIVIDUALS. THIS IS ANOTHER THING THAT WE MONITOR WHEN WE GO AROUND WITH THIS UNANNOUNCED MONITORING.

EVERY ROOM WHERE A TEST IS TAKING PLACE IS EITHER SUPPOSED TO HAVE WINDOW IN THE DOOR THAT IS NOT BLOCKED, OR THE DOOR IS AJAR ENOUGH SO THAT THE MONITOR CAN LOOK IN AND SEE WHAT'S TAKING PLACE TO MONITOR THE PROCTORING.

I'M NOT GOING TO TALK SO MUCH ABOUT--THESE ARE SOME OPEN QUESTIONS. A LOT OF PEOPLE HAVE TALKED ABOUT THE EXTENT TO WHICH ERASURE ANALYSIS IS VALID AND USEFUL, BUT THE EXTENT TO WHICH IT IS NOT AND THE INTER-RATER RELIABILITY ANALYSES, ISSUES ABOUT WHETHER

PROCTORS SHOULD BE FROM THE CLASSROOM OR NOT--IN OTHER WORDS, THE TEACHER IN THAT CLASSROOM OR NOT, AND AT WHAT AGE YOU CAN CHANGE THAT.

I WANTED TO TALK JUST A MOMENT ABOUT THIRD-DEGREE CHEATING BECAUSE I THINK IT GETS LESS ATTENTION THAN IT DESERVES, AND IT IS SOMETHING THAT CAN BE DEALT WITH. IF YOU GIVE TESTS OVER TIME, YOU CAN EXPECT THAT THERE WILL BE SLIPPAGE AND THAT THE TESTS WILL GET EASIER OR THAT THE SCORES WILL GET HIGHER OVER TIME. THIS IS A COMBINATION OF ALL OF THE OTHER KINDS OF CHEATING, AS WELL AS THE SLIPPAGE IN TERMS OF TEST ITEMS AND TEST PREP AND THINGS LIKE THAT.

THERE'S JUST WAYS TO LOOK OUT FOR THIS AND RESPOND TO THIS. I'VE CALLED THEM THE KORETZ TESTS BECAUSE DAN KORETZ AT HARVARD HAS DEVELOPED SOME OF THESE. ESSENTIALLY WHAT YOU DO IS YOU ANCHOR YOUR EXPECTED OUTCOMES TO SOMETHING THAT IS NOT SUBJECT TO CHEATING, LIKE HERE THE A.C.T. YOU CAN USE NAEP. YOU CAN USE GRADUATION RATES. THEN YOU TRY TO SEE IF IMPROVEMENTS IN TEST SCORES ARE MATCHED BY IMPROVEMENTS IN OTHER THINGS THAT SHOULD IMPROVE SOMEWHAT SIMILARLY.

THIS IS FROM KENTUCKY WAY BACK IN THE NINETIES. IT'S NOT TO VILIFY ANYBODY, BUT YOU CAN SEE THAT THERE WAS A TON OF SLIPPAGE HERE IN TERMS OF WHAT THE STATE TEST WAS SHOWING--THE DOTTED LINE AT THE

TOP--AND THE A.C.T. TEST WAS PRETTY STABLE OR GOING DOWN EVEN, AND THE DIFFERENCE BETWEEN THE TWO IS SUGGESTIVE OF A LOT OF SLIPPAGE THAT SHOULD RAISE SOME PROBLEMS. WE ASKED DAN KORETZ TO COME INTO NEW YORK CITY TO SEE AND ESSENTIALLY TO ANCHOR--WE LOOKED AT OUR EIGHTH-GRADE TEST SCORES AND LOOKED TO SEE IF THEY PREDICTED GRADUATION RATES 4 YEARS LATER SIMILARLY FROM ONE YEAR TO THE NEXT.

IF YOU START TO SEE THAT THE SAME GRADE PREDICTS A LOWER GRADUATION RATE, THEN YOU'VE GOT SLIPPAGE. I'M HAPPY TO SAY THAT THERE WASN'T MUCH FOUND. THIS WAS THE WORST OF IT, EIGHTH-GRADE MATH. YOU CAN SEE AT THE TOP, FOR THE HIGHEST-PERFORMING STUDENTS, IN ONE YEAR, THE LAST OF THE 3 YEARS, THERE WAS A LITTLE BIT LOWER GRADUATION RATE THAN YOU WOULD HAVE EXPECTED. THIS IS NOT A SERIOUS PROBLEM HERE, BUT IT CERTAINLY WAS ONE THAT STARTED US THINKING ABOUT WHAT WE COULD DO ABOUT IT, SO THERE ARE WAYS TO MONITOR THIS, AND I ENCOURAGE PEOPLE TO THINK ABOUT THOSE.

**JACK BUCKLEY:** THANK YOU VERY MUCH.

SO I'D LIKE TO SHIFT TO A MODERATED DISCUSSION JUST WITH THE PANEL BEFORE I OPEN IT UP TO THE BROADER AUDIENCE. IN CASE YOU DIDN'T NOTICE, THERE'S AN AWFUL LOT THAT JUST GOT COVERED. IF I COULD THINK OF A PLACE TO START, I GUESS IT WOULD BE WHETHER OR NOT WE CAN REACH SOME KIND OF CONSENSUS ON--I THINK WE STARTED TO PARTITION

THE WORLD INTO DIFFERENT TYPES OF CHEATING, OR AS GREG POINTED OUT, CHEATING DOES NOT EQUAL IRREGULARITIES, WHICH I THINK IS ALSO ANOTHER IMPORTANT POINT.

FROM THE STANDPOINT OF THE DEPARTMENT OR NCES, WE WERE THINKING REALLY MORE ABOUT, AUDREY, WHAT YOU CALL FIRST- AND SECOND-DEGREE CHEATING OR IRREGULARITIES, ALTHOUGH THE THIRD DEGREE IS VERY INTERESTING FROM A POLICY PERSPECTIVE. IF WE CAN STAY--WHAT ARE YOUR THOUGHTS AS A PANEL ON WHETHER OR NOT IT'S FAIR TO SAY THAT SOMETHING LIKE THE FIRST DEGREE, WHAT WE'RE REALLY TALKING ABOUT WITH PREVENTION, ARE OPERATIONAL MEASURES--THINGS LIKE CHAIN OF CUSTODY, SHRINK-WRAPPING. I WOULD PUT PROCTORING IN THAT CATEGORY.

THE SECOND DEGREE REALLY IS MORE, I THINK, WHAT A COUPLE OF FOLKS TOUCHED ON AS CULTURE, AS SOMETHING WHERE YOU'VE GOT TO PUT OUT LIKE SCOTT'S PROCEDURES OR POLICIES AND WHERE WE NEED TO THINK ABOUT THE ENTIRE SYSTEM OF ASSESSMENT AND NOT JUST SORT OF, "WHAT DID I DO FOR THIS OPERATIONAL STEP?"

DOES THAT SEEM LIKE A FAIR TYPOLOGY, OR WOULD SOMEBODY SUGGEST SOMETHING DIFFERENT OR MORE USEFUL?

**GREG CIZEK:** I MIGHT SAY THAT THE CULTURE ISSUE WOULD CUT ACROSS ALL 3 OF THOSE LEVELS. I THINK THERE'S A SENSE THAT SOMEHOW--IT'S

SOMEWHAT OF AN UNFAIRNESS. 179 DAYS OF THE YEAR, YOU ASK A TEACHER TO BE AN ADVOCATE FOR A KID, YOU ASK PRINCIPALS TO BE FOR THE KID'S SUCCESS AND ALL THAT KIND OF THING, AND THEN ON A DAY WHEN THERE'S A HIGH-STAKES TEST GIVEN, YOU ASK THEM TO BE AN INVIGILATOR, A PROCTOR, SOMEONE WHO IS NOT THERE ADVOCATING FOR THE GREATEST SUCCESS FOR THE STUDENT. THAT CULTURE CUTS ACROSS AND ENGENDERS ALL 3 OF THOSE LEVELS OF INAPPROPRIATE PRACTICE.

**JIM LIEBMAN:** I WOULD AGREE WITH THAT. I THINK THERE'S A PROPORTION OF BOTH. I THINK YOUR PROPORTIONS ARE RIGHT, THAT PROCEDURES AND OPERATIONS ARE MORE IMPORTANT FOR THE FIRST DEGREE, BUT I THINK CULTURE CUTS ACROSS THERE AS WELL, AND I THINK, EVEN AT THE CLASSROOM LEVEL, ONE OF THE THINGS THAT OUR HANDBOOKS DO THAT ARE REALLY GOOD IS JUST AN FAQ THAT JUST RAISES ALL THE THINGS THAT CAN HAPPEN AND JUST GOES THROUGH IT, SO A TEACHER HAS READ THAT AND SORT OF IS READY FOR, YOU KNOW, THE KID RAISES HIS HAND, HE WANTS TO GO TO THE BATHROOM. WELL, WHAT DO YOU DO ABOUT THAT? I THINK JUST TO HAVE SOME PROCEDURES THAT YOU'VE THOUGHT ABOUT BEFOREHAND EVEN FOR THAT KIND OF THING REALLY MAKES IT EASIER FOR PEOPLE TO JUST FEEL LIKE THEY'RE DOING THE RIGHT THING.

**SCOTT NORTON:** I'LL JUST BE BRIEF. I LIKE THE TAXONOMY. I THINK THAT'S HELPFUL AND A GOOD WAY OF THINKING ABOUT IT. FOR THE SECOND LEVEL IN OUR STATE, WE DO SOME OF THAT. WE IMBED THAT IN THAT NOT-SO-

EASY-TO-READ TEST ADMINISTRATION MANUAL THAT--I THINK HE READ OUR MANUAL--THAT HE TALKED ABOUT, BUT I THINK WE COULD PROBABLY DO A BETTER JOB OF SHOWING FOLKS HOW THAT WORKS, WEBINARS OR HANDS-ON TRAINING IN WHAT IS CROSSING THE LINE, IF YOU WILL.

**AUDREY AMREIN-BEARDSLEY:** I'D LIKE TO ECHO WHAT GREG HAD SAID BEFORE ABOUT THE ROLES OF THE TEACHER EDUCATION COLLEGES IN THIS. I THINK THAT'S REALLY IMPORTANT. WE CAN TRAIN OUR FUTURE TEACHERS TO DO A MUCH BETTER JOB. MANY OF OUR COLLEGES OF EDUCATION HAVE DROPPED THE ASSESSMENT COURSES THAT USED TO EXIST, AND SO I THINK ALSO TURNING THAT INTO A PROFESSIONAL RESPONSIBILITY--NOT WHERE YOU'RE A POLICE OFFICER, BUT THIS REALLY IS YOUR PROFESSIONAL RESPONSIBILITY TO LET YOUR STUDENT SHINE, AND YOU WANT TO EMBRACE THAT, VERSUS TRYING TO CASUALLY HELP THEM AND SO FORTH. THIS IS SOMETHING THAT'S MORE OF A PROFESSIONAL NEED ON THEIR PARTS.

**JACK BUCKLEY:** JUST TO GET BACK TO SOMETHING YOU SAID, GREG, YOU MENTIONED OR BRIEFLY TOUCHED ON AN IDEA ABOUT SOME KIND OF A THIRD-PARTY QUALIFICATION TOOL OR CERTIFICATION SYSTEM, FOR LACK OF A BETTER WORD. WE'VE HEARD JIM'S EXAMPLE OF A VERY DETAILED ADMINISTRATION GUIDE, SCOTT'S EXAMPLE OF ANOTHER VERY DETAILED SORT OF STATE LEVEL GUIDE, AND YOU ALSO BRIEFLY TOUCHED ON THE NCME GUIDANCE, WHICH WE'D LIKE TO HEAR MORE ABOUT, I'M SURE. SO WHERE IN THOSE DIFFERENT LEVELS WOULD SUCH A QUALIFICATION OR CERTIFICATION

SYSTEM FALL? WHAT SHOULD IT COVER? WHAT WOULD YOU LIKE TO SEE IN SOMETHING LIKE THAT?

**GREG CIZEK:** I WISH IT WERE UNNECESSARY BECAUSE, LIKE AUDREY, I WISH SCHOOLS OF EDUCATION WERE DOING A BETTER JOB PREPARING PEOPLE IN THEIR PROFESSIONAL RESPONSIBILITIES. WHERE THEY'RE NOT, I THINK IT WOULD BE AN APPROPRIATE THING FOR A STATE TO WORK WITH A VENDOR TO HELP WITH SOME SORT OF PROFESSIONAL DEVELOPMENT TOOL THAT WOULD ACTUALLY BE REQUIRED FOR PEOPLE TO PARTICIPATE IN, IN ORDER TO BE A QUALIFIED TEST ADMINISTRATOR. IT COULD COVER EVERYTHING LIKE WHAT'S PROPER FACILITATION OF STUDENT PERFORMANCE, WHAT'S INAPPROPRIATE INTERVENTION IN STUDENT PERFORMANCE--THE ENTIRE GAMUT OF THINGS COVERED IN, FOR EXAMPLE, THE NCME DOCUMENT THAT YOU HAVE, AS WELL AS OTHER GUIDANCE. I'M INTRIGUED TO SEE THE FAQ THAT YOU HAVE PUT TOGETHER. I'VE NOT SEEN THAT, BUT I SUSPECT THERE'S A LOT OF GOOD THINGS IN THERE THAT COULD BE COVERED, AS WELL.

**AUDREY AMREIN-BEARDSLEY:** IT WOULD POSSIBLY EVEN BE INCLUDED IN THE CERTIFICATION TYPES OF EXAMS, AS WELL, SO WHEN THE FUTURE TEACHERS ARE TESTED ON CONTENT AND PEDAGOGICAL SKILLS, IT WOULD POSSIBLY FIT SOMEWHERE IN THAT TYPE OF ADMINISTRATION, AS WELL.

**GREG CIZEK:** I ACTUALLY WANT THIS TO BE REAL FUZZY AND WARM AND FRIENDLY, AND I WANT PEOPLE TO DO THIS IN THEIR PJs AT HOME ON THEIR

HOME COMPUTER, QUALIFYING IN THEIR OWN TIME, AND NOT HAVE IT BE AN ONEROUS KIND OF THING THAT WOULD BE OPPRESSIVE BECAUSE I THINK TESTING FOLKS HAVE A HARD ROAD TO GO IN TERMS OF HELPING PEOPLE TO SEE THAT IT'S NOT AN OPPRESSIVE, BIASED, DISCRIMINATING, MARGINALIZING, CONTROLLING ACCESS TO WEALTH AND POVERTY IN AMERICA, BLAH, BLAH, BLAH--YOU KNOW, ALL THAT STUFF THAT TESTS ARE ACCUSED OF. I THINK WE HAVE A LONG WAY TO GO TO HELP PEOPLE SEE THAT THERE ARE REALLY BENEFITS OF GETTING QUALITY INFORMATION ABOUT TESTS AND THAT EDUCATORS HAVE A ROLE IN HELPING TO ENSURE THAT QUALITY.

**JACK BUCKLEY:** SO IF I WERE WORKING ON AN SEA AND WANTED TO TAKE THE NEXT STEPS, LET'S SAY, IN TERMS OF PREVENTION--AGAIN FOCUSING ON PREVENTION--WHAT ARE SOME OTHER EXAMPLES? WHO SHOULD I TURN TO IN TERMS OF IDENTIFYING BEST PRACTICES WITH RESPECT TO THIS KIND OF MANUAL OR TRAINING, IF I WANTED TO IMPLEMENT A REQUIREMENT LIKE THAT, MAYBE WITHOUT TURNING TO A VENDOR FIRST? I THINK WE'VE HEARD A COUPLE TODAY, BUT DOES ANYONE HAVE ANY OTHER IDEAS ABOUT TO WHOM TO TURN?

**SCOTT NORTON:** I DO WANT TO TIE THAT BACK TO THE LAST COMMENT. I THINK THAT'S AN INTRIGUING IDEA, AND IT'S SOMETHING WE MAY TRY TO DO BEFORE IT'S COMMON PRACTICE. WE DO REQUIRE THAT TEST PROCTORS AND TEST MONITORS BE TRAINED, BUT THAT IS A VERY LOOSE DEFINITION, AND

EVEN DESPITE OUR EFFORTS, WE KNOW THAT IT'S PROBABLY INADEQUATE. SOME TYPE OF CERTIFICATION ONLINE--YOU GUIDE YOURSELF THROUGH A TRAINING OR SOMETHING LIKE THAT--COULD BE DEVELOPED, I THINK, RELATIVELY EASY ENOUGH, AND THEN FOLKS WOULD BE REQUIRED TO DO THAT JUST SO THAT THEY WOULD BE AWARE WHAT'S NEEDED. I DON'T THINK IT'S BEEN EXPLICITLY STATED YET, SO I'LL GO AHEAD AND MENTION, I THINK WHAT WE'RE DOING HAS BEEN ADEQUATE ENOUGH FOR THE PAST FEW YEARS, BUT WITH THE ADVENT OF THE HIGHER-STAKES TEACHER EVALUATION SYSTEMS THAT ARE UPON US--SOME ALREADY HAPPENED, SOME HAPPENING SOON--I THINK WHAT HAS BEEN ADEQUATE IN THE PAST IS SOON GOING TO BE INADEQUATE.

**AUDREY AMREIN-BEARDSLEY:** MAYBE YOU COULD USE KIND OF WHAT GREG WAS TALKING ABOUT, TOO--THE INSTITUTIONAL REVIEW BOARD PROCEDURES THAT WE DO AT THE UNIVERSITY LEVEL TO DO RESEARCH. MAYBE THAT WOULD BE SOMETHING THAT WOULD ALSO BE CONSIDERED. GO THROUGH AN IRB TYPE OF PROCESS, LEARN ABOUT THE HISTORY OF THIS, AND ACTUALLY EARN YOUR CERTIFICATE TO ACTUALLY BE A PROCTOR OR SOMEBODY IN CHARGE OF A CLASSROOM THAT'S DELIVERING TESTS.

**GREG CIZEK:** YOU KNOW, THERE'S A DATABASE OF THOSE, RIGHT? ONCE YOU QUALIFY, YOU QUALIFY, AND AN EDUCATOR AND A STATE COULD QUALIFY. THERE WOULD BE A DATABASE THAT WOULD SAY, NO MATTER DISTRICT YOU'RE

IN, YOUR NAME IS LOGGED AS SOMEONE WHO'S COMPLETED THE QUALIFICATION, AND YOU'RE GOOD TO GO.

**JACK BUCKLEY:** ALL RIGHT. JUST TO SWITCH GEARS--ANOTHER THING THAT, SCOTT, BOTH YOU AND JIM DISCUSSED, AND I THINK MAKES A LOT OF SENSE, IS AN INSPECTION PROGRAM, SOME SORT OF MONITORING BEYOND JUST, SAY, THESE PROCTORS. LET'S SAY THEY'RE BETTER TRAINED OR THEY'RE CERTAINLY AWARE OF WHAT THEY'RE DOING INSIDE THE ADMINISTRATION ROOM. I THINK THERE'S AN ISSUE HERE IN TERMS OF COST. SCOTT, I THINK YOU MENTIONED SOMETHING LIKE 125 STAFF? I ASSUME THAT'S NOT THEIR FULL-TIME JOB, MONITORING, ALTHOUGH IT COULD BECOME ONE, I'D IMAGINE.

**SCOTT NORTON:** I'LL JUST SAY A WORD ABOUT THAT. THOSE ARE 125 PEOPLE WHO HAVE OTHER JOBS MOST OF THE YEAR. FOR A WEEK, WE GET THEM TRAINED UP AND SEND THEM OUT INTO THE FIELD TO HELP MONITOR THESE TESTS. WE TRY TO RELY ON THE FOLKS WHO HAVE EDUCATIONAL BACKGROUNDS AND ARE AT LEAST USED TO BEING AT SCHOOLS FOR SOME PURPOSE. I WAS TALKING FAST, BUT I'LL SAY JUST A WORD MORE ABOUT THAT. THAT'S A RELATIVELY NEWER THING THAT WE'VE STARTED DOING FROM THE STATE LEVEL, AND I THINK IT'S BEEN A GOOD THING. WE TRY TO BE UNOBTRUSIVE AND, YOU KNOW, NOT HEAVY-HANDED ABOUT IT. I THINK IT DOES SEND A STRONG SIGNAL WHEN A STATE MONITOR WALKS IN THE DOOR OF THE SCHOOL THE MORNING OF TESTING, AND YOU MAY OR MAY NOT BE EXPECTING THEM TO

BE THERE. FRANKLY, I'D LIKE TO SEE US DO A LITTLE BIT MORE OF THAT. IT'S A PRETTY MECHANICAL, GET IN YOUR CAR, DRIVE-AROUND-THE-STATE KIND OF THING, BUT IT HAS SERVED A GOOD PURPOSE FOR US, I BELIEVE.

**JIM LIEBMAN:** IN NEW YORK, WE DO THE SAME THING. WE USE CENTRAL STAFF WHO VOLUNTEER FOR THIS, SO IT DOESN'T COST ANYTHING FOR THOSE PEOPLE. IT COSTS THEM SOMETHING. WE ACTUALLY USE SUBWAY RATHER THAN CAR, BUT WHAT WE TRY TO DO IS TO ASSIGN PEOPLE TO SCHOOLS THAT ARE NEAR WHERE THEY LIVE. BECAUSE WE WANT THEM TO GET THERE VERY EARLY IN THE MORNING, WHEN THE SHRINK-WRAP IS OPENED, WE WANT THEM TO START OUT PRETTY CLOSE TO WHERE THEY ARE. WE'VE GOTTEN TO THE POINT WHERE THIS IS MUCH MORE REGULARIZED AND THERE'S A VERY GOOD CHECKLIST NOW AND EVERYBODY IS TRAINED AND EVERYBODY DOES THE SAME THING. IN THE EARLY YEARS, IT WAS NOT AS RIGOROUS AS IT IS NOW. I ACTUALLY THINK THAT WHERE YOU COULD GET A COMBINATION OF STATE AND LOCAL MONITORING, SO THAT IF THE STATE CAN COVER ITS QUARTER OR 30% AND THE DISTRICT IS COVERING ANOTHER 10% OR 15%, YOU CAN GET INTO AN AWFUL LOT OF SCHOOLS, AND MOST SCHOOLS CAN EXPECT TO SEE SOMEBODY EVERY FEW YEARS AT LEAST, AND I THINK THAT'S REALLY THE MAIN THING, THAT ONCE YOU SEE THAT HAPPENING, IT HAS A BIG EFFECT, AND I THINK IT'S IMPORTANT TO CIRCULATE AROUND THE SCHOOL SO THAT EVERYBODY IN THE SCHOOL WHO'S INVOLVED IN THE TESTING DOES SEE IT WHEN YOU'RE THERE.

**JACK BUCKLEY:** WELL, GIVEN THE TIME, I THINK AT THIS POINT WE'D LIKE TO OPEN IT UP BOTH TO THE OTHER SETS OF PANELISTS BUT ALSO TO THE AUDIENCE AND TO ANYBODY ONLINE VIEWING THE WEBCAST, SO PLEASE, FEEL FREE.

[INDISTINCT]

[SPEAKER'S MICROPHONE IS NOT ON]

**WAYNE CAMARA:** A POINT I WANTED TO MAKE BRIEFLY IS JUST THAT I THINK WE DO KNOW A LOT ABOUT SOME OF THE BEST PRACTICES AND THE DEVELOPMENTAL POLICIES THAT SHOULD BE ESTABLISHED. MORE CAN BE DONE THERE, BUT I THINK ONE OF THE AREAS WE REALLY HAVE TO LOOK CLOSELY AT IS IMPLEMENTATION, NOT DEVELOPMENT.

I WORK WITH A NUMBER OF STATES, AND THERE'S A LOT OF VARIATION IN TERMS OF STATE AUTHORITY, BOTH LEGAL AND POLITICAL, AND WORKED WITH ONE STATE WHERE THEY'VE HAD A NUMBER OF THESE KIND OF CHEATING ISSUES, AND ONE OF THE FIRST QUESTIONS I ASKED OF THE TAC MEMBER WAS, DO WE HAVE SOME GOOD DATA ON THE TRAINING OF PROCTORS AND TEST ADMINISTRATORS IN THAT STATE WHO ACTUALLY WENT THROUGH THE TRAINING? AND I WAS TOLD THAT THEY CAN'T ENFORCE TRAINING. YOU CAN'T MAKE AN LEA TEST COORDINATOR DO ANYTHING. WE CAN ENCOURAGE THEM TO GO TO TRAINING, WE CAN ENCOURAGE PROCTORS TO TAKE THE

WEBEX, BUT WE CAN'T MAKE THEM DO THIS. AND I THOUGHT, OF ALL THE THINGS TO NOT BE ABLE TO MANDATE, HOW CAN YOU MANDATE THE TEST BUT NOT THE ADMINISTRATIVE PROCESSES? SO THE WHOLE POINT--I THINK SOME OF THE EXAMPLES AUDREY HAD IN TERMS OF THE MATERIALS IN THE CLASSROOM AND WHAT YOU CAN DO AND WHAT YOU CANNOT DO, THAT CUTS ACROSS LOCAL AUTHORITY AND THE DESIRE FOR LOCAL FLEXIBILITY, AND THE WHOLE IDEA WITH ASSESSMENTS IS FOUNDED IN STANDARDIZATION. FAIRNESS IS FOUNDED IN EQUAL OPPORTUNITY, EQUAL ACCESS, AND PLAYING BY THE SAME RULES, AND IT'S NOT FAIR TO HAVE A LOCAL DISTRICT, FOR WHATEVER REASON--LOCAL CONTRACT, LOCAL NEGOTIATION WITH THE TEACHERS UNION, LOCAL POLITICAL POSITIONS--HAVE AUTHORITY TO DICTATE WHAT IS ON THE WALLS IN ONE SCHOOL, WHEN ANOTHER SCHOOL IS INTERPRETING IT DIFFERENTLY. I'M NOT SURE TO WHAT EXTENT WE'RE TALKING ABOUT THE LACK OF GOOD ADMINISTRATIVE PROCEDURES VERSUS THE POLITICAL IMPERATIVE IN NOT ALL STATES, BUT SOME STATES, TO HAVE AN AWFUL LOT OF LOCAL CONTROL WHERE THE STATE HAS NO AUTHORITY, AND I THINK AS LONG AS THOSE ISSUES ARE NOT RESOLVED, WE'RE NOT GOING TO GET TO THE BOTTOM OF THIS.

THIS IS ALL ABOUT PREVENTION, NOT ABOUT DETECTION. IF WE'RE NOT ABLE TO MANDATE GREATER STANDARDIZATION, WILL WE EVER BE ABLE TO REALLY PREVENT SOME OF THE THINGS WE'RE SEEING?

**JIM LIEBMAN:** AND THAT'S WHERE...I THINK IT'S TRUE THAT DIFFERENT STATES HAVE THE DIFFERENT LAWS, BUT WHERE THERE'S A WILL, THERE'S A WAY, AND I THINK A LOT OF IT COMES DOWN TO THE STATES REALLY ACCEPTING RESPONSIBILITY. I THINK THE FEDERAL GOVERNMENT HAS SHOWN THAT YOU CAN USE THE DISTRIBUTION OF MONEY AND GRANTS AS A WAY TO GET PEOPLE TO DO THINGS THAT YOU CAN'T REQUIRE THEM TO DO. YOU CAN ALSO JUST USE TRANSPARENCY--YOU KNOW, WHO'S DOING IT AND WHO'S NOT, AND THESE ARE BEST PRACTICES, AND YOU CAN GO LOOK AND SEE WHO'S DOING IT. I'VE JUST WATCHED IN THE STATE OF NEW YORK, OVER THE LAST 5 OR 6 YEARS, THAT THERE'S JUST BEEN A TREMENDOUS CHANGE IN THEIR SENSE OF RESPONSIBILITY FOR THIS, AND AS THEY'VE INCREASED THEIR SENSE OF RESPONSIBILITY, THEY'VE FOUND MANY WAYS TO INCREASE THE ABILITY TO ACHIEVE TEST SECURITY WITHOUT REALLY CHANGING THE LAW AT ALL.

**JACK BUCKLEY:** ANYONE ELSE ON THE PANEL COMMENT? REALLY WHAT YOU'VE INTRODUCED, FOR THE FIRST TIME, THAT WE HAVEN'T TALKED ABOUT YET, IS THE QUESTION OF BARRIERS, WHAT'S OUT THERE THAT'S IN THE WAY OF PREVENTION.

WE'RE THINKING SORT OF FROM AN OPERATIONAL AND CULTURAL STANDPOINT, BUT YOU RAISE THE POLITICAL ECONOMY ASSESSMENT. WHAT ELSE IS OUT THERE THAT WE'VE SEEN, THAT STATES AND DISTRICTS HAVE EXPERIENCED, IN TERMS OF BARRIERS TO PREVENTION, BUT THEN ALSO WAYS

THAT THEY'VE OVERCOME? CERTAINLY NEW YORK IS NOT A STRANGER TO POLITICAL PRESSURE, AND CLEARLY, YOU WERE ABLE TO OVERCOME, I IMAGINE, JIM, A BROAD SET OF OPPOSITION.

**JIM LIEBMAN:** YEAH. I MEAN, I THINK THAT THERE PROBABLY ARE, IN ANY GIVEN STATE, REALLY GOOD THINGS HAPPENING IN CERTAIN DISTRICTS, AND I THINK TO ESTABLISH A SORT OF RACE TO THE TOP IDEA, THAT IF IT'S HAPPENING IN SOMEPLACE AND IT'S A GOOD STANDARD, THEN IT OUGHT TO BE THE STANDARD FOR EVERYBODY ELSE, AT LEAST FOR DISTRICTS THAT ARE COMPARABLE IN SIZE OR WHATEVER OTHER CONSTRAINTS THEY MIGHT HAVE.

BUT I THINK THE STATES CAN REALLY DO A LOT, AND I THINK A FEW YEARS AGO, SOME OF US IN NEW YORK FELT LIKE THE STATE WAS HOLDING THE DISTRICTS BACK IN CERTAIN WAYS. WE REALLY WANTED TO GET THOSE HIGH-SCHOOL TESTS GRADED OUTSIDE THE HIGH SCHOOL, BUT SINCE THE STATE HAD BLESSED THEM BEING GRADED BY THE TEACHERS IN THE CLASSROOM, IT WAS VERY HARD TO MAKE THAT CHANGE AT THE LOCAL LEVEL. SO IT CAN KIND OF GO BOTH WAYS.

**JACK BUCKLEY:** JOHN?

**JOHN FREMER:** I'VE HAD THE EXPERIENCE OF AUDITING THE SECURITY PRACTICES

OF QUITE A NUMBER OF STATES AND DISTRICTS. WHEN WE GET TO THE END, WE MAKE RECOMMENDATIONS, AND ONE OF THE THINGS THAT ALMOST ALWAYS IS ASKED OF ME, IS THERE ANYTHING THAT'S EASY TO DO AND IMPORTANT? THERE ARE A LOT OF THEM THAT AREN'T, BUT ONE THAT PRETTY MUCH IS ALWAYS POSSIBLE--IT'S A GOOD ONE FOR ME TO MENTION--IT'S VERY COMMON FOR SENIOR PEOPLE, PEOPLE LIKE THE CHANCELLOR, TO TALK ABOUT, OUR GOAL IS IMPROVING TEST SCORES. SHE DIDN'T SAY THAT, BUT IT'S VERY COMMON TO READ THAT, AND YET, THAT'S NOT THE GOAL. THE GOAL IS TO HELP STUDENTS LEARN MORE AND BE BETTER EQUIPPED FOR WHAT THEY WILL DO IN LIFE. SO TO WORK ON THAT, TO NOT JUST CONTINUE WORKING WITH THOSE WHO ARE ADVISING THE PERSON IN CHARGE, TO GET US TO ALWAYS TALK ABOUT HELPING STUDENTS LEARN MORE AND NOT TO TALK ABOUT IMPROVING TEST SCORES, BECAUSE SOME OF THE THINGS THAT IMPROVE TEST SCORES ARE EXACTLY THE THINGS YOU'RE WORRIED ABOUT. THE THINGS THAT WILL IMPROVE LEARNING, WE CAN ALL GET BEHIND.

**JIM LIEBMAN:** I'D SAY SOMETHING ABOUT THAT. I THINK THAT THAT'S EASY AND YOU NEED TO DO IT, BUT THERE ARE HARD PARTS ABOUT THAT THAT ARE JUST AS IMPORTANT, WHICH IS TO DEVELOP A CULTURE IN THE CLASSROOM THAT THE DATA IS REALLY USEFUL AND HELPFUL, AND ONCE THAT HAPPENS, THEN THE RELIABILITY OF THAT DATA BECOMES IMPORTANT, INCLUDING TO EVERYBODY WHO'S USING IT, AND THEN ANYTHING THE EDUCATORS UNDERSTAND TO AFFECT THE RELIABILITY CAN BECOME A PROBLEM FOR THEM, AND AT THAT POINT, I THINK IT'S EASIER TO GET PEOPLE TO UNDERSTAND

WHY IT'S IMPORTANT, ESPECIALLY FOR SOME OF THE SECOND-DEGREE KINDS OF THINGS THAT ARE EASIER TO LET SLIP.

**JACK BUCKLEY:** ANYONE ELSE ON THE PANEL WITH A COMMENT ON THAT ONE?

**AUDREY AMREIN-BEARDSLEY:** WELL, I'LL PANEL--"I'LL PANEL." I WILL RESPOND AS A PANELIST.

**JACK BUCKLEY:** YOU'RE IMPANELED ALREADY.

**AUDREY AMREIN-BEARDSLEY:** I'LL PANEL ON THIS ONE.

I WOULD AGREE WITH THAT, TOO, BECAUSE I THINK THAT FITS INTO THE TAXONOMY, THE LAST POINT. IF PROFESSIONALLY, AS A CULTURE, WE CAN ALWAYS REMEMBER THAT IN THE END, IT'S ABOUT THE STUDENT LEARNING, AS INDICATED BY THE ACHIEVEMENT TEST, THAT'S REALLY WHERE WE WANT TO GO. PROFESSIONALLY, EVEN PROFESSIONAL DEVELOPMENT WITH TEACHERS AND ADMINISTRATORS, IS ALWAYS ABOUT THE STUDENT LEARNING. SO WHEN YOU FEEL TEMPTED TO MARGINALIZE THE GRAPHING OF INEQUALITIES BECAUSE YOU KNOW IT'S NOT GOING TO BE ON THE TEST, YOU'VE GOT TO GO BACK TO THAT FUNDAMENTAL LAW-- IS THIS ABOUT STUDENT LEARNING? AS INDICATED, HOPEFULLY, BY THE STUDENT TEST, THE ACHIEVEMENT TEST.

**JACK BUCKLEY:** I'M SORRY. JUST A REMINDER TO OTHER FOLKS AT THE TABLE THAT DON'T HAVE THESE MICROPHONES, JUST TO ASK YOU TO PLEASE PUSH TO TALK IF YOU'RE GOING TO SPEAK.

DO WE HAVE A QUESTION FROM THE WEB AUDIENCE?

**JAMES ELIAS:** WE DO. SEVERAL PEOPLE WHO ARE WATCHING THE WEBCAST WOULD LIKE TO KNOW ABOUT THE FEASIBILITY OF THIRD-PARTY CERTIFICATION, GIVEN THE FACT THAT TESTING WINDOWS ARE SO SHORT IN MOST STATES.

**GREG CIZEK:** I CAN ANSWER THAT.

**JACK BUCKLEY:** I HAVE A FEELING THAT'S NOT--

**GREG CIZEK:** I'M NOT SURE I UNDERSTAND THE QUESTION. I THINK THE SHORT TESTING WINDOWS ARE GREAT. I LOVED HEARING JIM TALK ABOUT ONE DAY OR SOMETHING. I'LL BE INTERESTED, WHEN TONY TALKS ABOUT SMARTER BALANCED, TO HEAR HOW SECURITY WILL BE DEALT WITH IN A LONGER COMPUTER-ADAPTIVE TESTING WINDOW. BUT YOU KNOW, SOME QUALIFICATION, WHETHER IT'S THIRD PARTY OR DIRECTED BY THE STATE, CAN BE DONE AT ANY TIME PRIOR TO THE ADMINISTRATION OF AN ASSESSMENT. SO I DON'T THINK THAT THE SIZE OF THE WINDOW IS RELATED

TO THE NEED OR THE TIMING OF THE QUALIFICATION THAT I'M TALKING ABOUT.

**SCOTT NORTON:** THAT WAS MY UNDERSTANDING, AS WELL--THAT THE CERTIFICATION WOULD BE SOMETHING THAT HAPPENED WELL BEFORE THE TESTING WINDOW.

**JACK BUCKLEY:** YES. IS THERE A MIKE?

**AUDIENCE MEMBER:** THANK YOU. I'M FROM THE DISTRICT OF COLUMBIA. I SERVE ON THE STATE BOARD OF EDUCATION. I HAVE 2 QUESTIONS. ONE IS VERY SPECIFIC. WHAT IS THE STATE OR DISTRICT OBLIGATION, IF ALLEGATIONS OF CHEATING PROVE TO BE TRUE, TO INFORM THE PARENTS WHOSE CHILD MAY HAVE SCORED PROFICIENT OR ADVANCED? AND THEN MY SECOND QUESTION IS MORE PHILOSOPHICAL. THE NATIONAL RESEARCH COUNCIL, THE NATIONAL ACADEMY OF SCIENCES' ARM, STUDIED 10 YEARS' WORTH OF BASICALLY HIGH-STAKES TESTS AND DETERMINED THAT OUR ABILITY TO MONITOR FOR GAMING OF THE SYSTEM WAS TOO LOW, BUT MORE IMPORTANTLY, THE MINUTE WE USE THE TEST DESIGNED TO MEASURE STUDENT LEARNING FOR ANY OTHER PURPOSE, WE INVALIDATE IT AS A MEASURE OF STUDENT LEARNING. SO I'D LIKE TO KNOW HOW THE PANEL CAN HELP US AS POLICYMAKERS GET OUT OF THAT CONUNDRUM, BECAUSE IF WE'RE INVALIDATING THE MEASURE, IF THE YARDSTICK IS WARPED, WE HAVE TO USE SOMETHING ELSE TO MEASURE STUDENT LEARNING AND EVALUATE

TEACHERS OR REWARD THEM OR WHATEVER. THOSE ARE MY 2 QUESTIONS.

THANKS.

**SCOTT NORTON:** I'LL TAKE THE FIRST ONE, IF THAT'S OK, AND TALK ABOUT THAT JUST A LITTLE BIT. IT'S A TRICKY ISSUE. WE HAVE INVESTIGATIONS. WE TALKED ABOUT THAT. THERE ARE PROCEDURES, AND SOMETIMES SCORES ARE INVALIDATED OR GIVEN A ZERO, SO THEN YOU REALLY DON'T HAVE A TRUE MEASURE OF THE CHILD'S ABILITY AT ALL. IT GETS VERY COMPLICATED VERY FAST IN OUR STATE. SOME TESTS ARE HIGH-STAKES, FOR GRADUATION OR PROMOTION. IN THOSE CASES, WE HAVE A RETEST OPPORTUNITY THAT'S PROVIDED--AGAIN, A LOT OF DETAILS AROUND THAT--SO THAT A VALID SCORE CAN BE REPORTED BACK. IF THEY'RE NOT HIGH-STAKES, WE DON'T DO THAT, BECAUSE OF COST, BUT IT IS AN ISSUE TO LET PARENTS KNOW WHAT IS THE TRUE MEASURE THAT WAS PROVIDED BY THE STATE, OTHER THAN JUST A ZERO BECAUSE THERE WAS A TEST SECURITY VIOLATION. WE TRY TO WORK AROUND THAT, BUT IT HAS BEEN A PROBLEM FOR US.

**JIM LIEBMAN:** NEW YORK CITY JUST PERFORMED A REALLY EXTENSIVE AUDIT OF, LIKE, 15% OF ITS HIGH SCHOOLS WHERE THE POSSIBILITY OF PROBLEMS WAS THE GREATEST. IT WASN'T JUST ON THE REGENTS EXAM, BUT IT WAS ALSO ON GRADUATION AND ALL OF THAT, AND IT WAS DETERMINED THAT THERE WERE A NUMBER OF STUDENTS WHO WERE GRADUATED WHO DID NOT ACTUALLY MEET THE STANDARD FOR GRADUATION, AND STEPS TAKEN THERE

WERE TO INFORM AND TO GIVE THEM THE RIGHT TO COME BACK AND GET WHATEVER IT WAS THAT THEY MISSED, BUT NOT TO PUNISH BY TAKING AWAY THEIR HIGH-SCHOOL DIPLOMA. IT WASN'T THEIR FAULT. SO THAT'S ONE WAY TO TRY TO BALANCE THAT PROBLEM.

**GREG CIZEK:** I'LL GO AHEAD AND WADE INTO THE MORASS OF THAT SECOND PART. AND I...FIRST OF ALL, I WOULD SAY, OF COURSE; WHEN THERE'S RELIANCE ON AN INDICATOR, PEOPLE WILL FIND WAYS TO GAME THE INDICATOR, AND SO IT BECOMES SOMEWHAT LESS PURE OF AN INDICATOR. I GUESS I WOULD QUIBBLE WITH THE TERM INVALIDATED. IT'S NOT VALID OR NONVALID. THERE'S A CONTINUUM OF VALIDITY THERE TO THE INDICATOR. BUT I ALSO COP TO SORT OF A BAD ATTITUDE HERE. IF YOU THINK BACK TO WHEN HIGH-STAKES TESTS FIRST WERE INSTITUTED, THEY WERE INSTITUTED LARGELY TO ADDRESS A PERSONNEL EVALUATION SYSTEM THAT WAS BROKEN. THERE WAS A CONCERN THAT WHAT WAS HAPPENING IN CLASSROOMS WAS NOT WHAT SHOULD BE HAPPENING IN CLASSROOMS, AND MY FIRST CHOICE, IF I WERE TO BE THINKING ABOUT PERSONNEL EVALUATION SYSTEMS, WOULD NOT BE TO GO GIVE KIDS ACHIEVEMENT TESTS. SO WE'VE MORPHED THAT INDICATOR BECAUSE IT HAPPENED TO BE THE QUICK AND EASY AND POLITICALLY RESPONSIVE THING TO DO, BUT LARGELY, TESTING HAS SUFFERED. ACHIEVEMENT TESTING HAS SUFFERED FOR A VERY LONG TIME BECAUSE IT'S BEEN USED AS A QUICK FIX TO ADDRESS BROKEN PERSONNEL EVALUATION SYSTEMS.

**JIM LIEBMAN:** I WANT TO SAY SOMETHING, TOO. I DISAGREE. I THINK THAT YOUR TESTING, YOUR ASSESSMENT, SHOULD BE A PACKAGE, AND YOU SHOULD BE DOING SUMMATIVE ASSESSMENT AND A LOT OF FORMATIVE ASSESSMENT AND A VARIETY OF THINGS IN BETWEEN, AND EACH OF THOSE ASSESSMENTS PERFORMS A DIFFERENT FUNCTION. I THINK ALL OF THEM HELP STUDENTS. I THINK WHAT THE SUMMATIVE TEST TELLS YOU IS, HERE'S AN AREA WHERE THIS STUDENT MAY NOT BE DOING SO WELL; GO LOOK AND SEE WHAT YOU CAN FIND OUT, AND THAT REQUIRES INQUIRY. WE HAVE THESE INQUIRY TEAMS IN ALL OF OUR SCHOOLS, AND THAT'S WHAT THEY INQUIRE INTO, AND YOU USE THEN THE FORMATIVE ASSESSMENTS, AND THEN YOU HAVE TO DO YOUR OWN SELF-ASSESSMENT--YOU KNOW, CLASSROOM ASSESSMENT AND ALL SORTS OF THINGS AND LOW-INFERENCE OBSERVATION AND EVERYTHING ELSE. SO WITHIN A SCHEME LIKE THAT, I THINK THAT THE IDEA THAT PUTTING STAKES ON AN ASSESSMENT TAKES AWAY ALL OF ITS VALUE, INCLUDING FOR PURPOSES OF PROVIDING INFORMATION ABOUT WHERE THINGS ARE GOING WRONG, IS JUST NOT THE CASE. WHEN WE PUT A GRADE AND STARTING GIVING SCHOOLS A GRADE, I WOULD GO AROUND AND TALK TO THE "F" SCHOOLS, AND THEY WOULD TELL ME, "WE WERE JUST SCREWING UP IN MATH. WE HAD A BAD MATH CURRICULUM." AND THAT'S SOMETHING THAT THEY PICKED UP FROM SOME STAKES PUT ON STATE TESTING, BASICALLY. WHAT THEY DID ABOUT THAT AND HOW THEY RESPONDED TO IT REQUIRED DEEPER TOOLS AND MORE SOPHISTICATED TOOLS, BUT TO KNOW THAT AND TO KNOW THAT THE PUBLIC KNOWS IT--REALLY IMPORTANT.

**AUDREY AMREIN-BEARDSLEY:** I THINK SOMETHING TO REMEMBER, AS WELL, IS IN 2000, THE AERA, WHICH IS THE AMERICAN EDUCATIONAL RESEARCH ASSOCIATION, THE NATIONAL COUNCIL FOR MEASUREMENT IN EDUCATION, AND THE AMERICAN PSYCHOLOGICAL ASSOCIATION CAME OUT WITH STANDARDS FOR TESTING, AND THE FIRST STANDARD IS THAT NO TEST SHOULD EVER BE USED IN ISOLATION TO MAKE ANY TYPES OF DECISIONS UNLESS A--OR HAVE A HIGH-STAKES CONSEQUENCE ATTACHED. I THINK THAT IF WE KEEP THAT IN THE FOREFRONT CONTINUOUSLY AS WE DEVELOP THESE PROGRAMS THAT YOU HAVE TO HAVE, AND THAT GOES BACK TO WHAT--WE LOOKED AT THE KORETZ ISSUE, WITH THE VALIDITY. WE HAVE TO MAKE THE CASE FOR CRITERION LEVELS OF VALIDITY. RECENTLY, IN THE LAST WEEK, THERE'S BEEN 2 ARTICLES PUBLISHED, ONE OUT OF TENNESSEE, ABOUT THESE VALUE-ADDED MEASUREMENT SYSTEMS, WHICH WE'RE BASICALLY NOW LOOKING AT NOT SNAPSHOTS, BUT LOOKING AT THE GROWTH OVER TIME, AND HOW THESE ARE TRUMPING THE SUPERVISOR EVALUATION SCORES. SO ADMINISTRATORS IN TENNESSEE, SPECIFICALLY, ARE GETTING IN TROUBLE IF THEIR OBSERVATION SCORES DO NOT MATCH THE VALUE-ADDED SCORES OF THESE TESTS, SO THEY'RE TAKING THAT DUAL APPROACH, THE MORE HOLISTIC APPROACH, AND SAYING, THESE TESTS, IT REALLY SHOULD BE A ONE-INDICATOR TYPE OF SYSTEM, AND I THINK THAT'S WHERE WE'RE GOING TO GET IN TROUBLE--IF WE START LOOKING MORE NARROWLY AT THESE SINGLE INDICATORS.

**JACK BUCKLEY:** OTHER QUESTIONS? EUNICE.

EUNICE GREER: THE ISSUE OF FORMATIVE ASSESSMENT HAS BEEN PUT ON THE TABLE, AND THE WHOLE NOTION OF ASSESSMENT BEING AN ONGOING PRACTICE, NOT JUST A ONE-TIME SHOT, AND I WOULD LOVE TO HEAR GREG AND AUDREY TALK MORE ABOUT THE TEACHER PREP ROLE IN THAT, BECAUSE THE RESEARCH THAT I'VE READ SUGGESTS THAT TEACHERS REALLY STRUGGLE WITH THEIR ROLE IN FORMATIVE ASSESSMENT. AND IT'S A FABULOUS PROCESS, BUT WHAT DO YOU SEE NEEDING TO CHANGE IF IT'S GOING TO TAKE THE OPPORTUNITY IT OFFERS US IN THE WHOLE SYSTEM OF ASSESSMENTS FOR TEACHER ED?

**AUDREY AMREIN-BEARDSLEY:** PROBABLY THE BIGGEST PROBLEM I SEE IS THAT THE TEST SCORES ARE ADMINISTERED AFTER THE STUDENTS LEAVE. I THINK THAT'S WHAT THE TEACHERS TALK ABOUT MOST, IN TERMS OF HOW ARE THEY SUPPOSED TO USE THE DATA IF THE STUDENTS ARE ALREADY GONE? THEY CAN USE IT FORMATIVELY TO TALK ABOUT GENERAL THEMES, BUT NOT ACTUALLY FOCUSED ON STUDENTS UNLESS THEY FORWARD THE SCORES INTO THE FUTURE. JAMES POPHAM, WHO'S AN ASSESSMENT PROFESSOR FROM UCLA, EMERITUS, HE ALSO TALKS ABOUT THE FURTHER THE TEST GETS AWAY FROM THE CLASSROOM, FOR FORMATIVE PURPOSES, THE WORSE IT GETS. SO HE REALLY ADVOCATES FOR THE CLASSROOM LEVEL, THE DISTRICT BENCHMARK,

HOPEFULLY VALIDATED DISTRICT BENCHMARK ASSESSMENT--USING THOSE FOR FORMATIVE PURPOSES, WHERE REALLY THE BIG, LARGE-SCALE, HIGH-STAKES TESTS ARE MORE APPROPRIATE FOR THE CONSEQUENTIAL SUMMATIVE PURPOSES.

**GREG CIZEK:** THE MORE I HEAR PEOPLE TALK ABOUT TESTING, SUMMATIVE OR FORMATIVE OR WHATEVER, THE MORE CONVINCED I AM THAT WE HAVE TO INTERVENE MUCH SOONER WITH TEACHERS AND PRINCIPALS TO HELP THEM UNDERSTAND THAT ASSESSMENTS HAVE DIFFERENT PURPOSES, AND THEY'RE DESIGNED TO SERVE DIFFERENT PURPOSES, AND THEY FULFILL THOSE PURPOSES DIFFERENTIALLY WELL. THE IDEA THAT EVERY TEST CAN HAVE A FORMATIVE PURPOSE IS PROBABLY NOT AN ACCURATE CONCLUSION. THE LARGE-SCALE END-OF-YEAR TEST THAT EVERY KID TAKES IN LOUISIANA OR INDIANA OR ANYWHERE IS VERY GOOD AT SYSTEM MONITORING, AGGREGATE KIND OF PERFORMANCE INFORMATION. IT'S OUTSTANDING AT THAT. IT'S NOT VERY GOOD AT PROVIDING INDIVIDUAL STUDENT REMEDIATION PLANS. FORMATIVE ASSESSMENT IS VERY USEFUL FOR HELPING TEACHERS KNOW WHAT TO DO WITH A KID THE NEXT TIME THAT TEACHER INTERVENES BUT TERRIBLE AT SYSTEM MONITORING KINDS OF FUNCTIONS. IF WE CONTINUE TO INSIST THAT EVERY TEST SERVE A MULTITUDE OF PURPOSES, NONE OF THEM ARE GOING TO DO THAT EQUALLY WELL, AND SO THE IDEA OF TAILORING THE TESTING ENTITY TO THE PURPOSE WE HAVE IN MIND IS, I THINK, ESSENTIAL TO UNDERSTANDING THE ROLE THAT LARGE-SCALE AND CLASSROOM TESTS PLAY.

**JACK BUCKLEY:** THERE'S A QUESTION IN THE BACK OF THE ROOM.

[WOMAN SPEAKING INDISTINCTLY]

**AUDIENCE MEMBER:** SORRY. HI. I'M WITH THE AMERICAN INSTITUTES FOR RESEARCH, AND I USED TO WORK FOR A SCHOOL DISTRICT WHERE ONE YEAR WE ASKED TEACHERS TO SIGN A PIECE OF PAPER THAT SAID, "I HAVE READ THE TESTING PROCEDURES, AND I AGREE TO FOLLOW THE RULES," AND THE UNION SAID, "NO. DON'T SIGN THAT." SO I'M WONDERING, AS WE'RE DEVELOPING BEST PRACTICES, HOW DO WE INVOLVE THE UNIONS, AND HOW DO WE LEVERAGE THE UNIONS?

**JIM LIEBMAN:** I THINK, YOU KNOW, IT'S A VERY IMPORTANT ISSUE, AND I THINK THAT I WOULD SAY, "HOW DO WE INVOLVE THE TEACHERS," AND THE UNIONS ARE A PART OF THAT. BUT I THINK IT'S VERY IMPORTANT, IN DESIGNING EVERYTHING THAT WE'VE TALKED ABOUT, THAT THEY SHOULD BE DESIGNED IN PART BY TEACHERS AND WITH THE INVOLVEMENT OF TEACHERS, BECAUSE THEY'RE GOING TO BE IN THE CLASSROOM AND THEY'RE GOING TO TELL YOU AN AWFUL LOT ABOUT HOW IT'S GOING TO WORK OR WHY IT WON'T WORK AND THINGS LIKE THAT. I THINK THAT ONCE YOU'VE DESIGNED THESE THINGS AND ROLLED THEM OUT FOR A YEAR, THEN YOU NEED TO START HEARING BACK FROM TEACHERS ABOUT WHAT THEY EXPERIENCED IN THE CLASSROOM WITH ALL OF THE PROCEDURES AND THEN KIND OF WORK BACK

FROM THAT. I THINK IF YOU CAN MAKE THAT KIND OF AN EFFORT WITH THE TEACHERS, IT WILL HELP YOU AN AWFUL LOT WITH THE UNION, BECAUSE THEN THEY'LL UNDERSTAND THIS TO BE SOMETHING THAT'S WORKING WITH TEACHERS, NOT AIMED AT TEACHERS. BUT, YOU KNOW, THE UNIONS HAVE A POSITION ON TESTING THAT MAKES THEM, IN SOME CASES, SOMEWHAT DUBIOUS ABOUT WHAT THE TESTING IS ALL ABOUT, AND THEY CAN RAISE SOME QUESTIONS THAT MAKE IT DIFFICULT, AND I THINK THAT ONE OF THE THINGS THAT--WE TALKED ABOUT HOW STATES CAN BE A LITTLE MORE PASSIVE THAN THEY NEED TO BE SOMETIMES. I THINK DISTRICTS SOMETIMES ARE A LITTLE MORE PASSIVE THAN THEY NEED TO BE BECAUSE THEY'RE WORRIED ABOUT WHAT THE UNION WOULD DO. I THINK YOU JUST WANT TO GET OUT IN FRONT OF THAT AND WORK WITH THE UNION AS MUCH AS POSSIBLE, BUT NOT USE IT AS AN EXCUSE NOT TO DO WHAT'S NECESSARY.

**GREG CIZEK:** ONE OF THE MAIN POINTS THAT I TRY TO MAKE IS THIS NOTION OF SHARED RESPONSIBILITIES, AND ONE OF THE GROUPS I TALKED ABOUT WAS PROFESSIONAL ASSOCIATIONS. I HIGHLIGHTED MY OWN PROFESSIONAL ASSOCIATION, BUT IT IS SOMEWHAT OF AN EMBARRASSMENT THAT EDUCATORS REALLY HAVE NOT BEEN THE MOTIVATION FOR WHY WE'RE EVEN HERE TODAY. IT'S LARGELY BEEN THE MEDIA. YOU HEARD EARLIER IN SOME OF THE INTRODUCTORY REMARKS THAT AN ARTICLE IN THE "USA TODAY" PROMPTED THE CHANCELLOR TO LOOK AT WHAT ARE WE DOING, HOW ARE WE DOING IT. IT WAS AN ARTICLE IN THE "DALLAS MORNING NEWS" THAT PROMPTED DALLAS AND THE STATE OF TEXAS TO START LOOKING INTO STUFF.

IT WAS AN ARTICLE IN THE "ATLANTA JOURNAL-CONSTITUTION" AND RESEARCH DONE BY MEDIA THERE THAT PROMPTED--YOU KNOW, EDUCATORS REALLY SHOULDN'T BE--BY EDUCATORS, I MEAN ALL OF US INVOLVED IN THE EDUCATIONAL ENTERPRISE, PROFESSIONAL ASSOCIATIONS INCLUDED-- SHOULDN'T BE REACTING TO OUTSIDE INFORMATION. WE SHOULD BE THE PRODUCERS AND ACTORS UPON THAT INFORMATION, THAT OUR ACTIONS GET REPORTED ON, NOT CONSTANTLY IN RESPONSE TO OUTSIDE PRESSURES.

**SCOTT NORTON:** IN MY COMMENTS, I TALKED A LITTLE BIT ABOUT THE OATH THAT WE HAVE IN LOUISIANA THAT EVERYBODY SIGNS OFF ON, AND I SAID I'M NOT SURE WHETHER THOSE THINGS REALLY DO A LOT OF GOOD OR NOT. WE'VE OPTED TO CONTINUE THAT. I DON'T THINK IT COULD HURT, BUT I WILL SAY I'M NOT SURE WHAT THE POINT WOULD BE OF ENCOURAGING TEACHERS TO NOT SIGN A STATEMENT THAT SAID THEY DIDN'T VIOLATE THE RULES. I DON'T REALLY GET THAT. IT'S A LITTLE BEWILDERING TO ME.

**JACK BUCKLEY:** YES.

**BOB WILSON:** I'M SORT OF THE OUTSIDE GUY ON THIS PANEL. I'M NOT A PROFESSIONAL EDUCATOR OR CONNECTED FULLY TO THE EDUCATIONAL SYSTEM. I'M A SCHOOL BOARD LAWYER, BUT NOT MANY PEOPLE AROUND HAVE BEEN AS DEEP INTO THE BOWELS OF A SCHOOL SYSTEM AS I HAVE. AND THE PREVENTION, WHICH DR. CIZEK MENTIONED--AND HE WAS OUR EXPERT IN THAT CASE AND WAS TREMENDOUSLY HELPFUL TO US; I CAN'T THANK HIM

ENOUGH--STARTS WITH YOUR SUPERINTENDENT OR CHANCELLOR. I UNDERSTAND THE QUESTION ABOUT THE UNION. IT'S A VERY GOOD QUESTION. BUT YOU'RE TALKING ABOUT AN HONOR CODE, AND IF YOU GO TO WEST POINT, YOU'RE NOT EXPECTED TO DO ANYTHING EXCEPT BE HONORABLE AND TAKE YOUR TEST FAIRLY AND REPORT IF YOU SEE SOMETHING WRONG, BUT THEY STILL HAVE TO AGREE TO THAT HONOR CODE. AND WHOEVER THAT SUPERINTENDENT OR CHANCELLOR IS FOR THAT SYSTEM, IN THE PREVENTIVE WORLD AND SHARED RESPONSIBILITY THAT DR. CIZEK MENTIONED, IS THE PERSON THAT, I THINK--AND I WANT YOUR REACTION TO IT--THAT ABSOLUTELY, POSITIVELY, WITHOUT DOUBT SETS THE TONE FOR THAT SYSTEM, AND IT HAS TO BE THAT THEY TALK ABOUT TESTING WITH HONOR AND THE RESPONSIBILITY TO EVERYBODY, STUDENTS AND PROFESSIONALS IN THE SYSTEM, WHETHER IT'S THE CAFETERIA WORKER WHO SEES SOMETHING OR A JANITOR, JUST AS MUCH AS A TEACHER OR A STUDENT. AND THEY SET THAT TONE, AND THEY EXPECT NOTHING LESS, AND THERE WILL BE DIRE CONSEQUENCES IF YOU VIOLATE IT, BUT EVERYBODY HAS A STAKE IN IT, AND WHAT I FOUND A LOT OF PLACES--WE TALK ABOUT THE TEST ALL THE TIME. NEVER DO THEY TALK ABOUT TAKING THE TEST WITH HONOR. NOW, YOU CAN SIT THERE AND SAY, "OH, I CAN'T BELIEVE THAT." NOBODY IN ATLANTA IN THE INSTRUCTIONAL PROCESSES, OR DOUGHERTY COUNTY IN SOUTH GEORGIA, WHICH WE ALSO INVESTIGATED, EVER MIXED THE TRAINING WHERE THE EMPHASIS WAS EQUALLY AS MUCH ON THE INTEGRITY OF TAKING THE SYSTEM--JUST THE RULES, POLICIES, AND PROCEDURES, WITHOUT REALLY UNDERSTANDING WHY. I WANT TO KNOW WHAT YOU ALL THINK OF

THAT, BECAUSE I BELIEVE IT DOES START AT THE TOP, BUT I WANT TO SEE IF YOU ALL AGREE.

**JIM LIEBMAN:** COULDN'T AGREE MORE. I MENTIONED THAT CHANCELLOR WALCOTT IN NEW YORK CITY, WHEN THERE WAS DISCUSSION OF CHEATING AND ALL OF THAT, JUST WENT ON TV, BASICALLY, AND SAID, "CALL ME. WE TAKE IT VERY SERIOUSLY." AND I THINK THAT'S REALLY IMPORTANT. NEW YORK CITY--I MENTIONED THIS AUDIT THAT IT RECENTLY DID. YOU MIGHT THINK IT'S SUICIDE. THEY ESSENTIALLY SAID, WE'RE GOING TO GO FIGURE OUT WHAT CAN GO WRONG IN HIGH SCHOOLS, AND WE'RE GOING TO START BY LOOKING AT THE 15% WHERE THE DATA THAT WE'VE JUST LOOKED AT TELLS US IT'S PROBABLY THE WORST, AND THEN THEY SAMPLED TEACHERS IN CLASSROOMS WHERE THEY THOUGHT IT WAS PROBABLY THE WORST OF THE WORST, AND THEN THEY MADE A PUBLIC REPORT ABOUT IT, AND I THINK THAT'S THE KIND OF THING THAT YOU NEED TO BE WILLING TO DO, EVEN THOUGH YOU KNOW YOU'RE GOING TO PUT YOURSELF OUT THERE, BECAUSE IT FORCES YOU TO CONFRONT WHAT THE PROBLEMS ARE. BUT IF YOU REALLY, REALLY CARE ABOUT YOUR DATA, AND I THINK MORE AND MORE PEOPLE REALLY DO NOW, FOR ALL OF THE REASONS THAT WE'VE TALKED ABOUT, YOU CAN'T HAVE BAD DATA, AND ONE OF THE WAYS YOU GET BAD DATA IS FROM CHEATING ON TESTS. SO I THINK IT CAN ALL WORK TOGETHER TO CREATE A CULTURE WHERE THAT'S JUST ACCEPTED.

**AUDREY AMREIN-BEARDSLEY:** I THINK THAT GOES TO THE SHARED RESPONSIBILITY, TOO, AND ALSO, WHAT BETTER LESSON FOR STUDENTS IN THE CLASSROOMS, THAT UNDERSTAND THAT THERE IS A CODE OF HONOR? THERE'S AN ETHICS OF HONOR IN TERMS OF WHAT WE'RE DOING, WHY WE'RE DOING THIS. WE HAD A LOT OF CASES IN THIS STUDY WHERE PARENTS ACTUALLY PARTICIPATED AND TALKED ABOUT THE STORIES THAT THEY HAD AFTER THEIR CHILDREN CAME HOME AND SAID, "MOMMY, THIS DIDN'T SEEM VERY RIGHT. THIS HAPPENED IN MY CLASS TODAY DURING THIS TEST." SO WE HAVE THEM. THEY'RE PART OF THIS, AS WELL, THE PARENTS AND THE STUDENTS. I THINK THAT--FRAMING THIS ISSUE WITH HONOR AND HOW WE'RE GOING TO APPROACH THIS APPROPRIATELY--I THINK THAT'S THE BEST LESSON FOR EVERYBODY.

**SCOTT NORTON:** I THINK YOU MAKE A GREAT POINT, AND I THINK OUR SYSTEMS, WHERE A STATE HAS ONE, DO A PRETTY GOOD JOB WHEN PEOPLE FOLLOW THAT CODE. IF SOMETHING HAPPENS, THERE'S A SYSTEM OF CHECKS AND BALANCES AND INVESTIGATION. WHERE WE'RE WEAKER, I THINK--IT WAS REALLY ON MY LAST SLIDE--IS WHEN THERE'S A BREAKDOWN IN THERE, WHETHER IT BE A SUPERINTENDENT, TEACHER, OR WHOEVER. THEY REALLY DON'T WANT TO REPORT UP, AND THEN WE'RE SORT OF LEFT HOLDING THE BAG AND TRYING TO FIGURE OUT WHAT REALLY HAPPENED. THAT'S WHEN IT BECOMES VERY DIFFICULT.

**JACK BUCKLEY:** ALL RIGHT.

**AUDIENCE MEMBER:** EXCUSE ME. I WORK FOR THE AFT, AND I'D LIKE TO SPEAK FOR THE UNION.

**JACK BUCKLEY:** I'LL GIVE YOU ONE MINUTE, BUT THEN I NEED TO WRAP UP AND MOVE ON TO THE NEXT PANEL.

**AUDIENCE MEMBER:** 30 SECONDS. LAST FALL, WE ADOPTED A VERY STRONG RESOLUTION ON CHEATING THAT'S AS STRONG AS ANYTHING, I SUSPECT, YOU WILL COME UP WITH TODAY. WE ENCOURAGE TEACHERS TO REPORT CHEATING THROUGH THEIR UNION SO THAT THEY CAN HAVE SOME SORT OF PROTECTION. IF YOU LOOK AT ATLANTA, IF YOU LOOK AT MANY OTHER PLACES AROUND THE COUNTRY, TEACHERS WHO REPORT CHEATING TEND TO BE THE ONES THAT GET FIRED OR GET IN TROUBLE, AND TO JUST SAY THAT UNIONS ARE, WHAT, THE BAD GUYS IN THIS, IS JUST ABSOLUTELY NOT TRUE.

**JACK BUCKLEY:** I THINK I'VE HEARD A LOT OF THINGS THIS MORNING. I HAVEN'T HEARD ANYONE TRYING TO BLAME THE UNION FOR PROBLEMS WITH TESTING INTEGRITY, AND I'D ENCOURAGE YOU TO STICK AROUND FOR THE NEXT TWO PANELS, WHERE WE'LL HAVE A LOT MORE TIME TO LOOK AT THAT ISSUE. I THINK IN TERMS OF WHAT WE'VE HEARD THIS MORNING, VERY BRIEFLY, CERTAINLY I THINK THAT THERE'S SOMEWHAT CONSENSUS THAT THERE ARE DEGREES OF CHEATING OR TESTING IRREGULARITY AND THAT THOSE TWO CONCEPTS ARE NOT NECESSARILY THE SAME THING. I THINK

CERTAINLY I'VE HEARD FROM THE PANEL AND FROM THE AUDIENCE THAT MULTIPLE MECHANISMS ARE NEEDED FOR PREVENTION, BOTH AT THE OPERATIONAL LEVEL AND THEN SORT OF FURTHER REMOVED FROM THE ADMINISTRATION ROOM. IT SOUNDS LIKE THERE'S A CALL FOR MORE CONSISTENCY IN STANDARDIZATION IN DETECTION, AND THAT'S SOMETHING WE'LL GET TO IN THE NEXT PANEL.

I THINK A LOT OF COMMENTS THIS MORNING FOCUSED ON TRYING TO MAINTAIN A HEALTHY TESTING ENVIRONMENT, AND I THINK THE POINT IS WELL TAKEN THAT THAT STARTS AT THE TOP AND NOT NECESSARILY AT THE BOTTOM, BUT IT IS A SHARED RESPONSIBILITY AMONG STUDENTS AND EDUCATORS AND TEST DEVELOPERS AND PARENTS AND PROFESSIONAL ASSOCIATIONS. I THINK THERE WAS A LOT OF TALK THAT I DIDN'T NECESSARILY EXPECT ABOUT TRAINING AND CERTIFICATION IN THE PROCTORING OR ADMINISTRATION PROCESS, AND I THINK THAT'S SOMETHING THAT WE'LL HAVE TO INVESTIGATE FURTHER. IT WAS VERY INTERESTING TO HEAR. AND I THINK JUST SORT OF GENERAL CALL FOR MORE COORDINATION AMONG STATE AND LEAs, AND THAT'S TO BE EXPECTED, AS WELL. THERE'S A LOT OF PLAYERS HERE. IT'S A VERY COMPLICATED ENVIRONMENT. SO AT THIS POINT, WE'RE GOING TO TAKE A 5-MINUTE BREAK JUST SO WE CAN, AGAIN, ROTATE THE PANELISTS UP FRONT. SO, FOLKS ON PANEL TWO, PLEASE COME UP.

THANK YOU VERY MUCH. JOIN ME IN THANKING THE PANELISTS FOR A GREAT  
JOB.

[APPLAUSE]

PANEL II: DETECTION AND ANALYSIS OF IRREGULARITIES IN ACADEMIC TESTING

**JACK BUCKLEY:** ALL RIGHT, SO, WE'RE BACK, AND OUR SECOND PANEL FOCUSES ON DETECTION AND ANALYSIS OF IRREGULARITIES IN TESTING.

WE'LL START WITH BRIAN JACOB, WHO IS GONNA JOIN US REMOTELY FROM ANN ARBOR, WHERE HE'S THE WALTER H. ANNENBERG PROFESSOR OF EDUCATION POLICY, PROFESSOR OF ECONOMICS, AND DIRECTOR OF THE CENTER ON LOCAL, STATE, AND URBAN POLICY AT THE UNIVERSITY OF MICHIGAN'S GERALD FORD CENTER OF PUBLIC POLICY. HIS PRIMARY FIELDS OF INTEREST ARE LABOR ECONOMICS, PROGRAM EVALUATION, AND THE ECONOMICS OF EDUCATION. HIS CURRENT RESEARCH FOCUSES ON URBAN SCHOOL REFORM AND TEACHER LABOR MARKETS.

NEXT, WE'LL HEAR FROM CARSWELL WHITEHEAD, A DIRECTOR IN ETS' OFFICE OF TESTING INTEGRITY, WHERE SHE PLANS, DIRECTS, AND MANAGES SECURITY INVESTIGATIONS AND POLICIES DESIGNED TO PREVENT AND DETECT SCORE IRREGULARITIES FOR COLLEGE BOARD TESTING PROGRAMS.

DR. DAVID FOSTER, CURRENTLY THE CEO OF CAVEON TEST SECURITY, WILL BE OUR NEXT SPEAKER. HE'S ALSO THE CHIEF SCIENTIST AND EXECUTIVE VICE PRESIDENT OF KRYTERION, A UNIQUE INTERNET TEST ADMINISTRATION COMPANY WHICH HE FOUNDED. HE'S BEEN CREATING AND IMPROVING

COMPUTERIZED TESTING SYSTEMS SINCE 1982 AND HAS BEEN INFLUENTIAL OVER THE PAST 28 YEARS IN INTRODUCING MANY IMPORTANT INDUSTRY INNOVATIONS AND INITIATIVES.

AND THEN FINALLY, AS ASSOCIATE SUPERINTENDENT FOR ASSESSMENT AND ACCOUNTABILITY AT THE GEORGIA DEPARTMENT OF EDUCATION, MELISSA FINCHER OVERSEES THE DEVELOPMENT OF ALL OF GEORGIA'S ASSESSMENT AND ACCOUNTABILITY PROGRAMS. SHE'S RESPONSIBLE FOR ENSURING THESE PROGRAMS MEET HIGH STANDARDS FOR TECHNICAL DEFENSIBILITY. NOW I'LL TURN IT OVER TO BRIAN.

**BRIAN JACOB:** OK, HELLO.

I'M NOT SURE IF EVERYONE CAN SEE OR HEAR ME NOW, BUT I WILL PROCEED AS IF YOU CAN.

I THINK I SHOULD BE ON THE TITLE SLIDE HERE. THIS IS "DETECTION AND ANALYSIS OF IRREGULARITIES IN ACADEMIC TESTING." AND SO SLIDE 2 IS JUST A FOCUS OF THE TALK.

SO, AS AN ACADEMIC ON THE PANEL, I...AM FREE TO ADMIT THAT I HAVE LIMITED FAMILIARITY WITH THE PRACTICES AND POLICIES CURRENTLY USED BY SEAs AND LEAs--SOME, BUT NOT ALL. SO, THE LESSONS I'M GONNA SHARE TODAY COME FROM THE THEORY OF STATISTICAL ANALYSIS FOR THIS

TYPE OF DETECTION AND THEN FROM MY EXPERIENCE WORKING WITH SEVERAL  
LARGE URBAN SCHOOL DISTRICTS AROUND SOME OF THESE ISSUES.

I'M GOING TO FOCUS OUR REMARKS TODAY ON TEACHER AND/OR  
ADMINISTRATOR CHEATING AS OPPOSED TO STUDENT CHEATING OR THE  
BROADER RANGE OF OTHER ILLICIT TESTING ACTIVITIES AND  
IRREGULARITIES. I'M ALSO GOING TO FOCUS MY TALK ON STATISTICAL  
ANALYSES ONE CAN USE TO DETECT SUCH MANIPULATION AND RECOGNIZING  
THAT THIS HAS LIMITS  
AND THEN OTHER METHODS WILL BE NECESSARY TO COMPLEMENT THESE.  
MY TALK WILL FOCUS ON THE STATISTICAL ANALYSIS.

SO, NEXT SLIDE.

THE ROLE OF STATISTICAL ANALYSIS IN DETECTING IRREGULARITY.

STATISTICAL ANALYSES HAVE A VARIETY OF BENEFITS.

ONE--THEY'RE RELATIVELY LOW COST. CERTAINLY, RELATIVE TO IN-DEPTH  
FORENSIC INVESTIGATIONS, MONITORING OF HUNDREDS OF THOUSANDS OF  
CLASSROOMS. THEY CAN COVER THE ENTIRE POPULATION. UNLIKE SOME  
VERY LEGITIMATE METHODS, SUCH AS TIPLINES OR INDIVIDUAL CASE  
STUDIES OF CERTAIN SCHOOLS, THIS IS A WAY TO LOOK AT THE ENTIRE

POPULATION OF CLASSROOMS IN AN AREA. SOME OF THESE MEASURES, WHEN PROPERLY USED, CAN PROVIDE A BETTER UNDERSTANDING OF THE EXTENT OF THE IRREGULARITY—YOU KNOW, A MEASURABLE LIKELIHOOD THAT AN IRREGULARITY WOULD HAVE OCCURRED BY CHANCE IN THE ABSENCE OF ANY MANIPULATION. THEN IT CAN ALSO IDENTIFY SYSTEMATIC PATTERNS AS WELL AS INDIVIDUAL CASES OF CONCERN.

SO, ALTHOUGH STATISTICAL ANALYSES CERTAINLY HAVE LIMITATIONS, SO I THINK A USEFUL ANALOGY IS KIND OF THE CLINICAL SCREENING FOR RARE DISEASES THAT ONE WOULD DO IN MEDICINE. STATISTICAL ANALYSES TO DETECT CHEATING CAN HAVE A HIGH RATE OF FALSE POSITIVES AND/OR FALSE NEGATIVES. IT'S CRITICAL TO COMPLEMENT THEM WITH OTHER METHODS.

AND THEN IT REALLY IS IMPOSSIBLE, I THINK, TO DETERMINE THE ACTUAL INDIVIDUALS RESPONSIBLE FOR MANIPULATION, EVEN IF ONE IS CONVINCED THAT THERE WAS IRREGULARITY AND THIS WAS DUE TO SOME ILLICIT HUMAN INTERVENTION, WHETHER IT WAS THE ACTUAL TEACHER OF RECORD, THE COUNSELOR, THE ADMINISTRATOR. ONE CAN'T REALLY DETERMINE FROM A STATISTICAL ANALYSIS. SO, NEXT SLIDE.

IN GENERAL, THERE ARE 3 DIFFERENT APPROACHES FOR DETECTING TESTING IRREGULARITIES USING STATISTICAL ANALYSES. ONE IS IN A RATIO ANALYSIS, AND THIS IS SOMETHING WHERE ONE WOULD DETERMINE

GENERALLY THE AVERAGE NUMBER OF WRONG TO RIGHT ERASURES IN A CLASSROOM. AND CERTAINLY PEOPLE ON THE PANEL AND IN THE AUDIENCE ARE PROBABLY MUCH MORE FAMILIAR WITH THIS THAN I AM.

TESTING COMPANIES HAVE, I THINK, FAIRLY WELL-DEVELOPED METHODS FOR SCANNING AND DETERMINING WHETHER SOMETHING IS A TRUE ERASURE AS OPPOSED TO A DIFFERENT TYPE OF MARK ON A TEST FORM.

THEN A SECOND METHOD IS ITEM RESPONSE ANALYSIS, WHERE YOU'D BASICALLY BE LOOKING FOR UNUSUALLY COMMON RESPONSE PATTERNS ACROSS STUDENTS WITHIN THE SAME CLASS. AND IN THIS TYPE OF ANALYSIS, YOU WOULD ALWAYS BE TRYING TO ACCOUNT FOR OTHER CHARACTERISTICS THAT WOULD CAUSE THIS RESPONSE PATTERN TO BE COMMON THAT WOULDN'T BE DUE TO CHEATING.

AND FINALLY, LOOKING AT THE ANALYSIS OF THE TEST SCORES THEMSELVES, LOOKING FOR UNUSUALLY LARGE GAINS IN ONE YEAR AND/OR VERY LARGE RELATIVE DECLINES IN THE SUBSEQUENT YEAR OR UNEXPECTED JUMPS IN THE DISTRIBUTION OF TEST SCORES. NEXT SLIDE.

THIS IS A GRAPH BASED ON SOME WORK THAT COLLEAGUES OF MINE AND I DID, LOOKING AT THE NEW YORK STATE REGENTS EXAM A YEAR OR SO AGO. AND THIS IS AN EXAMPLE OF HOW ONE COULD USE STATISTICAL ANALYSIS TO

IDENTIFY SYSTEMIC IRREGULARITIES. THIS WAS NOT AN ANALYSIS TO IDENTIFY WHICH PARTICULAR SCHOOLS OR CLASSROOMS MIGHT HAVE BEEN... INVOLVED IN MANIPULATION. BUT THIS IS THE NEW YORK STATE REGENTS HIGH SCHOOL EXAM REQUIRED FOR GRADUATION.

YOU SEE A GRAPHIC HERE BY SCALED SCORE ON THE VERTICAL AXIS. IT'S THE NUMBER OF STUDENTS IN THE STATE THAT RECEIVE THIS ACTUAL INDIVIDUAL SCALED SCORE. AND THE TWO SOLID VERTICAL LINES AT 55 AND 65 INDICATE THE PASSING THRESHOLDS FOR A REGULAR AND A MORE KIND OF BASIC DIPLOMA IN NEW YORK STATE.

SO, WHAT IMMEDIATELY JUMPS OUT AT YOU IS THE NUMBER OF STUDENTS SCORING A 63 OR 64 IS MUCH, MUCH LOWER THAN THE NUMBER OF STUDENTS SCORING EXACTLY A 65, WHICH IS WHAT'S NECESSARY FOR PASSING AND GRADUATING HIGH SCHOOL IN NEW YORK STATE. AND YOU SEE A SIMILAR BLIP AT THE 54 VERSUS 55 MARGIN.

AND SO BASED ON SOME STATISTICAL ANALYSES THAT UNDERLIE THIS GRAPH, YOU CAN CONCLUDE THAT THIS IS VIRTUALLY IMPOSSIBLE TO HAVE OCCURRED BY CHANCE AND SUGGEST THERE MAY HAVE BEEN SOME MANIPULATION IN THE GRADING OF THESE EXAMS, WHICH IS DONE BY THE TEACHERS IN THE ACTUAL HIGH SCHOOLS FOR THEIR OWN STUDENTS. SO AGAIN, THIS IS AN EXAMPLE USING STATISTICS TO IDENTIFY SYSTEMIC IRREGULARITIES. NEXT SLIDE.

SO, THIS IS AN EXAMPLE BASED ON SOME WORK I DID WITH STEVE LEVITT OUT IN CHICAGO IN THE EARLY 2000s LOOKING AT TESTING IRREGULARITIES USING ACHIEVEMENT GAINS AND RESPONSE PATTERNS.

THIS IS AN EXAMPLE. EACH ROW IS A DIFFERENT STUDENT. EACH COLUMN IS THEIR RESPONSE CATEGORY FOR DIFFERENT ITEMS. NEXT SLIDE.

SO, IF YOU LOOK HERE, THE FIRST THING YOU NOTICE IN THIS CLASS WAS THAT THERE WAS A VERY LARGE INCREASE IN SCORES--AN AVERAGE 4.1 GRADE EQUIVALENCE JUMPING TO 5.8 GRADE EQUIVALENCE IN THAT SUSPECTED YEAR. AND THEN IF YOU FOLLOW THE STUDENTS TO THE YEAR AFTER WHEN THEY WERE NOT IN THIS SUSPECTED TEACHER'S ROOM, THEIR TEST SCORES ACTUALLY GO DOWN, EVEN THOUGH THEY HAD A YEAR EXTRA OF EDUCATION. SO, THIS PATTERN HERE CLEARLY RAISES SUSPICION. AND THE NEXT SLIDE?

HERE, THERE'S ALSO SOME STATISTICAL ANALYSES UNDERLYING THIS, WHERE YOU CAN LOOK AND IDENTIFY A NUMBER OF THE ITEMS IN THIS CLASSROOM. THERE WAS--THE CORRELATION ACROSS STUDENTS WAS EXTRAORDINARILY LARGE, EVEN CONDITIONAL ON WHAT YOU WOULD EXPECT GIVEN THE STUDENT CHARACTERISTIC. NEXT SLIDE.

SO, FINALLY, I THINK IT'S IMPORTANT, THEN, REGARDLESS OF THE STATISTICAL ANALYSIS, TO UNDERSTAND THE TESTING IRREGULARITY DOES

NOT NECESSARILY IMPLY CHEATING. A FEW EXAMPLES I HAVE LISTED HERE:  
FREQUENT ERASURES MAY BE MORE COMMON AMONG CERTAIN STUDENTS THAN  
OTHERS. COMMONALITY OF ITEMS AND RESPONSE PATTERNS ACROSS  
STUDENTS MIGHT BE DUE TO CURRICULAR INSTRUCTIONAL PURPOSE OF THE  
TEACHER. AND OBVIOUSLY LARGE TEST SCORE GAINS COULD BE THE RESULT  
OF AN EXTREMELY EFFECTIVE TEACHER, WHICH WOULD BE THE LAST THING WE  
WOULD WANT TO DISCOURAGE.

SO WE ALWAYS NEED TO COMPLEMENT THESE ANALYSES WITH OTHER METHODS  
OF PROTECTION. NEXT SLIDE.

AND SO, WHAT CAN WE DO WHILE WE ARE ANALYZING SOME OF THESE TESTING  
DATA? ONE, IS THAT IT'S CRITICAL TO ACCOUNT FOR STUDENT AND/OR  
SCHOOL CHARACTERISTICS THAT MIGHT BE ASSOCIATED WITH ERASURES OR  
UNUSUAL TEST SCORE PATTERNS. I THINK THIS IS SOMETHING THAT IS LESS  
COMMON

BASED ON MY EXPERIENCE IN THE TESTING INDUSTRY, WHERE THERE'S  
USUALLY A COMPARISON OF ONE CLASSROOM OR SCHOOL TO, LET'S  
SAY, THE STATE AVERAGE, EVEN IF THE STATE AVERAGE MIGHT BE A VERY  
DIFFERENT GROUP OF STUDENTS OR TYPE OF STUDENTS THAN THE STUDENTS  
IN THE SUSPECTED CLASSROOM.

SECOND, USING METRICS OR ROBUST TO OUTLIERS, SO MAYBE THE MEDIAN AS OPPOSED TO THE MEAN, WE WOULDN'T WANT TO BASE A JUDGMENT ON A CLASSROOM WHERE ONE OR TWO STUDENTS WERE DRIVING AN ENTIRE CLASSROOM MEASURE, UTILIZING MULTIPLE APPROACHES WHEN POSSIBLE. I THINK CONDUCTING SELECTIVE RETEST AUDITS MAY BE SOMETHING THAT IN CERTAIN WELL-CHOSEN CASES CAN BE VERY EFFECTIVE.

THIS IS SOMETHING THAT WAS DONE BY THE ADMINISTRATION IN CHICAGO AS A RESULT OF SOME OF THE WORK THAT STEVE LEVITT AND I DID. TAKING SUSPECTED CLASSROOMS SEVERAL WEEKS AFTER THE REGULAR TESTING, RETESTING THEM IN CONTROLLED SITUATION, AND THEN RETESTING SOME RANDOMLY SAMPLED CLASSROOMS TO THE EXTENT THAT ONE FINDS THAT THE RETEST RESULTS FOR THE SUSPECTED CLASSROOM ARE MUCH LOWER THAN THE ORIGINAL RESULTS, ONE MIGHT HAVE MORE EVIDENCE THAT THERE WAS SOME INITIAL MANIPULATION IN THE ORIGINAL EXAM ITSELF.

AND THEN FINALLY, CONDUCTING IN-DEPTH QUALITY OF INVESTIGATIONS OF SUSPECTED INDIVIDUALS. NEXT SLIDE.

SO, THERE ARE CERTAINLY SOME PRACTICAL CHALLENGES. I'M SURE ALL OF THE OTHERS IN THE AUDIENCE AND THE PANEL WHO DEAL WITH IT DESCRIBE THEM BETTER THAN I CAN, BUT THERE'S ISSUES OF DISTRICT AND STATE CAPACITY. AGENCIES OBVIOUSLY CRUNCH RESOURCES NOW, AND I THINK FINDING THE TIME AND MONEY TO DO THIS IS OBVIOUSLY DIFFICULT.

I THINK THERE ARE ISSUES THAT HAVEN'T BEEN FULLY WORKED OUT ABOUT DATA SHARING BETWEEN TEST VENDORS, STATE AGENCIES AND LOCAL DISTRICTS THAT CAN, IN CERTAIN CASES, HAMPER THE DETECTION, ANALYSIS, AND EVENTUAL DEALING WITH TESTING IRREGULARITIES. AND THEN THERE'S THE STATISTICAL FLUENCY OF THE INTENDED AUDIENCE, WHETHER THAT'S DISTRICT OR SCHOOL LEADERS OR...EVENTUAL KIND OF ADMINISTRATIVE BODIES OR EVEN COURTS. AND LAST SLIDE.

SO, THAT IS IT. THANK YOU. I LOOK FORWARD TO HEARING THE OTHER PANELISTS AND THE QUESTIONS FROM THE AUDIENCE.

**JACK BUCKLEY:** BRIAN, THANK YOU VERY MUCH.

CARSWELL?

**CARSWELL WHITEHEAD:** GOOD MORNING. THANK YOU, JACK, AND THANK YOU TO THE FOR INVITING ME HERE TODAY.

IN THE OFFICE OF TESTING INTEGRITY AT EDUCATIONAL TESTING SERVICE, WE HAVE A 3-STEP SECURITY PROCESS TO SAFEGUARD THE MILLIONS OF TESTS THAT WE ADMINISTER ANNUALLY, AND DETECTION IS A CRITICAL PART OF

THIS PROCESS, BECAUSE WE KNOW THAT EVEN WITH THE BEST PREVENTATIVE MEASURES THAT CHEATING WILL HAPPEN, SO ONE MUST BE PREPARED TO DETECT IT WHEN IT DOES. TODAY, I'LL SHARE WITH YOU THE DETECTION PROCESS.

ALL HIGH-STAKES TESTS ARE SUBJECT TO TEST TAKERS AND PROFESSIONALS WHO SEEK TO GAIN OR PROVIDE AN UNFAIR ADVANTAGE. EXPERIENCE HAS TAUGHT US THAT THIS TYPE OF BEHAVIOR CAN BE DRIVEN BY ONE'S DESIRE TO ACHIEVE HIGHER TEST SCORES IN ORDER TO GAIN ENTRANCE INTO COLLEGE. IT CAN BE ALSO DRIVEN BY ONE'S DESIRE TO HAVE ONE'S STUDENT TO DISPLAY PROGRESS AS A RESULT OF HIGH SCORES. IT ALSO CAN BE TRIGGERED BECAUSE OF PERSONAL OR FINANCIAL GAIN.

TO PROTECT AGAINST TEST COMPROMISE, PROCESSES MUST BE BUILT INTO THE TESTING PLAN TO DETECT IRREGULARITIES. THE DETECTION PLAN SHOULD INCLUDE PROCEDURES TO DETECT IRREGULARITIES THAT OCCUR BEFORE THE TEST, DURING THE TEST, AND AFTER THE TEST.

THE PRE-ADMINISTRATION PROCESS. IRREGULARITIES THAT OCCUR PRIOR TO TESTING USUALLY INVOLVE TEST MATERIAL COMPROMISE AT THE TEST CENTER. THESE IRREGULARITIES CAN BE FAR-REACHING. EARLY DETECTION OF TEST SECURITY BREACHES IS ESSENTIAL TO THE CONTAINMENT OR WIDESPREAD COMPROMISE. A GOOD TEST SECURITY PLAN SHOULD INCLUDE AN

AUDIT PROCESS TO DETECT TAMPERING BY STUDENTS AND PROFESSIONALS BEFORE THE TEST IS ADMINISTERED.

THE AUDIT SHOULD OCCUR BEFORE TEST DAY OR ON THE MORNING OF THE TEST. AND IN ORDER FOR THE AUDIT TO BE EFFECTIVE, ALL TEST MATERIALS SHOULD BE INDIVIDUALLY INVENTORIED TO INSPECT FOR MISSING TEST BOOKS OR SIGNS OF TEST REPRODUCTION. SCHOOL OFFICIALS SHOULD BE NOTIFIED IMMEDIATELY OF ANY DETECTION OF IRREGULARITIES SO THAT THE OFFICIALS WILL BE ABLE TO MAKE A DECISION REGARDING TESTING.

AND LASTLY, THE MOST EFFECTIVE PROCESS IS AN UNANNOUNCED AUDIT. THIS IS BECAUSE YOU'RE ABLE TO CATCH THE PERPETRATOR OFF-GUARD.

TEST ADMINISTRATION. AN UNFAIR ADVANTAGE THAT CAN BE OBTAINED DURING TESTING EITHER TO GIVE AN UNFAIR ADVANTAGE OR TO RECEIVE AN UNFAIR ADVANTAGE CAN BE DETECTED BY THE TESTING PERSONNEL IF ACTIVE PROCTORING OCCURS. THE PROCTOR SHOULD CHECK IDENTIFICATION TO CHECK AGAINST IMPERSONATION AND TO PROTECT AGAINST IMPERSONATION. AND MONITOR THE ROOMS FOR WANDERING EYES, SUCH AS COPYING OR COMMUNICATION, SHOULD ENSURE STUDENTS ARE WORKING ON THE CORRECT SECTION OF THE TEST WHEN THERE ARE SCRAMBLED FORMS OF THE TEST, BECAUSE A STUDENT WHO'S NOT WORKING ON A CORRECT SECTION OF THE TEST COULD BE COPYING OFF OF A STUDENT WHO HAS A SCRAMBLED

FORM. OR THE STUDENT COULD BE WORKING ADDITIONAL TIME ON A PARTICULAR SECTION IN ORDER TO OBTAIN AN UNFAIR ADVANTAGE.

THE PROCTOR SHOULD CHECK FOR UNAUTHORIZED AIDS, SUCH AS A CELL PHONE, TO TRANSMIT AN ANSWER KEY OR TO RECEIVE AN ANSWER KEY; SHOULD DOCUMENT ANY UNAUTHORIZED TEST MATERIAL ACCESS OR UNUSUAL BEHAVIOR BY SCHOOL PERSONNEL, SUCH AS WHAT MANY OF YOU MENTIONED EARLIER TODAY--SIGNALING TO STUDENTS THE CORRECT RESPONSE OR INTENTIONALLY OVERTIMING THE TEST.

THE PROCTOR SHOULD ALSO, WHEN COLLECTING THE ANSWER SHEET, INSPECT THE ANSWER SHEET TO CHECK FOR IMPERSONATION OR THE SWITCHING OF THE ANSWER SHEET.

SO YOU SEE, IN THIS PROCESS, ACTIVE PROCTORING IS KEY. I HEARD MANY OF YOU GUYS MENTION ABOUT THE TRAINING OF THE PROCTORS AND HAVING PROCTORS TO BE AWARE. SO OBTAINING AN UNFAIR ADVANTAGE ACTUALLY CAN BE A PREVENTIVE IF THE PROCTOR IS ABLE TO OBSERVE THAT DURING THE ADMINISTRATION. SO IT SHOULD BE A GOOD EMPHASIS MADE TO THE TESTING STAFF TO BE VIGILANT AND TO NOT DO NON-TEST-RELATED ACTIVITIES DURING THE TEST SO THAT THEY WILL BE ABLE TO DETECT THESE IRREGULARITIES.

THE POST-ADMINISTRATION PROCESS.

IRREGULARITIES CAN BE DISCOVERED AFTER THE TEST AS WELL THROUGH THE SCORING PROCESS. THE SECURITY PLAN SHOULD INCLUDE POST-ANALYTICS TO CHECK FOR IRREGULARITIES SUCH AS UNUSUAL ERASURES, WHICH MANY OF YOU GUYS MENTIONED TODAY, AND I'M LOOKING AT WIDESPREAD UNUSUAL ERASURES, WHICH COULD BE AN INDICATION OF POSSIBLE TAMPERING OF THE ANSWER SHEET.

HANDWRITING OR GRIDDING IRREGULARITIES, IF WIDESPREAD, CAN BE AN INDICATION, ALSO, OF TAMPERING OF THE ANSWER SHEET. ESSAY SIMILARITY, IF WIDESPREAD, COULD BE AN INDICATION OF PRE-KNOWLEDGE. INCONSISTENT

PERFORMANCE IN A PARTICULAR SECTION, IF THERE'S A UNIQUE SECTION OF THE TEST, COULD BE AN INDICATION OF TAMPERING OF THE ANSWER SHEET AND ALSO COPYING.

UNUSUAL ANSWER PATTERNS, IF WIDESPREAD, CAN BE AN INDICATION OF POSSIBLE PRE-KNOWLEDGE AND IT ALSO CAN BE AN INDICATION OF ANSWER SHEET TAMPERING. TEST COMPLETION TIME AND IRREGULARITIES--THIS IS FOR YOUR COMPUTER-BASED TESTING. IF A STUDENT COMPLETES A TEST EXCEPTIONALLY FAST, THIS MIGHT BE AN INDICATION THAT THAT STUDENT HAD AN ANSWER KEY.

AND LASTLY, A LARGE SCORE DIFFERENCE FOR A REPEATER TEST OR SPIKE IN THE MEAN TEST SCORE—THESE ANALYSES CAN BE USED, SOME OF THEM, INDIVIDUALLY, AND MANY OF THEM COULD BE USED COLLECTIVELY IN THE DECISION-MAKING PROCESS.

WHEN YOU LOOK AT THESE ANALYSES, YOU HAVE TO DETERMINE "WHAT DOES THIS DATA MEAN? WHAT IS THIS DATA TELLING ME?" SO, THAT'S IMPORTANT. IT'S JUST NOT ABOUT CONDUCTING THE ANALYSES, BUT WHAT IS IT TELLING ME. SO, FOR INSTANCE, IF YOU HAVE UNUSUAL ERASURES ON A LARGE NUMBER OF ANSWER SHEETS, YOU DON'T STOP THERE. YOU CHECK TO SEE IF THERE'S AN UNUSUAL ANSWER PATTERN, AND THAT MAY HELP YOU TO RENDER A DECISION, WHETHER OR NOT A STUDENT OR WHETHER OR NOT A PROFESSIONAL PROVIDED AN UNFAIR ADVANTAGE.

INTERPRETATION OF THE DATA.

THE TEST SECURITY PLAN SHOULD HAVE ESTABLISHED GUIDELINES TO ASSIST WITH THE DECISION-MAKING PROCESS. USE DETECTED IRREGULARITIES AS FLAGS. AND I AGREE WITH A LOT OF YOU GUYS.

THESE ARE FLAGS THAT REQUIRE ADDITIONAL INVESTIGATIONS TO DETERMINE IF THERE IS A CONCERN. AND WE ALL KNOW WHEN ANALYZING DATA THAT THIS IS AN ART. NO STATISTICAL EVIDENCE IS TRULY POSITIVE. WE NEED TO—IN REVIEWING THIS EVIDENCE, WE HAVE TO MAKE

JUDGMENT CALLS. AND THEN WE ALSO HAVE TO KNOW THAT BEHIND ANY ANALYSIS WILL BE POLICY QUESTIONS.

IN CONCLUSION, DO THESE DETECTION PRACTICES WORK? YES, THEY DO. I DON'T DO A LOT OF INVESTIGATION FOR K-12. I DO INDIVIDUAL INVESTIGATIONS FOR TEST PAPERS FOR SCORE INVALIDITY, BUT I HAVE BEEN INVOLVED IN INVESTIGATIONS FOR K-12. AND I CAN TELL YOU THAT THESE ANALYSES AND THESE PROCESSES FOR DETECTING IRREGULARITIES HAVE BEEN VERY EFFECTIVE AT CONTAINING TEST COMPROMISE--SO INSTEAD OF MILLIONS OF TEST TAKERS BEING IMPACTED, MAYBE IT'S LIMITED TO A PARTICULAR SCHOOL; FOR PROVIDING EVIDENCE TO CANCEL INVALID TEST SCORES; FOR MITIGATING FURTHER RISK; AND FOR HELPING TO REFINE PROCEDURES.

AND WHAT I WANT TO MENTION IS THAT WHEN AN IRREGULARITY IS DETECTED, IT'S KEY FOR THOSE WHO ARE IN CHARGE TO SHARE THIS INFORMATION TO EITHER THE SCHOOL DISTRICT OR TO THE STATE AND FOR AN INVESTIGATION TO OCCUR. BECAUSE THE ONLY WAY THAT YOU'LL KNOW IF ANYONE IS PROVIDING AN UNFAIR ADVANTAGE IS TO CONDUCT AN INVESTIGATION. AND EVEN WHEN I REVIEW TEST SCORES ON AN INDIVIDUAL BASIS, WE RECEIVE INQUIRIES INTO OUR OFFICE VIA OUR SECURITY HOTLINE. IT COULD BE AN INQUIRY FROM SOMEONE FROM THE SCHOOL DISTRICT. IT CAN BE AN INQUIRY FROM ANYONE WHO MAY HAVE KNOWLEDGE THAT SOMEONE PROVIDED AN UNFAIR ADVANTAGE.

WHEN WE RECEIVE THAT INFORMATION, WE CONDUCT AN INVESTIGATION. AND IF NOTHING IS FOUND, WE DON'T ACT ON IT. BUT IF SOMETHING IS FOUND, WE ARE OBLIGATED TO MAKE SURE THAT THOSE SCORES DO NOT STAND.

THANK YOU.

**JACK BUCKLEY:** THANKS VERY MUCH. DAVID?

**DAVID FOSTER:** I APPRECIATE THE OPPORTUNITY TO PRESENT TODAY AND PARTICIPATE IN THIS PANEL. IT'S BEEN VERY INFORMATIVE SO FAR. MY PRESENTATION'S GOING TO BE SOMEWHAT HIGH-LEVEL. I THINK I PRESENT BEST PRACTICES PROBABLY ONLY AS EXAMPLES AMONG THE SLIDES, BUT LET'S START OFF WITH--IN DETECTION, WHAT ARE WE LOOKING FOR?

WE'RE LOOKING FOR AN ACTUAL BREACH, WHETHER IT'S IN PROGRESS, OCCURRED SOME TIME IN THE PAST, OR IT'S JUST A THREAT--IT MAY COME DOWN THE PIPE PRETTY SOON. THE GOOD NEWS IS THERE ARE ONLY TWO TYPES OF THESE BREACHES AND THREATS.

THERE'S CHEATING, WHICH I'VE BEEN TALKING ABOUT AS INAPPROPRIATE ATTEMPTS TO INCREASE SCORES; AND THEFT, WHICH DOESN'T CARE A WHOLE LOT ABOUT THE IMMEDIATE PROSPECT OF INCREASING SCORES--LATER, OF

COURSE--BUT STEALING TEST CONTENT FOR GAIN. THE BAD NEWS IS THAT THERE ARE HUNDREDS AND PERHAPS THOUSANDS OF WAYS TO CHEAT ON EXAMS OR TO STEAL AND USE THE CONTENT.

SO, THESE ARE A FEW EXAMPLES OF SOME TYPICAL SECURITY THREATS OR BREACHES. ON THE CHEATING SIDE, TEACHER TAMPERING WITH ANSWER SHEETS, TEACHERS HELPING STUDENTS DURING THE EXAM, TEACHERS PREPPING THE STUDENTS PRIOR TO THE EXAM, HAVING CERTAIN STUDENTS REMAIN AT HOME ON TEST DAYS. THEN YOU CAN SWITCH THE ATTENTION TO STUDENTS. CHEATING ON TESTS USING PRE-KNOWLEDGE, ASSISTANCE, CHEAT SHEETS, OTHER TECHNOLOGIES OR OTHER METHODS. BUT THEN YOU ALSO HAVE THE THEFT THREATS. A TEACHER OR A STUDENT CAN STEAL BOOKLETS. A TEACHER OR STUDENT CAN RECORD QUESTIONS THROUGHOUT THE EXAM.

SO, TODAY, I'M GONNA DISCUSS EACH OF THESE IN ORDER--7 DETECTION PRINCIPLES. THERE COULD BE MORE. IN FACT, SINCE I'VE SUBMITTED THIS, I'VE THOUGHT OF 2 OR 3 OTHERS. THERE JUST WASN'T TIME. BUT FOCUS. CONCENTRATE ON THE HIGH-RISK THREATS FIRST. ALMOST IGNORE THOSE THAT ARE INFREQUENT OR ARE LESS OF A THREAT.

ADAPT. YOU NEED TO MATCH YOUR DETECTION METHODS TO THE THREAT. OFTEN WE LOOK IN DIFFERENT DIRECTIONS.

BACK-UP. USING SEVERAL LAYERS OF DETECTION METHODS.

PREDICT. WATCH OUT FOR THE FUTURE. WHAT CAN WE EXPECT NEXT? LET'S NOT GET SURPRISED.

FILTER. RULE OUT OTHER EXPLANATIONS.

EVALUATE. USE YOUR DETECTION METHODS TO EVALUATE WHAT YOU ARE DOING AND THEN PLANNING A LITTLE MORE SPECIFICALLY.

SO, LET'S GO THROUGH THESE.

SO, FOCUS. CONCENTRATE ON THE HIGH-RISK THREATS FIRST. WHAT CONSTITUTES A HIGH-RISK THREAT OR A HIGH-DAMAGING BREACH? THE LIKELIHOOD OF THE THREAT, HOW IMMEDIATE IT IS. WHAT'S THE AMOUNT OF POTENTIAL DAMAGE IT CAN CAUSE? HOW DIFFICULT IS IT TO PREVENT? MAYBE IT'S BETTER TO--IF YOU CAN'T PREVENT IT, YOU MAY WANT TO DETECT IT. IT'S DIFFICULT TO DETECT AND MITIGATE AND/OR IT USES VULNERABILITIES IN OUR SYSTEMS.

FOR EXAMPLE, ONE TEACHER CHEATING BRINGS MUCH GREATER DAMAGE IN A HIGHER RISK THAN ONE STUDENT TEACHING--I'M SORRY. ONE STUDENT CHEATING.

ADAPT. MATCH YOUR DETECTION METHODS TO THE THREAT. NOW, THREATS DEPEND ON A NUMBER OF CIRCUMSTANCES. THEY'RE NOT THE SAME AT GRADE LEVELS. THREATS AT THIRD GRADE ARE DIFFERENT THAN THREATS AT 11<sup>TH</sup> GRADE. THEY'RE DIFFERENT FOR SCIENCE AND FOR MATH. THE DELIVERY MODE, WHETHER IT'S PAPER OR PENCIL OR COMPUTERIZED TESTING AFFECTS THE RISK. THE TESTING HISTORY OF A SCHOOL, A TEACHER, A STUDENT MAY CHANGE THE WAY YOU APPROACH DETECTION.

AND MOTIVATION. OFTEN, TESTS THAT WEREN'T INTENDED TO BE HIGH STAKES HAVE BECOME HIGH STAKES. AND NOW YOU NEED TO FOCUS ATTENTION ON THAT. WE OFTEN TEND TO USE EITHER ONE TOOL FOR DETECTION OR FOCUS WHAT WE HAVE IN THE WRONG DIRECTION. AND I MENTION THAT EXAMPLE THAT THE THREATS AT 3<sup>RD</sup> GRADE MAY INVOLVE MORE TEACHER CHEATING, WHEREAS THREATS AT 11<sup>TH</sup> GRADE, BECAUSE THE STAKES ARE HIGHER FOR THE STUDENTS, MAY INVOLVE--THE RISK MAY BE MUCH MORE WITH STUDENT CHEATING.

BACK UP. VISA USES SEVERAL METHODS TO DETECT IF YOUR CARD IS BEING USED INAPPROPRIATELY. LIKewise, WE SHOULD USE SEVERAL LAYERS OF DETECTION METHODS, AND I THINK DR. JACOB'S ALSO MENTIONED THIS. I'LL DO IT IN A LITTLE BROADER LEVEL, THOUGH. USE FORENSIC STATISTICAL METHODS AS HAVE BEEN DISCUSSED.

CONTINUOUS OBSERVATION. WHAT DOES IT TAKE TO WATCH A STUDENT WHOM YOU MAY SUSPECT OF TAMPERING WITH ANSWER SHEETS? IT'S KIND OF LIKE PROCTORING THE PROCTORS. USE OF A TIPLINE. I'VE BEEN SURPRISED RECENTLY BY THE NUMBER OF...CHEATING AND TESTING AND THEFT PROBLEMS THAT HAVE COME TO OUR ATTENTION SIMPLY BECAUSE OF TIPS.

WEB MONITORING. QUESTIONS THAT ARE STOLEN ARE ALMOST ALWAYS SHARED ACROSS THE WEB. THAT'S ONE WAY TO FIND OUT ABOUT THEM. SO, A TEACHER CHEATING—I THINK ALL OF THESE 4 METHODS HERE AND MAYBE OTHERS THAT I HAVEN'T LISTED SHOULD BE ENLISTED TO DETECT THE PROBLEM.

AND JUST A LITTLE ASIDE, IT'S PROBABLY BETTER TO DETECT THESE THINGS EARLY THAN LATE. SOMETIMES WE DON'T DETECT PROBLEMS UNTIL SEVERAL WEEKS, MAYBE SEVERAL MONTHS AFTER THE PROBLEM HAS OCCURRED. IF YOU CAN DETECT THE THREAT AND MITIGATE THE THREAT OR DETECT THE PROBLEM IN PROCESS, THE BREACH IN PROCESS, SECURITY WILL BE ENHANCED.

PREDICT. WATCH OUT FOR NEW THREATS. THEY COME ALONG EVERY DAY. CHANGING STAKES WILL AFFECT THOSE. NEW TECHNOLOGIES MAKES CHEATING AND THEFT LESS DETECTABLE. I MEAN, THE SMALL LITTLE 2-WAY RADIOS IN THE EAR—THEY CAN'T BE SEEN BY TEACHERS, BUT A STUDENT CAN SPEAK AND

RECEIVE INFORMATION DURING AN EXAM. THE CREATIVITY OF CHEATERS AND THIEVES IS ALWAYS HIGH.

AND I JUST PUT AN EXAMPLE HERE. THEY'RE ON THE MARKET FOR LESS THAN \$100. YOU CAN BUY THESE DEVICES THAT HAVE LENSES IN THEM. THEY CAN HOLD UP TO 16 GIGABYTES OF INFORMATION AND ESSENTIALLY RECORD AN ENTIRE SET OF EXAM CONTENT AND BE PRETTY MUCH UNDETECTABLE BY ANYBODY WALKING AROUND THE ROOM. I DON'T HAVE A GOOD SOLUTION TO THOSE, BUT THAT'S NOT THE POINT OF MY TALK TODAY ANYWAY, SO...

FILTER. THIS HAS BEEN MENTIONED A COUPLE OF TIMES. RULE OUT OTHER EXPLANATIONS. FOR FAIRNESS PURPOSES, YOU DON'T WANT TO ACCUSE SOMEONE OF CHEATING OR SPEND A LOT OF EFFORT IN THAT DIRECTION WHEN, IN FACT, CHEATING DIDN'T HAPPEN.

TESTING IRREGULARITIES, FROM THE GRAPHIC ON THE SLIDE, ARE MANY. THOSE SHOULD BE FILTERED AND EFFORTS SHOULD BE MADE TO DETECT AND RESPOND TO ACTUAL CHEATING OR OTHER KIND OF TEST FRAUD EVENTS. AND THE EXAMPLE I APPLY HERE, I SAW JUST RECENTLY.

A STATISTICAL ANALYSIS DETECTED A KIND OF COLLUSION AMONG A NUMBER OF STUDENTS IN A CLASSROOM, WHEN REALLY ALL THEY DID WAS DECIDE TO FINISH UP THE TEST BY EVERYBODY ANSWERING "C." IT WASN'T

CHEATING. IT WAS BAD TESTING PRACTICE, BUT NOT AN EXAMPLE OF CHEATING. IT WAS CERTAINLY A TESTING IRREGULARITY.

EVALUATE. YOU SHOULD USE DETECTION METHODS TO EVALUATE YOUR SECURITY--BOTH HOW WELL YOU'RE DETECTING AS WELL AS OTHER METHODS. THIS IS A GRAPH OF A STATE GIVING THE SAME STATISTICAL ANALYSIS OVER 5 SEPARATE YEARS IN THE SPRING. AND IT SHOWS A DROP IN DETECTED PROBLEMS, BECAUSE THEY INSTITUTED SCORE INVALIDATIONS AND BETTER TRAINING PROCEDURES AFTER THAT FIRST YEAR.

PLAN. SET UP TO SUCCEED. THIS IS AN INTERESTING--YOU CAN MODIFY AND SHOULD MODIFY YOUR DETECTION METHODS TO DETECT PARTICULAR PROBLEMS. PRE-KNOWLEDGE IS VERY DIFFICULT TO DETECT. NO ONE CAN SEE THAT A STUDENT COMES TO A TEST WITH ALL THE ANSWERS IN THEIR HEADS. YOU CAN, HOWEVER, PUT IN SPECIFIC QUESTIONS IN THE EXAM TO TELL WHETHER THEY'RE USING PRE-KNOWLEDGE OR NOT.

NOW, IT'S A TEST-DESIGN COLLABORATION WITH THE SECURITY FOLKS TO GO AFTER A PARTICULAR THREAT. BUT IN THIS CASE, THEY WERE USED IN 88% OF THE EXAMS, AND THIS KIND OF METHOD COULD PRETTY MUCH DETECT EVERY INSTANCE.

JUST A LITTLE BIT OF AN EXAMPLE HOW THEY MIGHT DIFFER, THEN I'M DONE. HIGH RISK AT GRADE 3 MIGHT BE TEACHER TAMPERING WITH ANSWER SHEETS. I THINK THIS HAS BEEN MENTIONED SEVERAL TIMES TODAY. DIRECT DETECTION WOULD BE TO PHYSICALLY OBSERVE THE TEACHER. IF THE TEACHER CAN'T BE REMOVED AND SOMEONE ELSE REPLACED, THEN OBSERVE THAT TEACHER FROM THE TIME YOU GATHER UP THE ANSWER SHEETS TO WHEN THEY HAND THEM OFF. IF THEY ATTEMPT SOMETHING, YOU'LL DETECT IT. INDIRECTLY, YOU CAN SEE THAT THROUGH FORENSICS EFFORTS, UNUSUALLY HIGH RATES OF WRONG TO RIGHT ERASURES, AND UNEXPECTED SCORE GAINS WILL GIVE SOME INDIRECT EVIDENCE THAT'S HAPPENING.

ANOTHER INDIRECT METHOD IS TO INSTITUTE A TIPLINE, WHERE SOMEONE WILL LET YOU KNOW THIS IS HAPPENING. AND THE LAST ONE IS 11<sup>TH</sup> GRADE, JUST AS A DIFFERENCE. STUDENT CHEATING BY COPYING FROM ANOTHER STUDENT.

DIRECT. YOU USE TRAINED PROCTORS WHO VIGILANTLY OBSERVE THE STUDENTS DURING THE EXAM. THAT'S A GOLD STANDARD METHOD.

INDIRECTLY. SIMILARITY AND COLLUSION ANALYSES AND MAYBE LOOKING AT SOME ABERRANT RESPONDING. THEY DON'T ALWAYS GET TO SEE WHAT THEIR STUDENTS ARE DOING.

AND FINALLY, AGAIN, A TIPLINE IS OF VALUE. SO, THANK YOU VERY MUCH.

**JACK BUCKLEY:** DAVE, THANK YOU VERY MUCH.

AND MELISSA?

**MELISSA FINCHER:** OK.

GOOD MORNING. HOPE TO BE ABLE TO SHARE SOME OF THE EXPERIENCES WE'VE GONE THROUGH IN GEORGIA, AND I HOPE THE INFORMATION IS HELPFUL.

IN GEORGIA, WE HAD BROADBASED COMMITMENT OF MULTIPLE AGENCIES. OF COURSE, THE GEORGIA DEPARTMENT OF EDUCATION WAS INVOLVED, AS WE'RE RESPONSIBLE FOR THE DEVELOPMENT, ADMINISTRATION, SCORING, AND REPORTING OF THE STATE TESTS, INCLUDING SECURITY PROCEDURES AND PROTOCOLS. THE OFFICE OF STUDENT ACHIEVEMENT ACTUALLY LED THE WORK INTO THE DETECTION IN THE STATE OF GEORGIA. THEY ARE STATUTORIALLY CHARGED WITH THE AUDIT FUNCTION, AND WE WORKED VERY CLOSELY WITH THE OFFICE OF STUDENT ACHIEVEMENT.

THE GOVERNOR'S OFFICE BECAME INVOLVED, AND THAT WAS VERY IMPORTANT. THE GOVERNOR APPOINTED SPECIAL INVESTIGATORS, AND LATER ON ANOTHER PANEL, YOU'LL HEAR FROM ONE OF OUR SPECIAL INVESTIGATORS.

THE PROFESSIONAL STANDARDS COMMISSION IN THE STATE OF GEORGIA IS ALSO INVOLVED IN THIS WORK IN THAT THEY HAVE ETHICS STANDARDS FOR EDUCATORS, AND THEY ALSO ARE THE CREDENTIALING AGENCIES. AND IN GEORGIA, THERE ARE STAKES ASSOCIATED WITH INAPPROPRIATE PRACTICES DURING ASSESSMENT. WE DO REQUIRE THAT ONLY CERTIFIED EDUCATORS ADMINISTER OUR STATE ASSESSMENT PROGRAM AS ONE DETERRENT.

IN TERMS OF THINGS THAT STATES NEED TO CONSIDER, IT'S VERY IMPORTANT TO UNDERSTAND THE TYPES OF ANALYSES THAT ARE AVAILABLE AND WHAT INFORMATION THEY CAN CONTRIBUTE. YOU HAVE TO LOOK AT THE STRUCTURE OF YOUR ASSESSMENT PROGRAM, THE INFORMATION THAT YOU CAPTURE.

FOR INSTANCE, WE WERE ABLE TO KNOW WHO ADMINISTERED THE TEST. WHAT'S DIFFERENT IN GEORGIA--WE DO HAVE THE HEADER SHEETS THAT DR. CIZEK TALKED ABOUT, BUT WE DO ALLOW SYSTEMS TO GROUP THEIR ANSWER DOCUMENTS IN A WAY THAT IS MEANINGFUL FOR THEM, SO THAT THE RESULTS COME BACK IN A STRUCTURALLY MEANINGFUL WAY.

HOWEVER, SYSTEMS ARE RESPONSIBLE FOR MAINTAINING RECORDS ON WHO ADMINISTERED THE TEST AND HOW STUDENTS WERE GROUPED. IT'S ALSO IMPORTANT TO UNDERSTAND, AS YOU'VE HEARD THROUGHOUT THIS

MORNING, THAT THERE'S NO SINGLE ANALYSIS OR COMBINATION OF ANALYSES THAT IS DEFINITIVE. YOU DON'T GET THE ANSWER, ADDITIONAL INVESTIGATION IS GOING TO BE WARRANTED. THAT IS SOMETHING THAT YOU CAN COUNT ON.

HAVING A CLEAR RATIONALE FOR THE FLAGGING CRITERIA THAT YOU USE IS VERY IMPORTANT, AND YOU NEED TO THINK AHEAD OF TIME OF HOW YOU'RE GOING TO FLAG, AT WHAT LEVEL YOU'RE GOING TO FLAG.

THAT FLAGGING CRITERIA WILL, IN TURN, BECOME A WAY TO INTERPRET THE RESULTS, AND IT WILL BE USED BY DISTRICTS AND SCHOOLS IN UNDERSTANDING. WITH THAT SAID, IT'S VERY IMPORTANT TO BE VERY CLEAR ABOUT WHAT YOU'RE FLAGGING, AND THAT'S SOMETHING THAT WE'VE HEARD, AGAIN, FREQUENTLY THIS MORNING.

WHAT YOU'RE FLAGGING ARE DATA ANOMALIES. YOU'RE NOT FLAGGING SOMETHING THAT'S DEFINITIVE. YOU'RE FLAGGING SOMETHING THAT NEEDS TO BE INVESTIGATED FURTHER. THERE IS A REASON IN GEORGIA WE WANT STUDENTS TO USE NUMBER 2 PENCILS. WE DID ERASURE ANALYSES. WE WANT STUDENTS TO HAVE THAT ERASURE AS A TOOL THAT'S NEEDED IN A LEGITIMATE MANNER. AS YOU CAN IMAGINE, ONE OF THE FIRST REACTIONS THAT WE GOT WAS, "WE'RE JUST GONNA REMOVE THE ERASERS." THAT'S NOT A GOOD SOLUTION.

[LAUGHTER]

AS YOU'RE WORKING THROUGH THIS, AS YOU'RE CHOOSING YOUR CONTRACTOR, MAKE SURE THAT THEY HAVE SUFFICIENT QUALITY CONTROL MECHANISMS IN PLACE. YOU NEED TO PLAN FOR INDEPENDENT VERIFICATION AND REPLICATION OF THE ANALYSES. YOU CAN EXPECT TO BE CHALLENGED.

AS ASSESSMENT FOLKS, WE'RE USED TO THAT, BUT THERE'S AN EMOTIONAL COMPONENT TO THIS WORK AS WELL. AS MUCH AS YOU CAN--THIS IS COMPLICATED WORK, AND IT'S DIRTY WORK AND IT'S MESSY WORK--TRY YOUR BEST TO KEEP THINGS SIMPLE, STRAIGHTFORWARD, AND CONSISTENT.

COMMUNICATION IS KEY HERE. THESE ARE TERMS AND ANALYSES THAT WE'RE FAMILIAR WITH. WE DEAL WITH STANDARD DEVIATIONS QUITE A BIT IN OUR WORK, BUT AT THE SCHOOL LEVEL THEY DON'T. SO THEY NEED A REFRESHER COURSE AND NEED A LENS TO UNDERSTAND THAT WORK.

SO, THEY'LL NEED QUITE A BIT OF HELP. DISTRICTS AND SCHOOLS WILL NEED QUITE A BIT OF SUPPORT AND HELP IN UNDERSTANDING AND INTERPRETING THE DATA. OPEN AND COMPREHENSIVE COMMUNICATION IS ESSENTIAL AND BEING CONSISTENT WITH THAT COMMUNICATION. YOU CAN EXPECT THINGS TO RAMP UP AND GET COMPLICATED REALLY, REALLY QUICKLY. IT'S ALSO IMPORTANT THAT YOU'RE VERY CLEAR ON WHAT DATA CAN BE SHARED AND WHAT DATA CAN'T BE SHARED. THIS IS BASED ON

SECURE ASSESSMENT DATA, AND SO THERE ARE SOME PIECES THAT CAN'T BE SHARED.

IT'S IMPORTANT TO HAVE A VERY CLEAR PLAN FOR NEXT STEPS. THE ANALYSIS IS JUST ONE PART. SO, I MENTIONED THE FACT THAT FOLLOW-UP INVESTIGATIONS ARE GOING TO BE REQUIRED. HAVE A PLAN FOR WHO WILL CONDUCT THOSE. IF IT IS THE DISTRICT, YOU NEED TO EXPECT THAT THEY WILL NEED SUPPORT. THEY WON'T KNOW WHERE TO BEGIN, NECESSARILY, BECAUSE THIS WILL BE NEW WORK TO THEM. THEY'VE INVESTIGATED INDIVIDUAL CASES IN THE PAST, BUT MAYBE NOT NECESSARILY THINGS ON THE MAGNITUDE OF WHAT WE SAW IN THE STATE OF GEORGIA.

CERTAINLY, THE CHAIN OF CUSTODY IS CRITICAL, BUT THAT SHOULD NOT BE THE SOLE FOCUS. AND THEN, QUITE FRANKLY, YOU NEED TO HAVE A PLAN FOR WHAT WE ENCOUNTERED IN GEORGIA, AND THAT IS WHEN THE DISTRICT REFUSES TO INVESTIGATE OR DOESN'T PUT FORTH THEIR BEST EFFORT. MAKE SURE THAT YOUR ASSESSMENT POLICIES ADDRESS RECORDS RETENTION. I TALKED ABOUT THAT A LITTLE BIT ALREADY IN MAKING SURE THAT THOSE RECORDS ARE RETAINED. WHO ADMINISTERED THE TEST? TO WHICH STUDENTS? WHO PROCTORED? WE'VE SEEN SEATING CHARTS, SOMETHING THAT WE PROBABLY NEED TO ADD IN THE STATE OF GEORGIA. WHO HAD POSSESSION OF THE SECURE MATERIALS AND FOR HOW LONG? WE HAVE CHECK-IN AND CHECK-OUT PROCEDURES RECORDING TIMES SO THAT IF MATERIALS AREN'T RETURNED FOR AN HOUR AFTER ADMINISTRATION, THAT IS

DOCUMENTED AND THAT'S SOMETHING THAT CAN BE INVESTIGATED.

KNOWING WHO HAS ACCESS TO THAT SECURE LOCATION WHERE MATERIALS ARE STORED OVERNIGHT, THAT KIND OF THING. THOSE ARE VERY IMPORTANT PROTOCOLS THAT NEED TO BE IN PLACE, AND THOSE RECORDS NEED TO BE MAINTAINED.

THINK ABOUT THE CHANGES IN YOUR TEST PROCEDURES THAT WILL TAKE PLACE DURING THE NEXT ADMINISTRATION. IN GEORGIA, THE OFFICE OF STUDENT ACHIEVEMENT REQUIRED THE ROTATION OF TEACHERS IN SCHOOLS THAT WERE FLAGGED SEVERE AND MODERATE. WE SENT IN STATE MONITORS. WE'D ALWAYS DONE RANDOM CHECKS OF STATE MONITORS, BUT HERE WAS A CONCERTED EFFORT TO SEND IN STATE MONITORS FOR THE DURATION OF THE ADMINISTRATION. THE MONITORS OBSERVE THE TEST ENVIRONMENT, AND AT THE CONCLUSION OF TESTING EACH DAY, THEY SECURED THE ANSWER DOCUMENTS THEMSELVES AND WERE THERE IN THE MORNING TO OPEN AND RETRIEVE THE ANSWER DOCUMENTS.

SO WE RESTRICTED ACCESS TO THE ANSWER DOCUMENTS. AN IMPORTANT CONSIDERATION FOR STATES AND DISTRICTS GOING INTO THIS IN TERMS OF STATE MONITORS, THE SELECTION AND TRAINING OF THOSE MONITORS IS ESSENTIAL. CERTAINLY, WE DO THAT IN THE STATE OF GEORGIA, BUT IT'S IMPORTANT THAT THE MONITORS NOT BE OVERZEALOUS AND UNDERSTAND

THAT THE SCHOOL IS STILL RESPONSIBLE FOR THE ADMINISTRATION. THEY'RE THERE TO OBSERVE THE TEST ENVIRONMENT AND TO SECURE THE ANSWER DOCUMENTS AND THEN REPORT ANYTHING THAT IS OF CONCERN. YOU NEED TO TAKE IT TO THE NEXT LEVEL IF INAPPROPRIATE BEHAVIOR IS FOUND.

THINK ABOUT WHAT SUPPORTS WILL BE PROVIDED TO THE STUDENTS WHO RECEIVED INACCURATE SCORES. HOW WILL YOUR ACCOUNTABILITY DETERMINATIONS BE ADJUSTED? WHAT OTHER PROGRAMS AND DECISIONS ARE IMPACTED BY THIS BAD DATA? IT'S VERY IMPORTANT THAT YOU HAVE THAT PLANNING DONE UP FRONT. CONSIDER THE CAPACITY AT ALL LEVELS-- CERTAINLY, AT THE SCHOOL DISTRICT, STATE, AND CONTRACTOR LEVEL. THIS HAS REALLY PUSHED OUR CAPACITY. IT'S BEEN A GOOD THING, NONETHELESS. IT'S SOMETHING THAT NEEDED TO OCCUR.

CONSIDER THE TIMING OF THE REPORTS AND THE INVESTIGATIONS. IF YOU'RE LIKE MY STATE, THEY WANT THE DATA IMMEDIATELY FOLLOWING. THEY WANT IT NEXT DAY, FASTER, MORE, MORE, MORE, BECAUSE THEY NEED TO MAKE DECISIONS. WE HAVE PROMOTION AND RETENTION. SO, OFTEN, THE DATA WILL COME BACK AFTER SCORES ARE BEING REPORTED, AND IT'S VERY HARD TO INVALIDATE AND PULL BACK A SCORE ONCE IT'S BEEN RELEASED. SO, CONSIDER THAT DOMINO EFFECT.

WHO WILL NEED TO BE NOTIFIED OF THE SUBSTANTIATED DATA INACCURACIES. THERE ARE USERS OF THE DATA THAT HAVE ALREADY RECEIVED DATA FOR GRANTS OR WHAT-HAVE-YOU THAT NEED TO BE NOTIFIED THAT THE DATA IS NOT ACCURATE. AGAIN, KEEP DETAILED RECORDS OF ALL THE STEPS FOLLOWED, AND THEN THINK AS BROADLY AND AS COMPREHENSIVELY AS YOU CAN.

**JACK BUCKLEY:** MELISSA, THANK YOU VERY MUCH.

SO, AGAIN, AS BEFORE, WE'LL BEGIN BY ME TRYING TO THROW A COUPLE QUESTIONS AT THE PANEL AND BEFORE I OPEN IT UP TO THE REST OF US.

SO, DAVID, I THINK I'M STEALING THIS FROM YOU, BUT WE'RE TALKING ABOUT DETECTION. I THINK WE CAN CATEGORIZE METHODS INTO TWO BROAD FAMILIES.

FIRST, THEY'RE SORT OF DIRECT, RIGHT? SO, IN THIS CASE, WE'RE TALKING--WHAT I'M HEARING IS THINGS WE'RE DOING IN OBSERVATION OR MONITORING OR PROCTORING. BUT ALSO, CARSWELL, DISCUSSIONS ABOUT THINGS LIKE MATERIALIC INSPECTION, SO LOOKING AT SHRINK WRAP, ACTUALLY DOING PHYSICAL INSPECTION OF THE SPACE OR THE TEST MATERIALS.

AND SECONDLY, I THINK, DAVID, WHAT YOU CHARACTERIZE AS INDIRECT, WHICH IS PROBABLY WHAT MOST PEOPLE THOUGHT WE WERE GONNA SPEND MOST OF THE TIME TALKING ABOUT HERE, WHICH IS SORT OF STATISTICAL FORENSICS, TIPLINES, OTHER SORT OF REPORTING MECHANISMS.

I WANT TO GET TO MELISSA, SOMETHING YOU SAID, BECAUSE IT ALSO SPEAKS TO WHAT CHANCELLOR HENDERSON SAID THIS MORNING ABOUT TRANSPARENCY AROUND THE STANDARDS IN INDIRECT REPORTING, PARTICULARLY STATISTICAL FORENSICS.

SO YOU MENTIONED SOMETHING ABOUT THE IMPORTANCE OF COMMUNICATION AND MAKING SURE THAT YOU TELL PEOPLE ESSENTIALLY WHAT ARE YOU GONNA BE FLAGGING. WHAT IS YOUR THRESHOLD FOR DETECTING AN IRREGULARITY IN ADVANCE? I'M CURIOUS. HOW DO YOU DO THAT? HOW DO YOU SET THAT? WHAT KINDS OF COMMUNICATION METHODS DO YOU USE, AND DOES EVERYONE ELSE AGREE THAT SORT OF TRANSPARENCY THERE IS IMPORTANT, THAT WE WANT TO BE VERY CLEAR ABOUT WHAT WE'RE USING TO LEAs AND TO SCHOOLS, OR IS THERE A COUNTERPOINT WHERE MAYBE WE DON'T WANT TO TELL THEM NECESSARILY WHAT THE THRESHOLDS ARE IN ADVANCE?

**MELISSA FINCHER:** I THINK IN GEORGIA, IN ALL HONESTY--I DON'T MEAN TO SPEAK FOR THE OFFICE OF STUDENT ACHIEVEMENT, BUT THEY ARE THE ONES THAT SET THAT FLAGGING CRITERIA AND CARRY THE BURDEN OF MUCH OF THAT COMMUNICATION. TO A CERTAIN EXTENT, IT WAS TRIAL AND ERROR.

THE FIRST TIME OUT. IT WAS LOOKING AT THE DATA, TRYING TO MAKE SENSE OF IT, FIELDING QUESTIONS FROM THE FIELD IN WAYS TO HELP THEM BETTER UNDERSTAND. YOU HAVE TO LOOK AT THIS AS A PROCESS OF CONTINUAL IMPROVEMENT, BUT I THINK IT IS IMPORTANT SINCE THOSE FLAGS ARE GOING TO BE USED AS INTERPRETATION TO BE VERY CLEAR ABOUT WHAT THEY MEAN. I'M NOT SURE THAT WE HAVE IT RIGHT YET. WE USE 3 STANDARD DEVIATIONS IN THE STATE OF GEORGIA. I NOTICE THAT SEVERAL STATES USE 4 STANDARD DEVIATIONS. I DON'T KNOW THAT WE DON'T HAVE IT RIGHT, BUT HAVING A RATIONALE FOR WHAT YOU'RE USING IS VERY IMPORTANT.

**JACK BUCKLEY:** AND WHEN YOU SAY 3 STANDARD DEVIATIONS, WE'RE TALKING ABOUT--

**MELISSA FINCHER:** WRONG TO RIGHT.

**JACK BUCKLEY:** FREQUENCY OF ERASURES OR WTR, WRONG TO RIGHT ERASURES IN THAT CASE.

**MELISSA FINCHER:** YES.

**JACK BUCKLEY:** BRIAN, YOU MENTIONED THE NEED FOR ROBUST STATISTICS WHEN

YOU'RE SETTING THRESHOLDS. WE'RE WORRIED ABOUT OUTLIERS THAT MIGHT BE--I'M SORRY. IN ONE OF YOUR POINTS, YOU HAD MENTIONED THE NEED OR RECOMMENDED THE USE OF ROBUST STATISTICS IN TERMS OF STATISTICAL FORENSICS. SO ANY PARTICULAR RECOMMENDATIONS? AGAIN? SO, WE WOULD SEND SOMETHING--

**BRIAN JACOB:** [INDISTINCT] I THINK AS MUCH AS KIND OF ONE SINGLE MEASURE WOULD BE THE GENERAL RULE OF TESTING [INDISTINCT] THE SENSITIVITY OF FLAGGING OR OTHER ANALYSIS TO A VARIETY OF APPROACHES. FOR EXAMPLE, IF YOU'RE LOOKING AT CLASSROOM OR SCHOOL STATISTICS OR NUMBER OF WRONG TO RIGHT, PROBABLY THE MEDIAN RATHER THAN THE MEAN NUMBER OF WRONG TO RIGHT WOULD BE A MORE ROBUST STATISTIC. NOW, AT THE SCHOOL LEVEL, IT MADE A WEALTH OF DIFFERENCE FOR THIS VERY LARGE SAMPLE, BUT IN SOME SMALL CLASSES WHERE YOU ONLY HAVE 10 OR 11 STUDENTS, YOUR ONE STUDENT GOING OFF GRID AND HAVING TO ERASE HALF THE TEST CAN CHANGE YOUR AVERAGE WRONG TO RIGHT FOR THE ENTIRE CLASSROOM. NOW, YOU EITHER USE THE [INDISTINCT] LIKE THAT OR HOPEFULLY THAT SORT OF THING IS CAUGHT IN KIND OF A FOLLOW-UP INVESTIGATION WHERE YOUR [INDISTINCT] IRREGULARITY.

**JACK BUCKLEY:** ANY OTHER THOUGHTS ON TRANSPARENCY HERE IN TERMS OF WHAT SHOULD YOU COMMUNICATE TO THE FOLKS WHO ARE ACTUALLY

ADMINISTERING THE ASSESSMENT IN TERMS OF WHAT YOU WILL BE FLAGGING OR LOOKING AT? DAVID, WHAT DO YOU GUYS RECOMMEND?

**DAVID FOSTER:** I'M A BIG FAN OF TRANSPARENCY AND DISCLOSURE MAINLY BECAUSE IT DETERS AND PREVENTS CHEATING. IF THEY KNOW YOU HAVE A STRONG, SOPHISTICATED, COMPREHENSIVE SYSTEM IN PLACE, PEOPLE ARE LESS LIKELY TO TRY TO GAME AND GO UP AGAINST IT, BECAUSE THE FEAR WILL BE--SO, THE MORE YOU CAN PROVIDE INFORMATION ON HOW THAT'S DONE, THE MORE CONFIDENCE THE INDUSTRY AND PEOPLE HAVE IN WHAT YOU'RE DOING AND IT WILL AFFECT THESE BEHAVIORS WE'VE BEEN TALKING ABOUT. SO, I'M A BIG FAN OF IT.

**CARSWELL WHITEHEAD:** BUT AS IT RELATES AS FAR AS THE AUDITING PROCESS, THAT'S THE REASON WHY I SAID THAT AN UNANNOUNCED AUDIT IS MORE EFFECTIVE, BECAUSE YOU'RE CATCHING ANYONE WHO'S ATTEMPTING TO REMOVE TEST MATERIALS OR SHARE TEST MATERIALS WITH STUDENTS BEFORE THE TEST. YOU'RE ABLE TO DETECT THAT BY HAVING AN UNANNOUNCED AUDIT. WHAT YOU CAN ALSO HAVE IS AN AUDIT THAT WILL OCCUR AFTER THE TEST TO ENSURE THAT ALL OF THE MATERIALS HAVE BEEN RETRIEVED FROM THE TESTING STAFF AND ALL OF THE MATERIALS ARE ACCOUNTABLE, BECAUSE WHAT YOU'RE LOOKING AT NOW IS MAKING SURE THAT YOU HAVE ALL OF THE MATERIALS WILL THEN MITIGATE ANY TYPE OF RISK FOR A COMPROMISE AT A FUTURE ADMINISTRATION. AND I MENTIONED, TOO, THAT WHEN THESE TYPES OF IRREGULARITIES IN REGARDS TO TEST MATERIALS ARE DETECTED, YOU

MUST ACT VERY QUICKLY, BECAUSE YOU'RE NOT ONLY LOOKING AT IMPACTING THE LOCATION WHERE THE IRREGULARITY OCCURRED. YOU'RE LOOKING AT IMPACTING MANY SCHOOLS THAT WILL BE ADMINISTERING THAT TEST. TRAINING OTHER TESTING STAFF IS VERY KEY FOR THE TESTING STAFF TO BE VIGILANT. AND I ALSO THINK, TOO, AS FAR AS WITH STUDENTS, HAVING A WAY FOR STUDENTS TO BE ABLE TO SHARE IF THEY RECEIVED ANY TYPE OF PRE-KNOWLEDGE. IN MY OFFICE, WE HAVE A HOTLINE WHERE TEST TAKERS ARE ABLE TO CALL IN OR TEST TAKERS ARE ABLE TO SEND AN E-MAIL TO ADVISORS IF ANYONE RECEIVED AN UNFAIR ADVANTAGE. IF THEY HAVE THIS OPPORTUNITY TO SHARE THIS INFORMATION, THEY WILL BRING THIS INFORMATION TO OUR ATTENTION VERY EARLY IN THE PROCESS EVEN BEFORE SCORES ARE REPORTED, BECAUSE WHAT I'M HEARING IS THAT SOMETIMES YOU'RE NOT FINDING OUT ABOUT THESE IRREGULARITIES UNTIL SCORES ARE REPORTED. IF YOU HAVE A WAY FOR THIS INFORMATION TO BE CONVEYED TO YOU PRIOR TO SCORES BEING REPORTED, THEN YOU'RE ABLE TO PREVENT THOSE SCORES FROM BEING RELEASED.

**JACK BUCKLEY:** ON THAT TOPIC OF TIPLINES, DAVID, I THINK YOU RECOMMENDED THEM AS WELL IN THE K-12 ENVIRONMENT. IN A TYPICAL STATE TESTING PROGRAM, WHAT'S THE SORT OF VOLUME OF TIPS AND WHAT'S THE FALSE POSITIVE/FALSE NEGATIVE RATE LOOK LIKE, JUST OUT OF CURIOSITY, IF ANYONE HAS EXPERIENCED IT?

**MELISSA FINCHER:** I CAN'T TELL YOU ABOUT THE VOLUME. WE DEFINITELY GET IT. OFTENTIMES, IT COMES FROM PARENTS, FROM STUDENTS GOING HOME AND TALKING ABOUT THE TEST ENVIRONMENT AND TALKING, PERHAPS, ABOUT HOW "IT WASN'T AS HARD AS I THOUGHT IT WAS. IT WAS JUST EVERYTHING THE TEACHER COVERED THE DAY BEFORE" OR "THE TEACHER WAS VERY HELPFUL," THAT KIND OF THING. WE PROBABLY GET 20 TO 30 OF THOSE EACH ADMINISTRATION. WHAT WE'RE SEEING NOW IS--WHICH DOESN'T SOUND LIKE IT'S VERY MANY. WHAT WE'RE SEEING NOW IS THINGS COMING IN AFTER THE FACT, RATHER THAN DURING THE ADMINISTRATION. AND THAT'S A LITTLE BIT MORE DIFFICULT TO DEAL WITH.

**DAVID FOSTER:** AND I DON'T HAVE ANY COMPARATIVE EXPERIENCE LIKE MELISSA DOES, BUT I'D BE VERY CONCERNED ABOUT WHAT IS COMING IN. THERE MUST BE A LOT OF SOCIAL CONSTRAINTS TO NOT BLOW THE WHISTLE, NOT TO TIP, AND SO IT SHOULD BE ONE--AS OTHERS HAVE MENTIONED, TOO--IT SHOULD BE ONE OF MANY WAYS WE DO TO TRY TO DETECT THE KINDS OF THINGS THAT ARE GOING ON. I DON'T REALLY HAVE A FEELING FOR THE RELIABILITY OF TIPLINES AND HOW EFFECTIVE THEY REALLY ARE. MAYBE THAT'S A GOOD AREA OF RESEARCH, ESPECIALLY BECAUSE IT'S A MAJOR WAY OF DETECTION TODAY.

**CARSWELL WHITEHEAD:** WELL, IN MY WORLD, THE TIPLINE IS VERY EFFECTIVE. AND HOW ARE STUDENTS ABLE TO SHARE INFORMATION WITH THE SCHOOL IF THEY HAVE A CONCERN THAT THE TEST QUESTIONS THAT THEY SAW THAT

PARTICULAR DAY ARE TEST QUESTIONS THAT WERE REVIEWED WITH THEM THE DAY PRIOR?

**MELISSA FINCHER:** THEY WOULD GO TO THE PRINCIPAL, TYPICALLY, IF THAT'S THE CASE. IT DEPENDS ON THE LOCAL STRUCTURE. AS WE HEARD EARLIER TODAY AND AS YOU ALL PROBABLY KNOW, UNFORTUNATELY SOME OF THE EDUCATORS WHO REPORTED INAPPROPRIATE PRACTICE WERE PERSECUTED. SO, IT'S MAKING SURE THAT THE PROTECTIONS ARE THERE SO THE INFORMATION CAN BE FORTHCOMING.

**JACK BUCKLEY:** ALL RIGHT, SOMETHING ELSE THAT CERTAINLY EMERGED FOR ME AS A COMMON THEME ACROSS YOUR COMMENTS WAS THE NEED TO NOT FOCUS ON A SINGLE ANALYSIS. SO, I THINK YOU COULD CALL THIS SOME SORT OF DEFENSE IN DEPTH STRATEGY, SOMETHING WHERE WE NEED TO BE LOOKING AT A RANGE OF BOTH DIRECT AND INDIRECT DETECTION METHODS. GIVEN SORT OF REAL-LIFE SCARCE RESOURCE CONSTRAINTS, IF YOU'RE IMPLEMENTING AN ASSESSMENT PROGRAM, SAY, AT THE DISTRICT LEVEL, AMONG THIS DIFFERENT RANGE OF METHODS TO CHOOSE FROM, WHAT SHOULD BE THE FIRST FEW THAT YOU WOULD ABSOLUTELY RECOMMEND IN A K-12 SORT OF STANDARD ESEA-NCLB-TYPE TESTING ENVIRONMENT? WHAT ARE YOUR FIRST 3 GO-TO METHODS?

**CARSWELL WHITEHEAD:** I WILL LOOK AT THE ERASURES, IF WIDESPREAD, AND ALSO AN ANSWER PATTERN. LIKE I SAID, YOU CAN USE THOSE ANALYSES IN

CONJUNCTION. I HEARD SOMEONE MENTION EARLIER THAT IT MAY NOT BE UNUSUAL FOR A LARGE NUMBER OF ERASURES. YES, ON AN INDIVIDUAL BASIS, YES. BUT IF YOU SEE HUNDREDS OF ANSWER SHEETS THAT COME IN AND YOU SEE A LOT OF ERASURES--AND YOU WOULD HAVE TO HAVE A CRITERION. WHAT ARE YOU LOOKING FOR IN ERASURES? WHAT IS CONSIDERED UNUSUAL ERASURES? BECAUSE YOU HAVE TO HAVE STANDARDS IN ORDER TO USE. AND IF YOU SEE THAT YOU HAVE HUNDREDS OF ANSWER SHEETS WITH ERASURES, THAT DEFINITELY WOULD BE A FLAG. NOW, AS FAR AS MY PROCESS, AS FAR AS INDIVIDUAL CASES OR GROUP CASES, I WOULD LOOK AT PATTERN OF RESPONSES TO CHECK FOR EITHER PRE-KNOWLEDGE OR TO CHECK FOR SOME FORM OF COPYING OR COMMUNICATION. SO, THOSE ARE THE ONES THAT I WOULD LOOK AT. AND THEN IF YOU'RE LOOKING AT A COMPUTER-BASED TEST, IF STUDENTS ARE TAKING A TEST ON A COMPUTER AND YOU'RE SEEING THAT THEY'RE COMPLETING THE TEST UNUSUALLY FAST, THAT MIGHT BE AN INDICATOR THAT YOU WANT TO LOOK AT AS WELL.

**DAVID FOSTER:** I DON'T THINK I CAN ANSWER THE QUESTION. MAYBE IT'S MY LACK OF EXPERIENCE IN THIS AREA, BUT EACH OF THESE HAS THEIR ADVANTAGES AND THEIR DRAWBACKS. ERASURE ANALYSIS--AS THE WORLD IS GOING MORE TOWARD COMPUTERIZING EXAMS, IT'S A DATED METHODOLOGY. SO, ADVICE TO USE IT IS "USE IT FOR A WHILE. IT'S GONNA GO AWAY." AND REPLACE IT WITH OTHER STATISTICAL ANALYSES. I DON'T KNOW. AGAIN, EFFECTIVE SECURITY IS GOING TO BE ABLE TO CAPTURE--MAYBE YOU

COULD JUST USE EACH OF THEM ON A SMALLER SCALE, AND THAT MIGHT BE MORE EFFECTIVE. THAT'S PROBABLY THE WAY I WOULD ANSWER IT, I GUESS.

**BRIAN JACOB:** I WAS JUST THINKING... THE THING ABOUT COST AND COST BENEFIT. I THINK THERE'S PROBABLY VASTLY DIFFERENT COSTS FOR SOME OF THESE STRATEGIES. I WOULD SUGGEST THAT ERASURE ANALYSIS AND SOME OF THE STATISTICAL-BASED TEST SCORE ANALYSIS, I THINK, SHOULD BE VIRTUALLY COSTLESS ONCE THE SYSTEM IS PUT IN PLACE. I MEAN THERE'S A COMPUTER PROGRAM THAT SCANS THE TESTS THAT POPS OUT THE AVERAGE OR THE MEDIAN NUMBER OF WRONG TO RIGHT ERASURES. YOU NOW HAVE A PROGRAM THAT CALCULATES THE AVERAGE TEST SCORE GAIN FROM ONE YEAR TO THE NEXT. SO WE SHOULD BE DOING ALL OF THOSE THINGS THAT ARE CHEAP ALL OF THE TIME. AND THEN, FOR THINGS THAT ARE ACTUALLY EXPENSIVE, LIKE MANUAL INSPECTION, INVESTIGATION OF SPECIFIC IRREGULARITIES—THOSE THINGS, I THINK, WE WOULD HAVE TO RATION. I THINK DAVID'S POINT ABOUT FOCUSING ON HIGH-RISK CASES IS A GOOD CRITERIA OR LOOKING AT THE MEDICAL SCREENING MODEL. YOU DO THE LOW-COST THINGS FOR EVERYONE AND THEN YOU ONLY DO MORE IN-DEPTH INVESTIGATIONS WHEN THERE'S INDICATION OF IRREGULARITY EITHER FROM A TIP OR FROM ONE OF THESE LOW-COST SCREENS.

**JACK BUCKLEY:** BEFORE I OPEN IT UP, JUST ONE MORE QUICK POINT, BACK TO ONE OF THE THINGS YOU SAID AS A BEST PRACTICE, 'CAUSE IT'S SOMETHING THAT I'M NOT SURE THAT I'D HEARD ELSEWHERE. YOU DISCUSSED

ACTUALLY CONDITIONING. SO, IF YOU'RE PERFORMING SOME SORT OF STATISTICAL FORENSIC ANALYSIS TO DO THAT ANALYSIS CONDITIONAL ON SCHOOL AND/OR STUDENT CHARACTERISTICS, CAN YOU EXPLAIN THE RATIONALE FOR THAT AND WHY IF FOLKS AREN'T DOING THAT THAT THEY SHOULD BE, IN YOUR OPINION?

**BRIAN JACOB:** I THINK THE RATIONALE GETS BACK TO THE UNDERLYING CONCEPT BEHIND THE STATISTICAL ANALYSIS. YOU'RE LOOKING FOR A DEVIATION FROM WHAT IS EXPECTED OR NORMAL. AND THAT DEVIATION, YOU'RE TAKING AT LEAST AS AN INITIAL SUGGESTION OF IRREGULARITY, NOT NECESSARILY CHEATING, OF COURSE. AND...SO I GUESS THE QUESTION IS WHAT IS EXPECTED. AND IF THERE ARE CERTAIN CHARACTERISTICS OF STUDENTS OR SCHOOLS THAT ARE ASSOCIATED WITH ERASURES OR CERTAIN TEST SCORE PATTERNS OR TEST SCORE GAINS, THEN YOU WANT TO BE COMPARING TO THE EXPECTED FOR THAT GROUP AND NOT NECESSARILY FOR THE EXPECTED FOR ALL GROUPS. I THINK IN MANY CASES, IN MOST CASES, IF THE FLAGGING CRITERIA IS SUFFICIENTLY HIGH--I THINK, CERTAINLY, 3, 4, 5 STANDARD DEVIATIONS--AND THERE'S NOT A HUGE CORRELATION BETWEEN STUDENT OR SCHOOL CHARACTERISTICS AND THESE MEASURES, IN PRACTICE, IT PROBABLY WILL NOT HAVE A LARGE IMPACT IN MANY CASES. BUT I THINK IT'S CERTAINLY EASY ENOUGH TO DO AND CONCEPTUALLY SHOULD BE A COHERENT THING TO DO.

**JACK BUCKLEY:** TO OUR OTHER PANELISTS AND, OF COURSE, THE AUDIENCE.  
JOHN.

**JOHN FREMER:** I WANT TO PUT OUT A PERCEPTION AND SEE IF THERE'S AGREEMENT OR DISAGREEMENT ON IT. MY PERCEPTION IS THAT THERE IS NOT A DISAGREEMENT ABOUT WHETHER YOU SHOULD DO MULTIPLE FORENSICS ANALYSES FIRST DAY ASSESSMENT. I THINK THAT IS NOW THE STANDARD, THE BEST PRACTICE. THE ISSUE IS HOW TO GET TO THAT, GIVEN THAT YOU HAVE THESE ONGOING PROGRAMS, YOU HAVE THESE CONTRACTS WITH VENDORS, YOU HAVE THE CONSORTIUM TRAINS COMING DOWN THE TRACK. BUT FROM MY CONVERSATIONS WITH PEOPLE, FROM THE THINGS I READ, FROM THE CONFERENCES I GO TO, I VIEW THAT AS ALREADY SETTLED. YOU NEED TO BE DOING THAT AND IF YOU'RE NOT DOING THAT, YOU PROBABLY ALREADY KNOW IT AND YOU HAVE IT IN YOUR HEAD OR YOU'RE HAVING MEETINGS OR PLANS OF HOW TO GET TO THAT POINT.

**BRIAN JACOB:** MAYBE I COULD FOLLOW UP TO THAT AND DO A QUESTION. FOR STATE OR DISTRICT CONTRAST WITH TESTING COMPANIES AND VENDORS, HOW COMMON ARE THE PROVISION OF ERASURE STATISTICS TO THE DISTRICT OR STATE? I WOULD THINK, GIVEN THE EASE, THAT THAT SHOULD JUST HAVE ALWAYS BEEN AN AUTOMATIC SOMETHING THAT IS INCLUDED EVERY TIME JUST AS A MATTER OF COURSE. IS THAT THE CASE, OR IS IT VERY UNUSUAL OR ADDITIONALLY EXPENSIVE FOR SOME REASON FOR A STATE TO GET ERASURE

STATISTICS ALONG WITH THE OTHER TEST STATISTICS IT GETS BACK FROM THE VENDOR?

**STEVE FERRARA:** I CURRENTLY WORK FOR PEARSON, AND I WORKED FOR OTHER TESTING COMPANIES. WE ALL OFFER A SUITE OF FORENSIC ANALYSES. AND THE QUESTION IS WHETHER THE RFP CALLS FOR IT OR WHETHER THE STATE ASKS FOR IT IN AN ADDENDUM. SO THAT'S MY RESPONSE TO THAT. I'D ACTUALLY LIKE TO RAISE ANOTHER POINT, UNLESS YOU WANT TO CONTINUE THIS DISCUSSION.

**JACK BUCKLEY:** PLEASE.

**STEVE FERRARA:** ALL RIGHT, SO THERE'S BEEN SOME DISCUSSION OF ROBUST STATISTICS. AND JACK, I THINK YOU ASKED MELISSA A QUESTION THAT LED HER TO SAY GEORGIA USES 3 STANDARD DEVIATIONS IN THE WRONG TO RIGHT ANALYSIS. OTHER STATES USE 4. THAT'S CONSISTENT WITH WHAT I KNOW. AND I THINK IT BECOMES A MATTER OF POLICY AND PRACTICALITY IN MAKING DECISIONS ABOUT THE STANDARDS FOR FORENSIC ANALYSES AND OTHER DECISION-MAKING IN THE PROCESS. AND SOME PRACTICAL MATTERS CONCERN THINGS LIKE WHETHER YOU USE 3 OR 4 STANDARD DEVIATIONS-- YOU CAN MAKE A DECISION LIKE THAT USING HISTORICAL DATA, DEPENDING ON THE FALSE POSITIVE AND NEGATIVE RATES THAT YOU GET IN HISTORICAL DATA. ANOTHER CONSIDERATION IS HOW MANY INVESTIGATIONS YOU CAN AFFORD TO CONDUCT. THEY'RE EXPENSIVE. THEY'RE HARD TO DO. THEY'RE

POLITICALLY SENSITIVE, AND THEY GET IN THE PAPERS. SO, IF YOU USE 2 OR 3 STANDARD DEVIATIONS, YOU'RE GONNA FLAG A LOT OF CLASSROOMS, A LOT OF TESTING GROUPS. YOU CAN'T NOT FOLLOW UP ON ANY OF THOSE. YOU HAVE TO FOLLOW UP ON EVERY ONE. IF YOU USE 4 STANDARD DEVIATIONS, YOU FLAG CONSIDERABLY SMALLER NUMBERS OF TESTING GROUPS. IT GIVES YOU MORE RESOURCES TO FOCUS ON THE ONES THAT IN SUBSEQUENT INVESTIGATION YOU NEED TO GO INTO THE SCHOOLS AND DO MORE. AND YOU ALSO HAVE THE BENEFIT OF BEING ABLE TO EXPLAIN "WE'RE USING 4 STANDARD DEVIATIONS. THIS IS A REALLY STRICT CRITERION. IT FLAGS VERY FEW CLASSROOMS." IT GIVES YOU SOME POLITICAL CAPITAL TO MAKE YOUR JOB EASIER OF DETECTING AND INVESTIGATING.

**JACK BUCKLEY:** ANY COMMENT FROM THE PANEL?

**AUDIENCE MEMBER:** GOOD AFTERNOON. MY NAME IS DR. ... I'M ON THE D.C. STATE BOARD OF EDUCATION. MY CONCERN IS THE YOUNG PEOPLE THAT--

[MICROPHONE CUTS OUT]

...MY CONCERN IS THAT-- I'D LIKE TO STAND--IS THE TESTING. THE KIDS ARE ACTUALLY BEING DAMAGED BECAUSE OF THE FACT, AS ADULTS, WE HAVE A MORAL OBLIGATION TO MAKE SURE THEY HAVE THE BEST QUALITY EDUCATION AND MAKE SURE THE SERVICE THAT WE GIVE THEM ARE THE BES

SERVICES THAT ARE GIVEN TO THEM. WE'RE DAMAGING OUR KIDS BECAUSE THIS HAS BEEN GOING ON FOR YEARS. I'M IN THE SCHOOL SYSTEM FOR OVER 15 YEARS, ALONG WITH ANNETTE. AND ONCE YOU GIVE THESE TESTINGS, WHAT'S IN PLACE FO THEM TO HELP THEM TO SAY THAT THEY'RE DOING THE RIGHT THING? BECAUSE THEY'RE LISTENING TO US, AND THEY HAVE THEIR TRUST IN US TO SAY THAT WE'RE GIVING THEM THE BEST TESTING— CURRICULUM THAT WE'RE ASSIGNED TO GIVE THEM. WHAT'S IN PLACE TO HELP THEM? BECAUSE THEY'RE MENTALLY BEING DISTURBED AND ALSO THEY'VE BEEN OUTCAST BY THEIR FRIENDS BY CALLING THEM "CHEATERS" AND STUFF. IT'S ALMOST THE SAME AS BEING BULLIED. SO WE NEED TO HAVE SOMETHING IN PLACE TO MAKE SURE THAT THEY'RE DOING THE RIGHT THING, BECAUSE THEY PUT THEIR TRUST IN US TO SAY THAT WE'RE DOING THE RIGHT THING. SO, WHAT'S IN PLACE? I KNOW DAVID TALKED ABOUT GIVING ANOTHER TEST WHEN THOSE DIDN'T DO RIGHT TO SAY WHATEVER. THAT'S DEVASTATING WHEN YOU HAVE TO GO BACK AND FORTH AGAIN TO SEE HOW THAT TEST IS BEING DONE WHEN YO THINK YOU DID A GREAT JOB AND YOU GOTTA GO BACK AND TAKE THE TEST ALL OVER AGAIN. BECAUSE TO ME, IF THE PACKAGE WAS DONE IN THE CORRECT WAY IN THE BEGINNING, COMING OUT THE TESTING, WE WOULDN'T HAVE THIS ISSUE. A LOT OF TIMES WHEN YOU GET--JUST LIKE YOU TAKE YOUR DRIVERS' LICENSE. FIRST TIME YOU MIGHT FAIL, YOU HAVE TO GO BACK AND TAKE THE TEST AGAIN. AND THE SAME THING WHEN YOU GET INSTRUCTIONS IN GIFTS OR CARS AND AIRPLANES. IF THOSE EXAMPLES IN THERE IS NOT DONE RIGHT, THEN WHAT HAPPEN? YOU BE PUTTING ALL THE GIDGITS TOGETHER AND YOU STILL FIND

SOME DEFAULT OR OTHER INFORMATION THAT'S NOT CORRECT. WE NEED TO HAVE SOMETHING IN PLACE THAT OUR KIDS DON'T HAVE TO BE GOING THROUGH THIS AND ESPECIALLY THE PARENTS, 'CAUSE THIS STAYS WITH YOU FOR A LIFETIME. SO, WHAT'S IN PLACE-- WHAT I WANT TO KNOW FROM STATE BOARD OF EDUCATION AND FROM PARENTS THAT I'M GETTING FEEDBACK--THAT WHAT ARE WE GONNA DO TO MAKE THIS BETTER? THANK YOU.

**JACK BUCKLEY:** I THINK THERE'S A LOT IN THERE. IF I COULD TAKE ONE PIECE OUT, IT WOULD BE REALLY THE QUESTION ABOUT FOLLOW-UP. SO, LET'S SAY WE'VE SET A FAIRLY RIGOROUS THRESHOLD FOR DETECTING IRREGULARITY. ONCE WE'VE DETECTED IT--WE'RE GONNA TALK ABOUT INVESTIGATION ON THE NEXT PANEL. BUT IN TERMS OF WITHIN THE NARROWER CONTEXT, WHAT DO THE DETECTION PEOPLE DO NEXT AND HOW OFTEN DO THEY DO IT? WHERE DO WE SEND THIS TO AND WHERE DO WE GO FROM THERE IN THE IMMEDIATE NEXT STEPS?

**DAVID FOSTER:** I'M GONNA MAKE A VERY GENERAL COMMENT, AND I DON'T KNOW HOW HELPFUL IT WILL BE, BUT I THINK IN MANY OF THESE AREAS WE'RE STILL IN THE INFANCY. OUR PROCTORING METHODS ARE BROUGHT INTO QUESTION. OUR STATISTICAL METHODS ARE FAIRLY NEW, AND THERE AREN'T AS MANY USED AS WE WOULD LIKE. TIPLINES. WE DON'T REALLY KNOW HOW EFFECTIVE THEY ARE. WHAT WE NEED TO DO FROM THE DETECTION SIDE IS KEEP REFINING AND IMPROVING THEM. THERE'S A LOT OF ROOM FOR IMPROVEMENT, AND THE PROBLEM IS, SAY WITH PROCTORING, IS WE GET

SET. TEACHERS ARE GREAT AND THEY'RE ALWAYS GONNA PROCTOR WELL AND THAT'S A STANDARD AND WE CAN FORGET ABOUT IT NOW FOR 50 YEARS. AND THAT'S REALLY NOT THE CASE. WE NEED TO EVALUATE EACH OF THE THINGS THAT WE DO SO THAT WE CAN MAKE SURE WE'RE DETECTING PROPERLY WITH FEW FALSE POSITIVES AND FALSE NEGATIVES. I THINK THERE'S ROOM FOR IMPROVEMENT. THAT'S A VERY GENERAL ANSWER AND PROBABLY DOESN'T SPECIFICALLY ANSWER ANY INDIVIDUAL CONCERN BUT--

**MELISSA FINCHER:** I THINK IT'S SOMETHING THAT AT THE STATE DEPARTMENT WE DON'T TAKE LIGHTLY. IT IS VERY DISHEARTENING WHEN YOU HAVE TO INVALIDATE A STUDENT'S SCORE THROUGH NO CAUSE OF THE STUDENT HIM OR HERSELF, THROUGH ACTIONS THAT IMPROPER ADMINISTRATION OR SOME OTHER FORCE HAS OCCURRED. CERTAINLY, AS A STATE, WE'RE WORKING TO IDENTIFY THOSE STUDENTS WHO DID NOT HAVE AN ACCURATE ADMINISTRATION AND MAKING SURE THAT THEY RECEIVE SUPPORT SERVICES THAT THEY NEED SO THAT THEY CAN BE SUCCESSFUL, BUT IT'S VERY DIFFICULT.

**JACK BUCKLEY:** SCOTT.

**SCOTT NORTON:** I HAVE A QUICK QUESTION, REALLY. I THINK IT'S MOSTLY FOR MELISSA, BUT ANYBODY COULD RESPOND. IF YOU WOULD, SAY A WORD OR TWO ABOUT THE PROS AND CONS OF HAVING THIS OTHER AGENCY THAT YOU DEAL WITH THAT TAKES A LOT OF THE BURDEN FOR THE INVESTIGATION. I

DIDN'T QUITE GET ALL OF IT, BUT THAT'S PROBABLY ATYPICAL COMPARED TO MOST? JUST FOR OUR INFORMATION, HOW DOES THAT GO?

**MELISSA FINCHER:** SURE. THE GOVERNOR'S OFFICE OF STUDENT ACHIEVEMENT IS A SEPARATE AGENCY FROM THE GEORGIA DEPARTMENT OF EDUCATION, AND THEY ARE STATUTORILY RESPONSIBLE FOR EDUCATIONAL AUDITING, AND SO THEY ARE THE ONES THAT EMBARKED ON THIS PATH IN THE STATE OF GEORGIA. AND IN MANY WAYS, IT HAS WORKED VERY NICELY FOR US, BECAUSE MY DIVISION CAN CONTINUE TO FOCUS ON THE DEVELOPMENT OF THE TEST, ON THE ADMINISTRATION SUPPORTING DISTRICTS. WE WORK VERY COLLABORATIVELY WITH THEM IN THAT THEY CONTRACT WITH OUR VENDOR AND OFTEN THROUGH OUR OWN CONTRACT, BUT THE DATA GO DIRECTLY TO THEM AND NOT TO MY OFFICE. WE REVIEW THE FINDINGS OF THE INVESTIGATIONS AND BUMP THAT UP AGAINST OUR PROTOCOLS. I WORK WITH THE OFFICE OF STUDENT ACHIEVEMENT FOR ANY FINDINGS THAT THEY HAVE OCCURRED, AND THAT'S ONE OF THE TIPLINES THAT I DIDN'T TALK ABOUT IS THAT OFFICE ALSO RECEIVES INQUIRIES AND TIPS FROM THE FIELD THAT THEY FOLLOW ON THAT I'M NOT PRIVY TO. BUT WE'RE CONSTANTLY--IT'S A RECURSIVE PROCESS. IN LOOKING AT OUR SECURITY PROTOCOLS AND, UM... MONITORING THEM AND REVISING THEM, WHAT WE HAVE FOUND AS A RESULT OF THE LATEST INVESTIGATION IS THAT THEY WERE USING SOME OF OUR PROTOCOLS AS AN EXCUSE TO DO SOME OF THE THINGS THEY WERE DOING. SO WE HAD TO ADJUST BASED ON THAT. BUT IT'S MADE MY LIFE, QUITE FRANKLY, SCOTT, A LOT EASIER HAVING THAT OUTSIDE AGENCY LEAD THAT

WORK BECAUSE EVEN WITH IT, IT'S BECOME A SECOND FULL-TIME JOB FOR ME.

**AUDIENCE MEMBER:** LET ME JUST SAY THIS. MY NAME IS STEVE JACKSON. I'M THE PRINCIPAL OF DUNBAR HIGH SCHOOL, AND I WAS ALSO THE PRINCIPAL OF TWO DIFFERENT HIGH SCHOOLS IN NEW YORK STATE, AND I'VE BEEN DOING THIS WORK FOR OVER 25 YEARS. NUMBER ONE, I THINK IT'S SAD THAT WE'RE HAVING THIS TYPE OF CONVERSATION INVOLVING INTEGRITY SIMPLY BECAUSE AS AN EDUCATOR, ALL OF US SHOULD HAVE SOME LEVEL OF INTEGRITY, ESPECIALLY AS IT PERTAINS TO TESTING. WE TALKED A LITTLE BIT ABOUT TRANSPARENCY A LITTLE EARLIER. I THINK IT'S IMPORTANT TO HAVE A LEVEL OF TRANSPARENCY, BUT MORE IMPORTANTLY I DO THINK THAT ADMINISTRATORS, DISTRICTS, AND OTHERS NEED TO BE TRAINED SPECIFICALLY ON TESTING INTEGRITY, BECAUSE IT'S CLEAR TO ME THAT THROUGHOUT THE COUNTRY, WE'RE HAVING A LOT OF PROBLEMS IN DIFFERENT DISTRICTS AND DIFFERENT SCHOOLS IN TERMS OF TESTING INTEGRITY AS WELL AS THE RESPONSIBILITY OF EDUCATORS. SO, FOR ME, THAT'S EXTREMELY IMPORTANT, BUT EVEN AS WE LOOK AT THE HIGH-STAKE TESTING AND AS WE MOVE TOWARD THE COMMON CORE, THERE ARE TEACHERS NOW THAT ARE BEING HELD RESPONSIBLE FOR WHETHER OR NOT THEIR PARTICULAR STUDENTS PASS OR FAIL. AND AS WE MOVE AROUND THE COUNTRY AND THEY'RE HELD RESPONSIBLE FOR WHETHER OR NOT THEIR STUDENTS PASS OR THEIR STUDENTS FAIL, THEN SOME OF THESE PEOPLE WHO MAY NOT HAVE INTEGRITY WILL CHEAT. SO I THINK IT'S IMPORTANT FOR US TO BEGIN TO

HAVE THESE CONVERSATIONS IN DISTRICTS THROUGHOUT THE COUNTRY SO THAT PEOPLE ARE REAL CLEAR NOT ONLY WHAT THEIR ROLE AND RESPONSIBILITY IS BUT ALSO WHAT SOME OF THE CONSEQUENCES MAY BE FOR CHEATING.

**JACK BUCKLEY:** AND I THINK THAT SUMMARIZES MUCH OF THE INTENT OF THIS ENTIRE DAY, SO WOULD AGREE. YES.

**AUDIENCE MEMBER:** EXCUSE ME.

[MICROPHONE CRACKLING]

I'D LIKE, UM--

[MICROPHONE FEEDBACK]

I JUST WOULD LIKE A REACTION FROM THE PANEL MEMBERS ABOUT THIS IDEA. THIS IS BASED ON THE COMMENT THAT WAS MADE BY DR. FOSTER THAT THERE ARE HUNDREDS, MAYBE THOUSANDS OF WAYS THAT THERE'S CHEATING AND STEALING AND ALSO DR. CIZEK'S COMMENT ON SHARED RESPONSIBILITY. ISN'T THE TIME THAT THERE IS SOME KIND OF EITHER A LIST OR A DATABASE OR SOME KIND OF CLEARINGHOUSE THAT IF THERE'S SOME NEW WAY OF CHEATING OR STEALING THAT EVERYONE'S MADE AWARE OF IT AND MAYBE WITH A CORRESPONDING WAY OF DEALING WITH IT? WHETHER

IT'S A NEW WAY OF PROCTORING, A NEW STATISTICAL INDEX, OR MAYBE EVEN NEW COMPUTER PROGRAM THAT MIGHT EVEN BE SHARED AS AN OPEN SOURCE. AND YES, AS MEMBERS OF NCME, WE DO PRESENT OUR RESEARCH ABOUT THESE THINGS ONCE A YEAR, BUT SOMETIMES THE NEED IS MORE IMMEDIATE. AND I'M THINKING ALONG THE SAME LINE AS FOR IN THE I.T. INDUSTRY WHEN THERE ARE NEW VIRUSES. NORTON KEEPS TRACK OF IT, AND THEN THEY HAVE A SOLUTION AND THEY UPDATE YOUR COMPUTERS WITH IT. SO I JUST LIKE A REACTION FROM THE PANEL MEMBERS. THANK YOU.

**CARSWELL WHITEHEAD:** AS FAR AS TEST CENTER, I CAN SPEAK IN REGARDS TO TEST CENTER SUPERVISORS BECAUSE THAT'S WHERE I'M PRIMARILY IN CONTACT WITH. WE DO SHARE THAT TYPE OF INFORMATION WITH THE SUPERVISORS DURING TRAINING SESSIONS SO THEY CAN BE VIGILANT DURING THE TEST AND BE ABLE TO DETECT THOSE TYPES OF IRREGULARITIES. SO IT'S AN ONGOING PROCESS WHERE THAT INFORMATION IS SHARED EITHER VIA ONLINE TRAINING OR IN PERSON WE SHARE THAT INFORMATION.

**DAVID FOSTER:** YOU KNOW, I AGREE THAT IT'S NOT VERY SYSTEMATIC. WHEN WE STARTED CAVEON 10 YEARS AGO, WE STARTED AT CHEATING IN THE NEWS, A NEWSLETTER, AND INTENDED TO SHOWCASE EVERY COUPLE OF WEEKS 5 OR 6 INCIDENTS THAT CAME OUT OF THE NEWS. EVERYBODY ENJOYED READING IT. YOU DO LEARN SOME THINGS, BUT IT ISN'T SYSTEMATIC LIKE YOU'RE SUGGESTING. IT'S NOT A DATABASE, AND I THINK SOMETHING LIKE THAT IS

NEEDED. HOW WE GATHER THAT INFORMATION AND PRESENT IT IN A REASONABLE WAY, I'M NOT REALLY SURE. BUT I CERTAINLY AGREE AS TO THE NEED FOR THAT.

**JACK BUCKLEY:** I THINK WE HAVE ONE QUESTION FROM THE WEB AUDIENCE.

**JAMES ELIAS:** WE DO, ACTUALLY. IT'S FROM JIM WOLLACK AT THE UNIVERSITY OF WISCONSIN. JIM WANTS TO KNOW ABOUT DEVELOPING A CRITERION BASED ON LIKELIHOOD OF COMMITTING A FALSE POSITIVE AND THEN USING STATISTICAL CORRECTIONS LIKE BONFERRONI TO ADJUST FOR THE TOTAL NUMBER OF ANALYSES BEING PERFORMED. HE WAS OFFERING THAT AS AN ALTERNATIVE TO USING 3 VERSUS 4 STANDARD DEVIATIONS.

**JACK BUCKLEY:** I'M HAPPY TO SPEAK TO THAT, BUT IF ANYONE ELSE ON THE PANEL WOULD RATHER... I THINK THIS IS SOMETHING WE'VE CERTAINLY SEEN IN MANY OF THE RFI COMMENTS THE DEPARTMENT HAS ALREADY RECEIVED. AT LEAST IN SEVERAL PLACES, THERE'S A QUESTION ABOUT PEOPLE UNDERSTANDING THAT IF YOU'RE DATA DREDGING OR NOT LOOKING TO TEST A PARTICULAR HYPOTHESES THAT YOU NEED TO CONSIDER VERY CAREFULLY THE FALSE POSITIVE AND FALSE NEGATIVE RATES. AND BRIAN, YOU TOUCHED ON THIS BEFORE. YOU CAN DO THAT BY SETTING YOUR THRESHOLDS VERY FAR UP, BUT YOU CAN ALSO DO THAT BY CONSIDERING THE ENTIRE SORT OF FAMILY-WISE ERROR RATE OR WHATEVER IT IS ACROSS YOUR DIFFERENT ANALYSES. I THINK YOU CAN SHOW THERE'S A VARIETY OF

METHODS THERE, AND SOME OF THEM ARE EQUIVALENT STATISTICALLY, AND SOME OF THEM SORT OF SERVE DIFFERENT PURPOSES. BUT I DON'T THINK THAT THERE'S ANY CLEAR STANDARDS YET THAT ANYONE HAS AGREED ON, WHICH IS SOMETHING I'M CERTAINLY HEARING TODAY. I DON'T KNOW IF ANYONE ELSE HAS A COMMENT ON THAT.

QUICK QUESTION? I GOT TIME FOR ONE MORE, THEN WE'VE GOTTA WRAP UP.

[MICROPHONE FEEDBACK]

**AUDIENCE MEMBER:** I'M CATE SWINBURN. I'M THE CHIEF OF DATA AND ACCOUNTABILITY IN D.C. PUBLIC SCHOOLS. NOT SURE IF MY MICROPHONE [INDISTINCT].

I WONDER IF ANY OF THE ACADEMICS WHO'VE DONE THE STATISTICAL ANALYSIS HAVE ALSO LOOKED AT TEACHER EVALUATION SCORES OR THIRD-PARTY EVALUATIONS, SO AS WE ARE STEPPING INTO THIS IN D.C. PUBLIC SCHOOLS THIS YEAR FOR SOME OF OUR ANALYSES, I'M LOOKING AT THOSE FLAGGED TEACHERS TO LOOK AT THEIR OBSERVATION SCORES ON OUR EVALUATION, WHICH FOR THOSE OF YOU WHO DON'T KNOW, FOR D.C. PUBLIC SCHOOLS, WE'VE GOT—TWO OF THE TEACHERS' OBSERVATIONS A YEAR ARE DONE BY A THIRD-PARTY EVALUATOR, SO IT'S NOT JUST A PRINCIPAL'S EVALUATION SCORE. ADDITIONALLY, OUR TEACHERS ARE EVALUATED ON CORE PROFESSIONALISM AS PART OF THEIR ANNUAL SCORE. SO, CERTAINLY, FOR

THIS YEAR, WE ARE INVESTIGATING EVERYONE. BUT THERE'S A PART OF ME THAT WONDERS IF OVER MULTIPLE YEARS IF WE CAN BE CORRELATING SORT OF DETERMINATIONS OF IMPROPRIETY ALONG WITH VERY HIGH OBSERVATION SCORES OR IF THERE'S THE ABILITY TO SAY IF SOMEBODY'S NOT HITTING FOR WRONG TO RIGHT BUT THEY'RE HITTING FOR BEING A HIGH FLYER ON GROWTH, BUT THEY ARE ALSO VERY HIGHLY RATED FOR OBSERVATIONS AND CORE PROFESSIONALISM, COULD WE EXCLUDE THEM? NOT THAT WE ARE NOW, BUT WONDERING IF ANYBODY'S ACTUALLY DONE ANY LOOK AT CORRELATING TEACHER PERFORMANCE EVALUATIONS TO IMPROPRIETIES AS PART OF MAKING THIS MORE ROBUST.

**JACK BUCKLEY:** MY GUESS IS NO, BUT IF ANYONE ELSE DOES KNOW AN EXAMPLE... I THINK THIS FALLS IN THE VERY BROAD CATEGORY OF USING ALL THE INFORMATION YOU CAN, AND I THINK THAT'S A GOOD IDEA.

**AUDREY AMREIN-BEARDSLEY:** I CAN ANSWER THAT A LITTLE BIT. I KNOW THAT THE CORRELATION IN THE OFFICES THAT ARE JUST STARTING TO BE CONDUCTED, PARTICULARLY ON WHAT YOU ALL ARE DOING IN OTHER STATES ON THE GROWTH AND THE VALUE-ADDED IS ONLY ON THE GROWTH AS CORRELATED WITH OBSERVATIONS RIGHT NOW. I'M NOT AWARE OF ANYTHING THAT'S LOOKED AT IMPROPRIETARIES. IT'S JUST WE'RE AT THE FIRST-LEVEL ANALYSES AT THIS POINT IN RUNNING THOSE CORRELATIONS.

**DAVID FOSTER:** AND I'LL ANSWER IT IN A DIFFERENT WAY. I AGREE THAT THERE'S A LOT MORE DATA THAT SHOULD BE GATHERED IN ON THESE ANALYSES. A RECENT ONE I'M AWARE OF--A PROBLEM WAS DISCOVERED THROUGH A FORENSIC ANALYSIS SIMPLY BECAUSE TWO TEST TAKERS SHARED THE SAME E-MAIL ADDRESS. IT'S REALLY A VERY SPECIFIC PIECE OF INFORMATION, BUT IT LED TO AN INVESTIGATION AND DISCOVERY THAT THERE WAS A PROXY TEST TAKING RING GOING ON, AND SO I BRING THAT UP JUST TO ILLUSTRATE THAT THERE'S A LOT OF DATA THAT CAN BE ASSOCIATED WITH THE TESTS THAT WE GENERALLY DON'T USE. AND THAT'S AN EXAMPLE OF ONE THAT HAPPENED TO CAUSE A HIT AT THAT TIME, BUT CERTAINLY WE SHOULD DO MORE OF THAT WORK.

**JACK BUCKLEY:** ALL RIGHT, SO, TO WRAP UP, CERTAINLY WE'VE HEARD FROM THIS VERY INTERESTING PANEL THAT WE NEED TO USE A RANGE OF STATISTICAL AND NON-STATISTICAL METHODS FOR DETECTION AND VERIFICATION. CERTAINLY NO ONE METHOD IS DEFINITIVE, AND WE NEED TO DETECT EARLY AND USE EARLY DETECTION METHODS TO EVALUATE SECURITY PROCEDURES AS WELL. THIS IS NOT JUST A POST-OP STATISTICAL PROCESS. CERTAINLY, I THINK SOMETHING WE'VE HEARD IN THE WRITTEN COMMENTS AND ALSO TODAY IS SOMETHING YOU'LL HEAR ALL THE TIME IN SECURITY THAT IT'S A PROCESS, NOT A PRODUCT. THERE'S ALWAYS NEW METHODS OF CHEATING DEVELOPED. THERE ARE ALWAYS SORT OF A CONSTANT RACE BETWEEN PEOPLE TRYING TO GET AWAY WITH SOMETHING AND PEOPLE TRYING TO DETECT THEM, AND THIS IS CERTAINLY NO EXCEPTION. I THINK THE

OTHER PIECE THAT I'M HEARING THAT'S VERY INTERESTING COMING OUT OF THIS PANEL IS, AGAIN, THE NEED FOR TRANSPARENCY AND CLEAR CRITERIA. WE DIDN'T GET A CHANCE TO TALK TO MAYBE ABOUT POLICY-WISE--STEVE, YOU WERE STARTING TO GET THERE-- BUT WHO SHOULD SET THOSE CRITERIAS. IS THAT A STATE BOARD? IS THERE A TECHNICAL ADVISORY PANEL? AND THAT'S SOMETHING I'D LIKE TO HEAR MORE ABOUT IF WE HAVE THE TIME, BUT CLEARLY WE NEED TO COMMUNICATE THESE THRESHOLDS TO THE AUDIENCE. SO, ON THAT NOTE, I'D LIKE TO THANK THIS PANEL FOR ANOTHER VERY INTERESTING SESSION.

[APPLAUSE]

WE'RE NOW GONNA SHIFT TO LUNCH ON YOUR OWN. SO, YOU GOT AN HOUR TO EAT LUNCH, AND WE WILL START PROMPTLY AT 1:30, BECAUSE WE HAVE TWO MORE VERY GOOD SESSIONS TO GET THROUGH THIS AFTERNOON.

THANK YOU VERY MUCH.

PANEL III: RESPONSE AND INVESTIGATION OF ALLEGED AND/OR ACTUAL  
IRREGULARITIES IN ACADEMIC TESTING

**JACK BUCKLEY:** WELCOME BACK, EVERYONE.

OUR THIRD PANEL OF THE DAY FOCUSES NOW ON RESPONSE TO AND  
INVESTIGATION OF ALLEGED AND/OR ACTUAL IRREGULARITIES IN ACADEMIC  
TESTING.

WE'LL START BY HEARING FROM BOB WILSON OF WILSON, MORTON, AND DOWNS  
AND A FORMER DISTRICT ATTORNEY FOR DEKALB COUNTY, GEORGIA, OF  
COURSE. BOB IS ONE OF THE SPECIAL INVESTIGATORS APPOINTED BY FORMER  
GOVERNOR SONNY PERDUE AND GEORGIA GOVERNOR NATHAN DEAL TO  
INVESTIGATE CHEATING ON THE CRCT IN THE ATLANTA PUBLIC SCHOOLS  
AND DOUGHERTY COUNTY SCHOOL SYSTEM.

NEXT, WE'LL BE JOINED BY TISHA EDWARDS, WHO'S SPEAKING ON BEHALF OF  
DR. ANDRES ALONSO, CEO OF BALTIMORE CITY PUBLIC SCHOOLS.

TISHA IS THE CHIEF OF STAFF OF BALTIMORE CITY, WHERE SHE  
COORDINATES THE ACTIVITIES OF THE SYSTEM OFFICERS FOR THE  
EFFICIENT OPERATION OF ALL SCHOOLS AND OFFICES. SHE'S ALSO THE  
FOUNDER OF BALTIMORE FREEDOM ACADEMY, A CHARTER HIGH SCHOOL IN THE  
CITY.

NEXT, WE'LL HEAR FROM STEVE FERRARA, WHO'S VICE PRESIDENT AND CO-DIRECTOR OF THE CENTER FOR PERFORMANCE ASSESSMENT AT PEARSON.

PRIOR TO JOINING PEARSON, STEVE WAS A PRINCIPAL RESEARCH SCIENTIST AT CTB/McGRAW-HILL, MANAGING RESEARCH DIRECTOR IN THE ASSESSMENT PROGRAM AT THE AMERICAN INSTITUTES FOR RESEARCH, AND STATE ASSESSMENT DIRECTOR IN MARYLAND.

AND FINALLY, LOU FABRIZIO IS THE DIRECTOR OF THE DIVISION OF DATA, RESEARCH, AND FEDERAL POLICY AT THE NORTH CAROLINA DEPARTMENT OF PUBLIC INSTRUCTION. HE CONTINUES TO SERVE AS THE STATE'S FEDERAL LIAISON WITH THE U.S. DEPARTMENT OF EDUCATION AND ALSO SERVES AS THE SENIOR ADVISOR TO STAFF IN THE ACCOUNTABILITY SERVICES DIVISION REGARDING THE STATE'S ABCs ACCOUNTABILITY PROGRAM, THE STATEWIDE TESTING PROGRAM, AND THE STATE'S ACCOUNTABILITY AND ASSESSMENT PLANS.

SO, BOB?

**BOB WILSON:** GOOD AFTERNOON.

IT'S ALWAYS GREAT TO BE FIRST AFTER LUNCH WHEN EVERYBODY WANTS TO TAKE A NAP. THE ATLANTA INVESTIGATION, WHICH I WAS HONORED ALONG

WITH A GENTLEMAN NAMED RICHARD HYDE AND MIKE BOWERS TO HEAD UP, WAS A MASSIVE INVESTIGATION, AND UNFORTUNATELY, IT REVEALED AN UGLY AND DISTURBING TRUTH.

TO GIVE YOU SOME CONCEPT OF THE SIZE, THE INVESTIGATIVE TEAM INCLUDED OVER 60 STATE AGENTS, ATTORNEYS, PARALEGALS. IN THE END, WE ACTUALLY USED OVER 100 DIFFERENT PEOPLE IN THAT INVESTIGATION. THAT INVESTIGATION, BY ANY STATISTICAL ASSESSMENT, WOULD NORMALLY HAVE TAKEN 14 TO 18 MONTHS TO 18 MONTHS, AND WE PRESSED IT INTO 10 MONTHS. SO IT WAS VERY EXHAUSTIVE FOR EVERYONE INVOLVED IN IT.

WE INTERVIEWED APPROXIMATELY 2,000 PEOPLE AND A TOTAL OF 2,200 INTERVIEWS. THAT'S A LOT OF INTERVIEWS.

WE WENT THROUGH OVER 800,000 DOCUMENTS.

THERE WERE 56 SCHOOLS THAT WERE INVOLVED IN THE INVESTIGATION, AND WE FOUND CHEATING AT 44 OF THEM. IN 30 OF THE SCHOOLS, 82 PRINCIPALS AND TEACHERS CONFESSED.

OVER 180 EDUCATORS WERE INVOLVED ULTIMATELY IN THE END BY NAME.

IT'S IMPORTANT TO NOTE THAT IN OUR REPORT, BASED ON ALL OF THE EVIDENCE THAT WE FOUND, WE ESTIMATED THAT THE NUMBER OF TEACHERS

AND ADMINISTRATORS INVOLVED WAS SOMETHING LIKE TWO TO THREE TIMES GREATER THAN THOSE WE COULD NAME.

THERE WERE 38 PRINCIPALS INVOLVED. IF YOU'RE IN THE SYSTEM OUT THERE IN AMERICA, YOU DON'T WANT TO BE WHERE ATLANTA WAS.

THE HEADLINES WERE AWFUL. YOU DON'T WANT TO READ HEADLINES LIKE, "WHISTLE BLOWERS GET PUNISHED," "A.P.S. SUPPRESSED SCANDAL," "ATLANTA SCHOOLS SOFT ON CHEATERS," "UNETHICAL BEHAVIOR ACROSS EVERY LEVEL"--THAT'S A HEADLINE--WITH PICTURES OF TOP ADMINISTRATORS.

IT HURTS THE SYSTEM. IT HURTS THE IMAGE OF EDUCATION IN THIS COUNTRY.

WHEN YOU LOOK AT THE FINAL POINT ON THIS PARTICULAR SLIDE, YOU SEE THAT IT'S TIME-CONSUMING, AND IT'S CERTAINLY OFF MISSION. IT ABSOLUTELY WILL DIVERT YOUR SYSTEM FROM WHERE IT OUGHT TO BE TO WHERE IT SHOULD NOT BE. IN THE ATLANTA SYSTEM...THEY WERE NOT HAPPY TO WORK WITH NO CHILD LEFT BEHIND. THEY WANTED TO CREATE TARGETS THAT WERE FAR ADVANCED BEYOND WHAT NO CHILD LEFT BEHIND REQUIRED.

THEY CREATED THEM, AND THE CUMULATIVE EFFECT WAS DISASTROUS AND DESTRUCTIVE FOR THEM. THEY NOT ONLY CREATED TARGETS THAT EXCEEDED

NO CHILD LEFT BEHIND, BUT THEY IMPOSED PRESSURES TO ACHIEVE THOSE TARGETS, AND THOSE PRESSURES WERE BOTH NEGATIVE AND POSITIVE.

THE NEGATIVES WERE, YOU WERE GOING TO BE PUBLICLY HUMILIATED IF YOU DIDN'T MEET TARGETS. IF YOUR SCHOOL DIDN'T MAKE IT, YOU, AS THE PRINCIPAL WERE GOING TO BE ON THE HOOK. IF YOU, AS A TEACHER, DIDN'T MAKE IT, THEN YOU WERE GOING TO BE ON THE HOOK, AND YOU WERE GOING TO BE PUBLICLY HUMILIATED IN FRONT OF OTHER PRINCIPALS OR OTHER TEACHERS, AND YOUR JOB WAS IN DANGER.

IF YOU MADE IT, WELL, YOU GOT RECOGNITION. YOU GOT BONUSES. YOU GOT PRAISE. YOU GOT JOB SECURITY. THE END RESULT WAS ABSOLUTELY A DISASTER.

THE SUPERINTENDENT SAID THE FOLLOWING. SHE QUOTED HERSELF AS TELLING HER PRINCIPALS WHEN SHE HIRED ONE, "YOU HAVE 3 YEARS TO MEET TARGETS, OR I WILL FIND SOMEONE WHO WILL. THERE WILL BE NO EXCEPTIONS AND NO EXCUSES." NOT ONLY DID SHE SAY THAT IN ONE INTERVIEW. SHE SAID IT IN YET A SECOND INTERVIEW.

WHAT MESSAGE DID THAT DELIVER?

IF I REQUIRE MY LAWYERS IN MY FIRM TO BILL A CERTAIN NUMBER OF HOURS A YEAR THAT IS OUTRAGEOUS AND TOO MUCH, THEY ARE EITHER GOING TO QUIT OR THEY ARE GOING TO LIE AND CHEAT.

WE MUST SET REASONABLE GOALS, NOT UNREASONABLE GOALS.

WE HAVE HEARD ABOUT SOME OF THESE TARGETS AND WARNINGS, SIGNIFICANT INCREASES IN SCORES OR SIGNIFICANT DECREASES IN SCORES, AND PROBABLY THE BETTER WORD IS UNUSUAL INCREASES OR DECREASES.

SOME PEOPLE NEVER THINK ABOUT THE DECREASES, BUT THE DECREASES THIS YEAR MIGHT TELL YOU SOMETHING ABOUT LAST YEAR--UNUSUAL ANSWER PATTERNS, HIGH WRONG-TO-RIGHT ERASURES, REPORT OF SUSPICIOUS CONDUCT, REPORT OF CHEATING.

THOSE ARE DIFFERENT THINGS.

PEOPLE TELL AND REPORT SOMETHING THAT THEY SAW THAT JUST DIDN'T SEEM RIGHT. PEOPLE SAW THINGS AND THOUGHT IT WAS CHEATING. YOU HAVE TO PAY ATTENTION TO THOSE THINGS.

GREG CIZEK AND I WERE TALKING DURING LUNCH. THE HIGH WRONG-TO-RIGHT ERASURES THAT WERE MENTIONED ARE SELDOM, IF EVER, SKEWED BY A SINGLE STUDENT TO THE EXTENT THAT IT WOULD CAUSE YOU TO THINK THAT

CHEATING HAD HAPPENED IN THAT SCHOOL OR EVEN THAT CLASSROOM.

WE LOOKED AT THOUSANDS OF TESTS. WE HAD STUDENT-LEVEL DATA. WE COULD GO TO EACH STUDENT'S TEST AND HOW MANY WRONG-TO-RIGHT ERASURES THAT PARTICULAR STUDENT HAD ON EACH OF THE TESTS.

WE NEVER FOUND ONE WHERE SOME STUDENT BY THEMSELVES BY GETTING OFF TRACK AND CORRECTING THEIR SITUATION JUST THREW THE WHOLE NUMBERS FOR THAT CLASS OFF SO BADLY THAT IT WOULD LEAD YOU TO BELIEVE THERE WAS JUST CHEATING THERE.

YOU CAN GET THE STUDENT-LEVEL DATA. THE IDEA THAT WE WON'T BE ABLE TO FOLLOW THIS IF WE GO TO COMPUTERIZED TESTS IS HOGWASH. COMPUTERS CAN DO ABOUT ANYTHING. THEY CAN TELL YOU WHEN PLUG IN AN ANSWER AND WHEN YOU CHANGED THE ANSWER AND HOW MANY TIMES YOU CHANGED IT. YOU WILL JUST HAVE A NEW FORM OF ERASURE, IS WHAT YOU WILL HAVE.

WHAT DO YOU DO WHEN YOU GET THESE REPORTS?

WELL, YOU DON'T WANT TO BE A PENN STATE. YOU REPORT IT, AND YOU KEEP REPORTING IT, AND YOU REPORT IT UP THE LINE: ONE, TO SCHOOL LEADERSHIP--THE PRINCIPAL, THE ASSISTANT PRINCIPAL, THE TEST COORDINATOR.

IF YOU ARE A STAFF MEMBER OR A TEACHER, YOU REPORT IT, AND YOU REPORT IT UP THE LINE UNTIL SOMEBODY PAYS ATTENTION.

THE SCHOOL HAS TO REPORT IT TO THE SYSTEM LEADERSHIP AND TAKE IT UP, ASSISTANT SUPERINTENDENTS AND SUPERINTENDENTS OF THE SYSTEM.

THE SYSTEM NEEDS TO LET THE LEGAL COUNSEL FOR THE SYSTEM KNOW WHAT THE PROBLEM MAY BE. THE LEGAL COUNSEL NEEDS TO BE INVOLVED.

YOU GENERALLY ARE GOING TO WANT TO NOTIFY THE STATE EDUCATION AUTHORITIES IN YOUR STATE. IT MAY NOT BE BECAUSE THERE IS ANYTHING MORE THAN YOU WANT TO LET THEM KNOW WHAT YOU ARE LOOKING INTO, AND DO THEY HAVE ANY ADVICE FOR COUNSEL FOR YOU?

YOU MAY WANT TO LET YOUR STATE LICENSING AUTHORITY KNOW, NOT BECAUSE YOU HAVE ANYTHING OTHER THAN TO LET THEM KNOW THAT THEY MAY, TOO, HAVE SOME ADVICE FOR YOU AS TO HOW YOU MIGHT GO ABOUT IT.

YOU MAY NEED TO LET YOUR SCHOOL RESOURCE OFFICERS OR LAW ENFORCEMENT KNOW. IT DEPENDS ON WHO YOU ARE GOING TO USE FOR YOUR INVESTIGATION, BUT YOU NEVER EVER GO WRONG BY REPORTING IT UP THE CHAIN, AND IF THE PERSON AHEAD OF YOU DOES NOT DO ANYTHING, GO ABOVE THEM AND BEYOND THEM.

KEY COMPONENTS.

NEVER EVER IGNORE ANY COMPLAINT. ALL REPORTINGS ARE IMPORTANT. THEY MUST BE VIGOROUSLY AND THOROUGHLY INVESTIGATED, AND IF IT IS SOMETHING PRETTY BIG, YOU ARE GOING TO NEED SOME SERIOUS ORGANIZATION AND PLANNING BEFORE YOU DO IT.

EVERY WITNESS MUST BE INTERVIEWED, EVERYONE WHO MAY KNOW SOMETHING, IF IT IS POSSIBLE THEY KNOW SOMETHING, AND IT IS NOT ABOUT WHAT THEY KNOW DIRECTLY.

IT IS ABOUT WHAT THEY KNOW IN ANY REGARDS.

WE FOUND THAT THE PRELIMINARY INVESTIGATION DONE BY ATLANTA, WE TELL THE PEOPLE, "WELL, WE JUST ONLY WANT TO KNOW WHAT YOU KNOW."

"WELL, I HEARD--"

"WELL, I DON'T KNOW WHAT YOU HEARD."

YOU WANT HEARSAY. YOU WANT GOSSIP. YOU WANT INFORMATION FROM WHATEVER SOURCE IF YOU'RE ACTUALLY INVESTIGATING IT. WHY? BECAUSE SOMEONE UNRELATED TO THE WITNESS YOU ARE TALKING TO MAY VERY WELL HAVE TOLD YOU THAT SAME KIND OF INFORMATION.

THESE PEOPLE HAVE NO CONNECTION. IT TELLS YOU THERE IS SMOKE OUT THERE.

YOU NEED TO FIGURE OUT WHERE IT IS COMING FROM.

SCHOOL PEOPLE ARE NOT REALLY WELL-EQUIPPED TO BE INVESTIGATORS. THAT'S NOT WHAT YOU ARE TRYING TO DO. YOU ARE TRYING TO EDUCATE. THAT'S WHY YOU GOT TO USE RESOURCES OF ATTORNEYS, INVESTIGATORS, LAW ENFORCEMENT WHERE APPROPRIATE, BUT EVERYBODY NEEDS TO BE INTERVIEWED.

WHEN INTERVIEWING STUDENTS, YOU GOT TO BE CAREFUL, NO DOUBT ABOUT IT, BUT STAFF, CUSTODIAL STAFF, THEY KNOW A LOT. THEY HAVE THEIR EYES OPEN. I'LL GIVE YOU SOME OF THE STORIES.

YOU JUST HAVE TO BE CAREFUL.

TEACHER HAD 6 DIFFERENT STUDENTS, WHO SAY SHE GAVE THEM ANSWERS. HER RESPONSE WHEN ASKED, "WHY ARE 6 STUDENTS TELLING US YOU GAVE THEM ANSWERS ON THOSE TESTS?" HER BEST RESPONSE WAS, "PROBABLY A CASE OF MISTAKEN IDENTITY."

NOW, DID THAT ANSWER TELL YOU ANYTHING?

IT CERTAINLY DID US.

REINTERVIEWS. THIS IS IMPORTANT TO NOTE. ANYBODY WHO REMEMBERS "COLUMBO," THIS IS HOW IT WORKS. YOU KEEP GOING BACK. GOOD INVESTIGATORS ACT LIKE THEY'RE DUMB AS A BRICK, BUT THEY KEEP GOING BACK, JUST LIKE HE DID.

IN THE ATLANTA INVESTIGATION, NOT ONE TEACHER OR ADMINISTRATOR EVER CONFESSED AND ADMITTED WRONGDOING ON THE FIRST INTERVIEW, NOT ONE, YET 82 ENDED UP CONFESSING.

YOU MUST GO BACK, AND THE MORE INFORMATION YOU GAIN, THE MORE PRESSURE YOU PUT ON AS THE QUESTIONS GO FORWARD. YOU HAVE TO CONSIDER POLYGRAPHS.

THESE ARE EDITORIAL COMMENTS BY SOMEBODY WHO WAS INVESTIGATED IN THE SYSTEM.

I'M NOT A PROFESSIONAL EDUCATOR. I AM A SCHOOL BOARD LAWYER. THE TESTS SHOULD BE ABOUT THE CHILDREN, NOT ADULTS. IF YOU TAKE THE ADULTS OUT OF IT WHERE THE TEST IS NOT ABOUT THEM, THEY HAVE NO REASON TO CHEAT.

WHEN THE TEST BECAME ABOUT TEACHERS AND SCHOOLS, TEACHERS AND SCHOOLS HAD SOMETHING AT STAKE, A REASON TO CHEAT.

BECAUSE THE TEST TIED TO ADULTS, IMAGE BECOMES IMPORTANT. IT WAS AN OVERRIDING FACTOR IN ATLANTA. IMAGE IS NOT MORE IMPORTANT THAN CHILDREN, AND NOTHING, ABSOLUTELY NOTHING, IS MORE IMPORTANT THAN CREATING A CULTURE OF INTEGRITY WHERE AN HONOR SYSTEM CONTROLS AND SUSPICIOUS ACTIVITY IS REPORTED AND ENCOURAGED TO BE REPORTED.

IN ATLANTA, THEY HAD JUST THE OPPOSITE. THEY HAD A CULTURE OF FEAR AND A CULTURE OF INTIMIDATION, AND PEOPLE WHO REPORTED GOT PUNISHED AND LOST THEIR JOBS. THAT SILENCES OTHER PEOPLE.

IN A SYSTEM WHERE PEOPLE REPORT AND SOMETHING IS DONE ABOUT IT, THAT MAKES OTHERS WILLING TO DO IT, AS WELL, AND THIS GOES FOR BOTH ADULTS AND STUDENTS. IT SHOULD BE AN HONOR SYSTEM.

THAT'S IT.

**JACK BUCKLEY:** THANKS VERY MUCH. TISHA?

**TISHA EDWARDS:** GOOD AFTERNOON.

MY PRESENTATION WILL FOCUS ON SOME LESSONS LEARNED FROM BALTIMORE CITY PUBLIC SCHOOLS TEST MONITORING AND INVESTIGATION PROCESS. I WILL PROVIDE A SENSE OF OUR SCHOOL SYSTEM'S BEST PRACTICES ON ITS TEST MONITORING AND INVESTIGATION PROCESS AND SHARE RECOMMENDATIONS FOR RESTORING CREDIBILITY.

WE HAD A PARTICULAR INCIDENT THAT REALLY RAISED OUR ORGANIZATIONAL AWARENESS AROUND ISSUES RELATING TO CHEATING IN OUR SCHOOL SYSTEM. ONE OF OUR SCHOOLS WAS ALLEGED TO HAVE ALTERED ITS TEST SCORES FROM THE MARYLAND STATE ASSESSMENT IN ORDER TO QUALIFY FOR BLUE RIBBON STATUS, AND THERE HAD BEEN PREVIOUS ALLEGATIONS CONCERNING THIS SCHOOL THAT WERE FOUND TO BE UNSUBSTANTIATED, AND TO BOB'S POINT, OUR ALLEGATIONS ACTUALLY STARTED FROM A PARENT, AND THIS PARENT WAS PRETTY ADAMANT ABOUT TAKING THE ALLEGATIONS TO THE TOP.

NOT ONLY DID SHE NOTIFY THE SUPERINTENDENT. SHE ALSO NOTIFIED THE STATE, AND SHE TOOK IT UPON HERSELF TO COME TO EVERY BOARD MEETING FOR ALMOST 6 MONTHS AND TESTIFY PUBLICLY, AND SHE HAD NO CONCERNS ABOUT BEING ANONYMOUS. SHE WANTED PEOPLE TO KNOW THAT SHE HAD THIS CONCERN, AND SHE PUT THE ORGANIZATION ON NOTICE THAT THERE WAS THIS ALLEGATION. THIS PARTICULAR INVESTIGATION LASTED ABOUT 18 MONTHS, WHICH REALLY PUT THE STRAIN ON THE ORGANIZATION BECAUSE THERE WAS THIS SENSE THAT WE WERE NOT DOING ANYTHING, BUT IN ALL ACTUALITY, WE WERE MOVING REALLY FEVERISHLY TO ASCERTAIN WHAT WAS HAPPENING,

AND WE WERE ENGAGED IN LOTS OF ANALYSIS, LOTS OF INTERVIEWING AND REINTERVIEWING, AS BOB COMMUNICATED IN HIS PRESENTATION.

#### CITY SCHOOLS' EXPERIENCE

CHALLENGES ESTABLISHING CULPABILITY. ONE OF THE REALLY DIFFICULT THINGS--AND I DON'T THINK THAT WE REALLY TALKED ABOUT THAT TODAY-- IS, ONCE YOU FIND OUT THAT THERE HAS BEEN WIDESPREAD CHEATING, IT IS A WHOLE OTHER SITUATION TO FIND OUT WHO IS THE CULPRIT IN THAT SITUATION, AND WE HAD A REAL HARD TIME ESTABLISHING CULPABILITY. PEOPLE WERE NOT TALKING.

THERE WAS OVERSENSITIVITY IN REPORTING OF ALLEGATIONS. THERE WAS BACKTRACKING OF ALLEGATIONS ONCE STAFF AND PARENTS WERE INTERVIEWED. AS YOU ALL MAY KNOW, SCHOOLS ARE, IN MANY CASES, FAMILIES. THEY ARE PART OF A COMMUNITY, AND ONCE PEOPLE FIND OUT THAT THERE IS GOING TO BE HIGH ACCOUNTABILITY, A LOT OF TIMES, THEY BACKTRACK FROM PREVIOUS STATEMENTS. SO THAT MADE IT VERY DIFFICULT FOR US.

BY CONDUCTING MULTIPLE ANALYSES--COMMUNICATION ANALYSIS, ERASURE ANALYSIS--CITY SCHOOLS ESTABLISHED A VERY CLEAR CASE THAT THERE WAS LIKELY SYSTEMIC CHEATING OCCURRING AT THIS PARTICULAR SCHOOL. OUR

LESSONS LEARNED IS THAT ANALYSIS IS NEVER ENOUGH TO ESTABLISH CULPABILITY.

UNLIKE ATLANTA, WE DIDN'T HAVE ANYBODY CONFESS, AND SO IT MADE IT REALLY DIFFICULT FOR US TO FIGURE OUT WHAT THE INDIVIDUAL ACCOUNTABILITY SHOULD LOOK LIKE, EVEN THOUGH WE WERE PRETTY SURE THAT CHEATING HAD OCCURRED. UNLESS COMPREHENSIVE INVESTIGATIONS OCCUR, THE CREDIBILITY OF DISTRICTS AND SCHOOLS SUFFER.

WHILE THE ORGANIZATION IS UNDER INVESTIGATION, WE KNEW THAT WE WERE GOING TO TAKE A HIT, BUT WE FELT LIKE THE PRICE WAS WORTH IT.

AGAIN, IT WAS AN 18-MONTH INVESTIGATION.

EVERY BOARD MEETING, WE HAD SOMEONE CALLING US TO THE CARPET.

WE HAD FOLKS, BECAUSE OF THE TIME IT WAS TAKING, DIDN'T FEEL LIKE WE WERE MOVING FAST ENOUGH, AND SO THERE WAS THIS MOMENT IN TIME WHERE IT WAS VERY DIFFICULT FOR THE ORGANIZATION, BUT OUR SUPERINTENDENT WAS VERY CLEAR THAT NOTHING MATTERED MORE THAN THE INTEGRITY OF OUR SCHOOL SYSTEM AND THAT WE WOULD NOT STAND BY AND ALLOW ANYONE TO CALL INTO QUESTION THE CAPACITY OF OUR CHILDREN.

HE REALLY SAW THIS AS AN ATTACK ON TEACHING AND LEARNING IN OUR DISTRICT.

WE, AFTER THIS PARTICULAR CASE, DID A LOT OF WORK TO STRENGTHEN ALL ASPECTS OF OUR INTERNAL CONTROLS. THE MARYLAND STATE DEPARTMENT OF EDUCATION HAS PUBLICLY SAID THAT BALTIMORE CITY HAS ONE OF THE MOST STRINGENT TESTING PROTOCOLS IN THE STATE AS A RESULT OF THIS INCIDENT. SO WE TRIED TO LEARN FROM IT, AND WE TRIED TO CREATE IT AS AN OPPORTUNITY TO BE BETTER AS AN ORGANIZATION.

WE DID CONDUCT MULTIPLE TYPES OF ANALYSIS IN ORDER TO ESTABLISH CULPABILITY WITH CERTAINTY, AND WE ARE CONTINUING TO ADHERE TO THE CODE OF MARYLAND STATUTES FOR CONDUCTING TESTING INVESTIGATIONS.

**JACK BUCKLEY:** THANK YOU VERY MUCH.

**TISHA EDWARDS:** I'M NOT FINISHED.

**JACK BUCKLEY:** I'M SORRY.

[LAUGHTER]

**JACK BUCKLEY:** SUCH A PREGNANT PAUSE, I THOUGHT--

**TISHA EDWARDS:** THANK YOU. THANK YOU.

SO WE TOOK STEPS FOR PREVENTING POTENTIAL VIOLATIONS. WE ESTABLISHED AND IMPLEMENTED INTERNAL CONTROLS THAT WOULD SIGNAL PROBLEMS. WE REALLY HAVE WORKED HARD TO BE PROACTIVE. WE USED SOME OF THE ANALYSES THAT YOU'VE TALKED ABOUT EARLIER TODAY--GROWTH ANALYSIS, BENCHMARK ANALYSIS, LEAVER ANALYSIS WHERE WE LOOK AT STUDENTS AND HOW THEY ACHIEVE ONCE THEY LEFT THE SCHOOL WHERE THEY WERE TESTED THE PREVIOUS YEAR, AND COHORT ANALYSIS.

WE HAVE ALSO DEVELOPED A VERY DETAILED MONITORING PLAN WHERE WE MONITORED ALL OUR HIGH-STAKES ASSESSMENT ANNUALLY. NONE OF THIS IS DONE BY STATE MANDATE OR DIRECTIVE. THIS IS ALL BASED ON ORGANIZATIONAL WILL AND OUR DESIRE TO PUT OURSELVES IN A POSITION WHERE WE NEVER HAVE THE SITUATION THAT HAS HAPPENED IN THE PAST TO HAPPEN AGAIN, AND WE'VE ENHANCED THE EDUCATION AND AWARENESS OF ALL OF OUR STAKEHOLDERS.

WE NOW PROVIDE MANDATORY TRAINING FOR ALL PERSONS PARTICIPATING IN TEST ADMINISTRATION. IN THE PAST, IT WAS ONLY LIMITED TO TESTING COORDINATORS, SCHOOL-BASED TESTING COORDINATORS. NOW ALL PRINCIPALS, ALL TEACHERS, ALL TEST MONITORS, EVERYONE WHO HAS ANY ROLE IN TESTING ADMINISTRATION MUST GO THROUGH AN ANNUAL TESTING TRAINING DONE BY OUR ORGANIZATION.

WE DISSEMINATE A VIDEO FOR EVERYONE IN THE SCHOOL SYSTEM TO LOOK AT EVERY YEAR. IT'S A VIDEO WHERE DR. ALONSO GETS ON CAMERA AND, FOR THE RECORD, COMMUNICATES HOW IMPORTANT TESTING INTEGRITY IS.

THIS IS A VIDEO THAT THE CUSTODIAN WATCHES, THE SECRETARY WATCHES, THE TEACHERS WATCH. IT HAS WIDESPREAD DISTRIBUTION THROUGHOUT THE ORGANIZATION.

ALL OF OUR PERSONNEL WHO ARE INVOLVED IN THE ADMINISTRATION ARE ALSO REQUIRED TO SIGN A NONDISCLOSURE FORM, AND WE COLLABORATE VERY STRONGLY WITH THE STATE TO ENSURE, AGAIN, THAT OUR TESTING PROTOCOLS ARE ALIGNED WITH BEST PRACTICES.

DETECTING POTENTIAL VIOLATIONS.

WE HAVE ESTABLISHED WELL-DEFINED PROTOCOLS. WE MADE AN INVESTMENT OF ABOUT HALF A MILLION DOLLARS A YEAR IN TESTING MONITORING AND ENSURING THAT WE HAVE TIGHT PROTOCOLS AND WE'VE COMMUNICATED THOSE AND MEMORIALIZED THEM, AND WE FLAG POTENTIAL CONCERN WITH SUPPORT FROM OUR STAKEHOLDERS. AGAIN, THE STATE IS A REAL WONDERFUL PARTNER IN THIS PROCESS.

THE STATE PROVIDES US WITH SCORING ALERTS, WHICH WE ACT ON VERY SWIFTLY, AND THE STATE IS ALSO RESPONSIBLE FOR PERFORMING THE ERASURE ANALYSES WHEN NECESSARY.

IN TERMS OF INVESTIGATING POTENTIAL VIOLATIONS, AS YOU'VE DISCUSSED EARLIER TODAY, WE DO HAVE A TIP LINE. WE GET INFORMATION BOTH ON THE STATE LEVEL AND AT THE LOCAL LEVEL.

WE TALKED ABOUT SOME OF THE PROTOCOLS, AND WE DO HAVE A STANDARD AND EXPECTATION THAT EVERY ALLEGATION IS INVESTIGATED. THERE IS NOT A SINGLE ALLEGATION THAT IS IGNORED OR PUT IN A FILE.

WE'VE DIRECTED THE RESOURCES SO THAT WE CAN INVESTIGATE EVERY ALLEGATION, AND WE'VE STARTED TO IMPLEMENT STRONG QUALITATIVE EVIDENCE IN HELPING US TO HAVE AN UNDERSTANDING OF WHAT'S HAPPENING IN SCHOOLS AND, IN PARTICULAR, TESTING ADMINISTRATIONS.

ONE OF THE THINGS THAT WE WOULD ADVISE OTHER LOCAL SCHOOL SYSTEMS IS THAT YOU REALLY NEED TO HAVE A STRONG INVESTIGATIVE REPORT. YOUR CONCLUSIONS NEED TO BE BASED ON SOLID DATA. WHERE POSSIBLE, YOU NEED TO TRIANGULATE THE DATA. WE'VE TALKED A LITTLE BIT ABOUT LOOKING AT DIFFERENT DATA SETS TO GET A SENSE OF HOW IS ONE

THING HAPPENING IN ONE ENVIRONMENT AND THEN THERE'S ANOTHER DATA SET THAT COMPLETELY CONTRADICTS THAT REALITY. SO TRIANGULATING THE DATA IS VERY IMPORTANT.

AND YOU NEED TO PRESENT YOUR FINDINGS TO PROPER INTERNAL AUTHORITIES. OUR CEO IS MADE AWARE OF EVERY INVESTIGATION ALONG THE WAY. WE MEET WITH HIM ON A REGULAR BASIS SO THAT HE UNDERSTANDS WHERE WE ARE IN THE INVESTIGATIVE PROCESS.

THE OTHER THING WE SHOULD SAY IS, YOU SHOULD PLAN FOR LITIGATION. THERE IS NO WAY THAT YOU ARE GOING TO HAVE A STRONG INVESTIGATIVE PROTOCOL AND PROCESS IN YOUR DISTRICT AND NOT BE INVOLVED IN FUTURE LITIGATION. SO YOU HAVE TO BE PREPARED FOR LITIGATION ON THE FRONT END.

MAINTAINING CREDIBILITY.

THIS IS SOMETHING, AGAIN, THAT BOB REFERRED TO IN HIS PRESENTATION. WE FELT LIKE BEING OUT FRONT, AND OUR SUPERINTENDENT ACTUALLY HAD A PRESS CONFERENCE AROUND THE ALLEGATIONS IN OUR SCHOOL SYSTEM.

IT WAS IMPORTANT THAT HE ESTABLISH CREDIBILITY FOR THE DISTRICT, AND EVEN THOUGH WE HAD THESE PARTICULAR INCIDENCES WHERE WE HAD CONCERNS, WE ALSO KNEW THAT WE HAD TO HAVE THE PUBLIC CONTINUE TO

BUY INTO OUR SCHOOL SYSTEM AND TO KNOW THAT WE WERE GOING TO DO THE  
RIGHT THING FOR OUR CHILDREN.

MY TIME IS UP.

**JACK BUCKLEY:** KEEP GOING. I OWE YOU--

**TISHA EDWARDS:** NO. I'M GOOD.

[LAUGHTER]

I'M GOOD.

**JACK BUCKLEY:** YOU ARE?

**TISHA EDWARDS:** YEAH. SURE.

**JACK BUCKLEY:** STEVE.

**STEVE FERRARA:** THANK YOU.

GOOD AFTERNOON, EVERYONE, AND THANK YOU, JACK AND EUNICE, FOR  
GIVING ME THE OPPORTUNITY TO SPEAK.

I'M GONNA TRY TO ADDRESS ISSUES AROUND INVESTIGATION, AND THIS IS BASED ON MY--OH, YEAH; THAT'S A GOOD IDEA--EXPERIENCE AS A FORMER STATE ASSESSMENT DIRECTOR. SOME PEOPLE WOULD SAY I'M ONE OF THE MANY RECOVERING STATE ASSESSMENT DIRECTORS.

SO...

**JACK BUCKLEY:** TRY THE ONE ON THE LEFT.

**STEVE FERRARA:** THERE WE GO. OK.

SO AS AN OVERVIEW, I WANT TO MAKE SOME BACKGROUND COMMENTS, AND I HOPE YOU'LL INDULGE ME. SOME OF THEM ECHO THINGS THAT YOU'VE HEARD BEFORE, BUT THEY'RE BACKGROUND TO WHAT I WANT TO SAY ABOUT CONDUCTING INVESTIGATIONS AND CONSIDERATIONS FOR CONDUCTING THOSE INVESTIGATIONS.

SO TO START WITH, I WANT TO DRAW A CONCEPTUAL DISTINCTION BETWEEN TEST SECURITY AND TESTING INTEGRITY, AND I WAS PLEASED TO SEE THAT NCES HAS TREATED THIS AS AN INTEGRITY ISSUE, NOT JUST A TESTING SECURITY ISSUE.

TESTING INTEGRITY OBVIOUSLY REQUIRES TEST SECURITY--SECURITY OF THE MATERIALS, APPROPRIATENESS OF THE PROCEDURES, AND SO FORTH. IT

ALSO REQUIRES PROFESSIONAL ETHICAL BEHAVIOR AND THINKING, AND WE HAVE HEARD REFERENCES TO THAT THROUGH THE MORNING AND THROUGH THESE PREVIOUS PRESENTATIONS.

WITHOUT PROFESSIONAL ETHICAL BEHAVIOR, YOU CAN'T REALLY HAVE TEST SECURITY. WHEN YOU HAVE THOSE TWO THINGS TOGETHER, WHAT YOU CAN GET IS WHAT WE ARE ALL LOOKING FOR--TRUSTWORTHY DATA AND SOMETHING THAT OFTEN GETS LOST. I'VE ONLY HEARD ONE OR TWO MENTIONS OF IT TODAY.

IT HELPS--BY HAVING TESTING INTEGRITY, YOU AVOID INAPPROPRIATE MESSAGES TO CHILDREN. WHEN YOU THINK ABOUT IT--I'VE HAD TO SAY THIS IN PUBLIC FORUMS BEFORE, INCLUDING A COURTROOM--WHEN A TEACHER CHEATS WITH STUDENTS, THE MESSAGE TO THE STUDENTS IS, "CHEATING IS OK, AND MAYBE THERE'S SOMETHING WRONG WITH ME. THAT'S WHY THE TEACHER IS CHEATING."

SO THE WAY I THINK OF THIS IS A MUTUALLY SUPPORTIVE TEST SECURITY SYSTEM AND A CULTURE OF PROFESSIONAL ETHICS. WE'VE HEARD REFERENCES TO CULTURE ALL DAY SO FAR. SO THE SECURITY SYSTEM PROVIDES RULES, GUIDELINES, AND SO FORTH. IT SHOULD REQUIRE TRAINING.

THERE SHOULD BE EXTENSIVE TRAINING LIKE THE TYPE THAT TISHA HAS REFERRED TO. I THINK YOUR SYSTEM IS PRETTY IMPRESSIVE, WHAT I'VE HEARD FROM YOU SO FAR.

THIS NOTION OF A CULTURE--SCHOOLS AND SCHOOL SYSTEMS ARE MINI CULTURES. AT LEAST, I VIEW THEM THAT WAY.

AND WHAT IS A CULTURE? WELL, IF YOU GO TO THAT--IF YOU LOOK UP DEFINITIONS OF "CULTURE," I CHOSE AN EASY ONE. IT IS A SET OF SHARED ATTITUDES, VALUES, GOALS, AND PRACTICES THAT CHARACTERIZE AN INSTITUTION OR ORGANIZATION.

I THINK THAT APPLIES VERY WELL TO WHEN YOU THINK ABOUT SCHOOLS AND SCHOOL SYSTEMS AS CULTURES. IN ORDER TO SUPPORT THE SECURITY RULES AND PROCEDURES AND SO FORTH THAT STATES AND LOCAL SCHOOL SYSTEMS SET UP, YOU NEED THAT CULTURE OF PROFESSIONAL ETHICAL THINKING AND BEHAVIOR,

AND YOU DON'T NECESSARILY HAVE TO PERSUADE PEOPLE TO BUY INTO THE CULTURE. YOU DO HAVE TO CONVINCING THEM THAT THEY HAVE TO ABIDE BY IT.

WHEN YOU HAVE THOSE TWO THINGS TOGETHER, THEN YOU HAVE THE CHANCE TO HAVE A SYSTEM WITH GREAT INTEGRITY AND MINIMIZE VIOLATIONS,

WHETHER THEY ARE INADVERTENT OR ADVERTENT. SO HOW DOES THE SYSTEM AND CULTURE WORKING TOGETHER SUPPORT INVESTIGATIONS? I'M GOING TO DO THIS VERY BRIEFLY.

YOU CAN THINK HYPOTHETICALLY ABOUT 3 GROUPS OF PEOPLE.

LEADERS IN THE SCHOOL WHO THINK AND PRACTICE IN PROFESSIONALLY ETHICAL WAYS, THEY BUY INTO THE NOTION. THEY ACT ACCORDINGLY, AND THEY'RE FAIRLY VOCAL ABOUT IT. THEY ARE LEADERS ABOUT THINGS LIKE PROFESSIONAL ETHICS.

THERE MAY BE A SECOND GROUP THAT ACTS APPROPRIATELY. THEY MAY NOT BUY INTO THE SYSTEM. THEY MAY NOT BE VOCAL ABOUT IT, BUT THEY GO ALONG WITH THE SYSTEM, AND THEN THERE ARE OTHER PEOPLE.

NOW, THE REASON I LIKE TO THINK ABOUT THOSE 3 GROUPS IS BECAUSE IT HELPS YOU THINK ABOUT WHEN YOU'VE DETECTED ANOMALIES AND YOU HAVE TO FOLLOW UP WITH INVESTIGATION. I AGREE YOU HAVE TO FOLLOW UP ON EVERYTHING. IT ENCOURAGES AND SUPPORTS PROFESSIONALLY ETHICAL THINKING AND BEHAVIOR, AND IT ENCOURAGES OTHER HELPFUL PARTS OF THE SYSTEM, LIKE REPORTING OF IRREGULARITIES AND COOPERATIONS WITH INVESTIGATIONS WHEN IRREGULARITIES MAY OCCUR.

SO WITH THAT AS A BACKGROUND, I'M NOT GOING TO TALK EVEN FASTER THAN I ALREADY AM. I WANT TO TALK ABOUT A COUPLE OF VIGNETTES THAT I THINK ILLUSTRATE A COUPLE OF POINTS. ONE OF THE THINGS THAT I LEARNED--I THINK YOU CAN JUST ARRIVE AT BY LOGIC--IS, WHAT HAPPENS WHEN A SCHOOL IS INVESTIGATED FOR POTENTIAL VIOLATIONS, WHETHER THE VIOLATIONS OCCURRED OR NOT, WHETHER ANYBODY IS CAUGHT AT SOMETHING INAPPROPRIATE OR NOT.

WHAT OFTEN HAPPENS IS, THE BEHAVIOR IN THE SCHOOL CHANGES, TYPICALLY IN A POSITIVE DIRECTION. YOU CAN SEE THAT IF YOU JUST GO BACK TO THE SCHOOL AFTER THE INVESTIGATION. YOU SEE THAT BEHAVIOR HAS CHANGED.

ONE OF THE THINGS THAT ALWAYS WORRIED ME IS THAT IT CAN DRIVE INAPPROPRIATE BEHAVIOR FURTHER UNDERGROUND. I HAVE SEEN INCIDENTS--AND I'M GOING TO TELL YOU A BRIEF EXAMPLE OF THIS--WHERE THE BEHAVIOR WAS ACTUALLY QUITE BRAZEN, AND SO WHEN THERE IS AN INVESTIGATION, THAT BRAZEN BEHAVIOR DISAPPEARS. THE QUESTION IS, WHAT SORT OF UNDERGROUND STUFF MAY BE GOING ON?

SO TWO QUICK VIGNETTES RELATED TO INVESTIGATIONS AND WHAT CAN GO ON AT SCHOOLS. WHEN I WAS A STATE ASSESSMENT DIRECTOR, I WENT TO VISIT A BLUE RIBBON SCHOOL TO WATCH AN ADMINISTRATION OF THE STATE

ASSESSMENT. IT WAS THE MARYLAND SCHOOL PERFORMANCE ASSESSMENT PROGRAM.

THE PRINCIPAL ASKED ME TO GO AND OBSERVE THE STATE TEACHER OF THE YEAR IN HER SCHOOL. SHE WAS OBVIOUSLY AND APPROPRIATELY VERY PROUD OF THIS WOMAN.

SO EVERYBODY KNOWS WHO I AM. I'M SITTING IN THE CLASSROOM. 10 MINUTES BEFORE THE ADMINISTRATION IS GOING TO START, THIS TEACHER OF THE YEAR WAS CHARISMATIC, ABSOLUTELY CHARISMATIC. SHE BROUGHT THE THIRD-GRADERS AROUND HER.

SHE SAT IN A ROCKING CHAIR. SHE READ THEM A STORY, AND SHE SAID, "I WANT YOU TO LISTEN CAREFULLY TO THIS STORY. IT'S GOING TO HELP YOU RELAX AND GET READY FOR THE ASSESSMENT," AND THEY WERE ENRAPT LISTENING TO THE STORY AND LISTENING TO THIS CHARISMATIC PERSON.

SHE PUT THE KIDS AT THEIR DESKS, DO THE TEST ADMINISTRATION. THEY BEGAN TO ADMINISTER THE ASSESSMENT.

I WAS ALREADY NERVOUS. THE STORY SHE HAD READ TO THE KIDS WAS RELEVANT TO THE THEME OF THE READING PASSAGES THE KIDS WERE READING IN THE ASSESSMENT. NOW, IN A CASE LIKE THAT, THERE WASN'T MUCH INVESTIGATION THAT I HAD TO DO--IT WAS A FOLLOW-UP--BUT I THINK IT

MAKES THE POINT OF HOW--I THINK IT ILLUSTRATES HOW DELICATE THINGS CAN BECOME WHEN YOU DETECT CHEATING AND WHEN YOU HAVE TO CONDUCT AN INVESTIGATION.

THIS IS ALL LEADING TO A POINT ABOUT INVESTIGATIONS.

A SECOND WAS SAME PROGRAM, MSPAP, ALL CONSTRUCTED-RESPONSE ITEMS. STUDENTS WROTE THE RESPONSES TO ALL ITEMS. WE DISCOVERED JUST BY DUMB LUCK THAT THE TEACHER HAD DICTATED RESPONSES TO ALL OF THE STUDENTS IN HER TESTING GROUP.

NOW, NO ONE SAW HER DO THAT. WE FOUND IT IN THE SCORING PROCESS. ALL STUDENTS HAD EXACTLY THE SAME RESPONSES TO ALL ITEMS OR ALMOST EXACTLY THE SAME RESPONSES TO ALL ITEMS.

THE INVESTIGATION THERE WAS NOT THAT DIFFICULT BECAUSE THE EVIDENCE WAS SO STRONG. NONETHELESS, INVESTIGATION WAS REQUIRED TO BUILD THE CASE WE NEEDED TO PROCEED.

SO AGAIN, THAT'S MORE BACKGROUND TO THESE CONSIDERATIONS THAT I'D LIKE TO BRING TO YOUR ATTENTION BEFORE I FINISH UP.

FIRST OF ALL, TYPICALLY, TEST DEVELOPMENT AND DELIVERY CONTRACTORS AREN'T NECESSARILY INVOLVED IN INVESTIGATIONS ONCE AN ANOMALY IS DETECTED.

NOT TYPICALLY, AS I SAID EARLIER IN RESPONSE TO JACK'S QUESTION, WE MAY BE INVOLVED IN TEST SECURITY DETECTION. WE ARE DEFINITELY INVOLVED IN PROTECTING SECURITY.

TYPICALLY, WE ARE NOT INVOLVED IN INVESTIGATIONS, AND I THINK IT IS IMPORTANT TO KEEP THE ROLES AND RESPONSIBILITIES CLEAR.

SECOND OF ALL, PRINCIPLES AND PROCEDURES FOR INVESTIGATIONS, THEY MUST BE CLEAR. THEY MUST WITHSTAND POTENTIAL CHALLENGES IN ADMINISTRATIVE LAW COURTS. EDUCATORS--PRINCIPALS, ASSISTANT PRINCIPALS, ASSISTANT SUPERINTENDENTS--WHO ARE ASKED TO CONDUCT THESE INVESTIGATIONS DON'T REALLY KNOW ANYTHING ABOUT INVESTIGATIVE PROCEDURES. THEY'RE NOT TRAINED TO DO THAT SORT OF THING. THEY DON'T HAVE THE MIND-SET THAT'S INVOLVED. I THINK WHAT WE NEED--THE KIND OF THING THAT TISHA REFERRED TO THAT'S GOING ON IN BALTIMORE CITY, THE KIND OF THING THAT BOB REFERRED TO--WE NEED TRAINING FOR THE PEOPLE WHO ARE GOING TO CONDUCT THOSE INVESTIGATIONS.

I'D SUGGEST THAT IT EVEN BE OUTSIDE PEOPLE.

THERE ARE IMPORTANT INVESTIGATION PROCEDURES, PROTOCOLS THAT PEOPLE FOLLOW. THERE ARE ALSO SUBTLE THINGS, LIKE FBI AGENTS AND OTHER INVESTIGATORS ARE TRAINED TO DETECT EVASIVE BEHAVIOR AND LYING, AND THAT CAN HELP MAKE A DIFFERENCE ON HOW MUCH FURTHER THE INVESTIGATION HAS TO GO.

VERY QUICKLY, THIS WILL COME UP IN THE NEXT PANEL, I'M SURE, BUT I WAS ASKED TO BRING IT UP. THERE ARE CHEATING RISKS WITH NEW DEVICES THAT ARE GOING TO HAVE A HUGE IMPACT ON HOW WE GO ABOUT INVESTIGATING WHEN WE FIND ANOMALIES. APPARENTLY, WHEN WE DO ONLINE TESTING, WE CAN LOCK THINGS DOWN. KIDS CAN'T GO OUTSIDE OF THE TEST. THEY CAN'T GO OUT ONTO THE WEB TO LOOK UP THINGS, AND WE CAN DEAL WITH WHO IS SITTING IN FRONT OF THE COMPUTER AND ALL THAT SORT OF THING.

TABLETS INTRODUCE A WHOLE NEW SET OF RISKS, LIKE TASK SWAPPING, GOING OUT ONTO THE WEB, AND SCREENSHOT CAPABILITIES.

THERE IS A VERY SIMPLE WAY TO TAKE SCREENSHOTS ON AN IPAD. SO EVENTUALLY, WE'LL HAVE TO WORK OUT HOW TO LOCK THAT STUFF DOWN.

THE PROBLEM WILL BE THAT NOT EVERY KID IS GOING TO HAVE SOME SORT OF DEVICE. IF THE SCHOOLS DECIDE TO ALLOW KIDS TO BRING THEIR OWN

DEVICES, THINK OF THE PROBLEMS OF TRYING TO INVESTIGATE WHEN THERE'S BEEN AN ANOMALY WHEN A STUDENT HAS USED HIS OWN DEVICE.

LAST COMMENT, SOMETHING BOB ALREADY SAID, IT'S NECESSARY TO INVESTIGATE, ENFORCE, AND SANCTION RIGOROUSLY AND CONSISTENTLY AND FOLLOW UP ABSOLUTELY ON EVERYTHING.

THANK YOU.

**JACK BUCKLEY:** THANKS VERY MUCH. LOU?

**LOU FABRIZIO:** ALL RIGHT.

GOOD AFTERNOON.

I WOULD LIKE TO THANK THE DEPARTMENT FOR GIVING ME THE OPPORTUNITY TO BE HERE TODAY. THIS HAS BEEN AN AMAZING DAY LISTENING TO ALL THE DIFFERENT PRESENTERS TALKING ABOUT ALL THESE VARIOUS ASPECTS OF TESTING INTEGRITY, BUT THERE OBVIOUSLY ARE LOTS OF THEMES THAT WE'RE HEARING IN THE DIFFERENT SESSIONS OVER AND OVER AGAIN, AND I'D LIKE TO JUST START OUT BY TALKING ABOUT THE RULES AND REGS IN THE ADMINISTRATIVE CODE.

IN NORTH CAROLINA, THESE ARE THE 3 MAIN THINGS THAT WE HOLD OUT AS THE ADMINISTRATIVE CODE WHICH ARE ALL STATE BOARD REGULATIONS, RULES AND REGULATIONS, AND THE FIRST ONE IS THE TESTING CODE OF ETHICS, AND I WANT TO BRING UP AN ISSUE THAT OCCURRED WHEN WE WENT THROUGH THE ADMINISTRATIVE PROCEDURES ACT IN NORTH CAROLINA.

THIS WAS IN 1997, AND IN THE MATERIALS THAT WE SUBMITTED TO GO THROUGH THE APA PROCESS, THE ATTORNEYS DIDN'T LIKE THE WORDING, WHICH WAS THAT THE STATE BOARD STRONGLY RECOMMENDS THE USE OF PROCTORS IN THE CLASSROOM, AND THE ATTORNEYS CAME BACK TO US AND SAID, "OK. LOOK. THIS IS ADMINISTRATIVE CODE. YOU EITHER HAVE TO SAY YOU HAVE TO HAVE THE PROCTORS, OR YOU DON'T MAKE ANY MENTION OF IT," AND WE SAID WE HAVE TO HAVE THEM, AND WE HAVE HAD THEM EVER SINCE.

IT HAS BEEN REQUIRED, AND EVERY YEAR, WE GET HATE MAIL FROM PEOPLE WHO ARGUE, "WHY DON'T WE TRUST THE TEACHERS? WHY ARE WE FORCING TEACHERS TO HAVE TO SPEND TIME PROCTORING IN OTHER CLASSROOMS?"

AND OUR RESPONSE IS, "BECAUSE WE WANT TO ENSURE THE TESTING INTEGRITY OF THE SYSTEM." SO I DID WANT TO BRING THAT UP AS THE MIND-SETS THAT YOU ARE HAVING TO CHANGE IN ORDER TO MAKE THIS ALL BE EFFECTIVE.

WE HAVE TALKED ABOUT THE FACT THAT WE NEED TO BE RECORDING AND MONITORING IRREGULARITIES, BUT I DIDN'T HEAR ANYONE TALK ABOUT A SPECIFIC SYSTEM. WE HAVE AN ACTUAL ONLINE SYSTEM WHERE DATA IS ENTERED ONLINE AND CAN BE TRACKED. ALL THE STEPS OF THE PROCESS ARE IN THAT SYSTEM. WE REQUIRE LOCAL INVESTIGATIONS BY THE L.E.A. WE HAVE STAFF WHOSE JOB IT IS TO ASSIST THE SCHOOL DISTRICTS IF THEY HAVE ANY QUESTIONS ABOUT HOW TO CONDUCT THAT INVESTIGATION, AND ANY SERIOUS IRREGULARITIES THAT GET IDENTIFIED REQUIRE, THEN, THE USE OF A CHECKLIST THAT MUST THEN BE SUBMITTED.

ONCE THE L.E.A. SUBMITS ITS REPORT, THEN IT IS UP TO THE STAFF IN ACCOUNTABILITY SERVICES TO REVIEW THE ACTUAL REPORT.

IF THE REPORT IS OF ONE OF THOSE SERIOUS IRREGULARITIES, WE THEN FORWARD THAT REPORT TO THE ATTORNEY WITH THE STATE BOARD, AND THAT PERSON, IN CONJUNCTION WITH THE ATTORNEY GENERAL'S OFFICE, MAKE DECISIONS AS TO WHETHER OR NOT A FURTHER INVESTIGATION NEEDS TO OCCUR OR WHETHER OR NOT THEY ARE OK WITH WHAT'S HAPPENED.

BUT ONE OF THE THINGS THAT WE HAVE NOT SPENT A LOT OF TIME TALKING ABOUT, ALTHOUGH IT HAS BEEN MENTIONED A LITTLE BIT, ARE THESE BARRIERS, AND ONE OF THE BARRIERS THAT I KNOW WE ARE NOT ALONE IN NORTH CAROLINA IS THE LIMITED RESOURCES.

EVERY YEAR, THE GENERAL ASSEMBLY CONTINUES TO MAKE CUTS TO THE BUDGET, AND THINGS HAVE TO BE ELIMINATED. WE DO NOT HAVE ANY PREASSIGNED STAFF WHOSE SOLE RESPONSIBILITY IS TO DO INVESTIGATIONS, AND WE DO NOT HAVE ANY SPECIFIC LINE ITEMS SET ASIDE FOR DOING INVESTIGATIONS, BUT BECAUSE OF THAT, THE FOLLOW-UP INVESTIGATIONS THAT DO OCCUR TEND TO TAKE A LITTLE LONGER THAN WE WOULD LIKE.

BALTIMORE IS A GOOD EXAMPLE, 18 MONTHS. THESE THINGS DO TAKE TIME.

BASED ON THE INVESTIGATIONS THAT ARE THEN DONE BY THE STATE BOARD ATTORNEY AND IN CONJUNCTION WITH THE ATTORNEY GENERAL'S OFFICE, SOME OF THE INDIVIDUALS THEN COME BEFORE WHAT'S CALLED THE STATE SUPERINTENDENT'S ETHICS ADVISORY COMMITTEE, AND THESE ARE INSTANCES WHERE REVOCATION OF LICENSES WILL OCCUR, AND SOME INVESTIGATIONS MAY ALSO RESULT IN CIVIL ACTION AND/OR, AND I ALREADY SAID, THE LOSS OF THE LICENSE.

ALL OF THAT KIND OF INFORMATION IS LAID OUT IN THAT TESTING CODE OF ETHICS THAT I STARTED OUT WITH. THAT TESTING CODE OF ETHICS SAYS WHAT THE SUPERINTENDENT IS SUPPOSED TO DO, WHAT THE TEACHER IS SUPPOSED TO DO, WHAT THE PROCTOR IS SUPPOSED TO DO. IT EVEN GIVES EXAMPLES OF UNETHICAL PRACTICES SO FOLKS CAN'T SAY, "WELL, I DIDN'T KNOW THAT THAT WASN'T ALLOWED," BECAUSE IT IS IN THAT CODE OF

ETHICS, AND THEN THAT CODE OF ETHICS ENDS WITH THE STATEMENT THAT CIVIL ACTION AND POSSIBLE REVOCATION OF LICENSE MAY OCCUR IF ONE IS ACCUSED OF THE VIOLATION OF THE TESTING CODE OF ETHICS.

SO ALL OF THAT STUFF IS LAID OUT IN THE PROCESS OR SYSTEM THAT WE HAVE. JUST SOME OTHER QUICK THOUGHTS.

ONCE WE GET THE WORD THAT THERE IS SOMETHING THAT APPEARS PRETTY DARN SERIOUS--AND THAT NORMALLY OCCURS BECAUSE ONCE SOMEBODY ENTERS ANYTHING INTO THAT ONLINE SYSTEM THAT WE HAVE, IT AUTOMATICALLY SENDS INFORMATION TO DIFFERENT INDIVIDUALS TO ALERT THEM--WHAT WE HAVE LEARNED IS THAT IF YOU HEAR ABOUT SOMETHING THAT YOU THINK IS GOING TO BE VERY SERIOUS, IT IS ALWAYS BEST TO LET THE STATE SUPERINTENDENT KNOW, TO LET OTHER MEMBERS OF THE TOP LEADERSHIP BE AWARE OF IT, ALSO TO MAKE OUR COMMUNICATIONS AND INFORMATION STAFF AWARE OF IT BECAUSE WHAT YOU DON'T EVER WANT IS TO BE CAUGHT OR FOR THOSE INDIVIDUALS TO BE CAUGHT ABOUT SOMETHING THAT THEY HAVE HEARD THROUGH SOMEBODY, AND IT IS ALWAYS GREAT WHEN WE ARE ABLE TO SAY, "YES. WE'RE AWARE OF THAT, AND WE ARE UNDERGOING A STUDY IN RELATIONSHIP TO IT."

I MUST HAVE HEARD THIS ON AT LEAST TWO OR 3 OTHER--OR SEEN IT ON TWO OR 3 OTHER SLIDES. PREVENTION IS REALLY THE BEST SOLUTION, WHICH

THEN GETS TO ANOTHER THING THAT YOU HAVE BEEN HEARING OVER AND OVER AGAIN.

I HAVE BEEN TALKING ABOUT THESE COMMON THEMES, THIS CULTURE OF HONESTY. TO US, IT IS FANTASTIC WHEN YOU HAVE STUDENTS THAT END UP SAYING THINGS LIKE, "BOY, THIS LOOKS JUST LIKE THE QUESTIONS WE HAD TWO WEEKS AGO," OR WHEN YOU HAVE TEACHERS THAT WILL SAY, "I THINK THE ASSISTANT PRINCIPAL HAS BEEN DOING CERTAIN THINGS THAT HE OR SHE SHOULD NOT BE DOING," OR WHEN YOU HAVE A PRINCIPAL REPORTING ON SOMETHING THAT'S HAPPENING.

IT IS ONLY WHEN THOSE KINDS OF SYSTEMS ARE IN PLACE WHERE THERE IS NO RETALIATION AND THERE IS NO PUNISHMENT FOR IT THAT I THINK IT REALLY WORKS WELL, AND WE HAVE BEEN WORKING VERY HARD.

SOMETIMES I HAVE TO END UP SAYING TO FOLKS THAT WE'VE MAYBE DONE SUCH A JOB OF MAKING FOLKS SO AWARE OF THE RIGIDITY OF THE SYSTEM THAT THEY ARE AFRAID TO TAKE ANY TINY DEVIATION WHICH THEY TECHNICALLY COULD, BUT THEY WILL CALL TO ASK IF IT IS OK, AND THEY WILL APOLOGIZE, AND OUR RESPONSE IS ALWAYS, "NO. WE WOULD MUCH RATHER YOU CALL TO CHECK IF IT IS OK THAN TO HAVE GONE AHEAD AND THEN FIND OUT AFTERWARDS THAT WAS NOT THE RIGHT THING TO DO."

SO THAT WHOLE ISSUE OF PREVENTION, THAT WHOLE ISSUE OF THE CULTURE OF HONESTY, I THINK, IS WHAT REALLY MAKES THE WHOLE SYSTEM WORK, AND IT'S NOT EASY. IT TAKES TIME.

WE DON'T KNOW ALL THE EXACT RIGHT THINGS TO DO, BUT MY TIME IS UP, SO I WILL STOP.

**JACK BUCKLEY:** HA HA! THANKS VERY MUCH.

ALL RIGHT. SO I THINK THERE'S CERTAINLY A FEW THEMES THAT I'M HEARING AGAIN THAT I'D LIKE TO ASK THE PANEL MAYBE TO ELABORATE ON.

THE FIRST ONE, I'M NOT SURE IF THERE'S CONSENSUS ON, BUT, BOB, YOU CAME DOWN PRETTY HEAVILY IN FAVOR OF THE NEED TO BRING IN OUTSIDE PERSONNEL, AND OTHERS TOUCHED ON THE FACT THAT CERTAINLY AT THE SCHOOL BUILDING LEVEL AND PERHAPS ALSO AT THE L.E.A. LEVEL, EXCEPT FOR OUR LARGEST AND MOST SOPHISTICATED ORGANIZATIONS, WE PROBABLY DON'T HAVE THE PERSONNEL RIGHT THERE ON THE SCENE WHO CAN ACTUALLY CONDUCT AN INVESTIGATION DISPASSIONATELY.

IS THAT FAIR TO CALL THAT CONSENSUS, OR DO WE THINK AT LEAST MAYBE FOR SOME TYPES OF INTEGRITY VIOLATIONS, THE LOCAL FOLKS CAN HANDLE IT, OR DO WE NEED TO BRING IN SOMEBODY ELSE?

**STEVE FERRARA:** I DON'T FULLY AGREE WITH THAT, I THINK. I MADE THE POINT THAT EDUCATORS ARE NOT TRAINED TO BE INVESTIGATORS, AND, IN FACT, THERE ARE REASONS WHY YOU DON'T WANT TO PUT A SCHOOL LEADER OR MANAGER IN THE POSITION OF HAVING TO INVESTIGATE ONE OF THEIR STAFF. I THINK YOU NEED PEOPLE WHO ARE TRAINED IN CONDUCTING INVESTIGATIONS.

MAYBE THEY ARE AN EMPLOYEE OF THE SCHOOL SYSTEM. MAYBE THEY'RE SHARED ACROSS LOTS OF LARGE SCHOOL SYSTEMS. MAYBE IT IS A STATE PERSON. THOSE ARE THINGS THAT NEED TO BE WORKED OUT, AND THERE ARE POLICY CONSIDERATIONS, BUT I DON'T THINK IT HAS TO BE AN OUTSIDE AGENCY.

**BOB WILSON:** I DON'T THINK IT HAS TO BE AN OUTSIDE AGENCY, AND I THINK THERE ARE SOME THINGS THAT COME UP THAT I MIGHT DISAGREE WITH STEVE JUST A LITTLE BIT THAT THE PRINCIPAL CAN TAKE A PRELIMINARY LOOK AT, BUT IF IT'S GOING TO BE SOMETHING WHERE IT LOOKS LIKE THERE MAY BE SOMETHING HERE, I THINK YOU NEED SOMEBODY FROM THE OUTSIDE OF THAT SCHOOL AND OUTSIDE OF THAT PARTICULAR ADMINISTRATION OF THAT SCHOOL HANDLING THAT INVESTIGATION, AND IT MAY BE YOUR LEGAL COUNSEL WHO DOES IT, BUT THERE ARE A LOT OF TRICKS TO THE TRADE TO PROPER INVESTIGATIONS, AND THAT IS JUST NOT WHAT EDUCATORS ARE TAUGHT TO DO.

EDUCATORS BY NATURE ARE COMPASSIONATE PEOPLE. INVESTIGATORS BY NATURE ARE...DIFFERENT. ARE DIFFERENT. IT IS A QUESTION-AND-ANSWER PROCESS THAT TAKES A LOT OF PATIENCE, AND IT TAKES A LOT OF TIME TO DO IT RIGHT.

LIKE I SAID, OUT OF ALL THOSE INTERVIEWS, WE DID NOT HAVE A SINGLE CONFESSION ON THE FIRST INTERVIEW, AND THOSE WERE BEING DONE BY STATE INVESTIGATORS, OUR GEORGIA BUREAU OF INVESTIGATION. IT WASN'T UNTIL WE UPPED THE HEAT, CHANGED THE ATMOSPHERE, GOT THEM OUT OF THE SCHOOLS, GOT THEM INTO LAW OFFICES, GOT THEM SURROUNDED BY SOME LAWYERS AND SOME OF THOSE STATE AGENTS AND BEGIN TO ASK THEM QUESTIONS THAT TRIANGULATED, AS TISHA SAID, SOMETIMES MORE THAN THAT, POINTS OF REFERENCE THAT THEY COULD NOT EXPLAIN THAT YOU BEGAN TO GET SOMEWHERE.

THAT'S JUST NOT WHAT EDUCATORS ARE UP TO. I THINK IT SORT OF DEPENDS. DOES IT APPEAR TO BE SOME ISOLATED SITUATION IN THE CLASSROOM THAT DOESN'T APPEAR TO BE TOO SERIOUS? FINE. YOU MIGHT BE ABLE TO HANDLE IT IN HOUSE, BUT ONCE IT LOOKS LIKE HONEST-TO-GOD CHEATING OF A RATHER SIGNIFICANT NATURE AND MAY INVOLVE MORE THAN ONE CLASS, I THINK YOU DEFINITELY GOT TO GO OUTSIDE. DOESN'T HAVE TO BE ANOTHER AGENCY. IT CAN BE A LAW FIRM. IT CAN BE PRIVATE INVESTIGATORS THAT WORK FOR THE SYSTEM, YOU KNOW, OR GET HIRED, CONTRACT EMPLOYEES, SOMETHING LIKE THAT.

**JACK BUCKLEY:** TISHA, WHAT DID YOU GUYS USE?

**TISHA EDWARDS:** I AGREE. WE DID NOT ALLOW ANY OF THE SCHOOL-BASED STAFF TO PARTICIPATE IN ANY OF OUR INVESTIGATIONS. WE DO HAVE AN INVESTIGATIONS UNIT IN OUR ACHIEVEMENT AND ACCOUNTABILITY OFFICE, AND WHEN WE RECEIVE ALLEGATIONS, PARTICULARLY THOSE THAT ARE VERY SUBSTANTIAL IN NATURE, WE DO GO OUT AND CONTRACT WITH INDIVIDUALS WHO HAVE EXPERTISE IN LAW ENFORCEMENT TO PARTNER WITH US AND TO BUILD THE CAPACITY OF OUR ACHIEVEMENT AND ACCOUNTABILITY OFFICE, BUT WE HAVE NEVER ALLOWED ANY SCHOOL-BASED STAFF TO PARTICIPATE IN THE INVESTIGATING OF ANY TESTING ALLEGATION.

**STEVE FERRARA:** BY THE WAY, SORT OF A SIDE COMMENT, I'M THINKING 5 OR 10 YEARS DOWN THE ROAD, AND WE HAVE ALL THESE PROCEDURES IN PLACE FOR CONDUCTING INVESTIGATIONS AND GETTING CONFESSIONS FROM PEOPLE LIKE HAPPENED IN ATLANTA. IF YOU READ THE SUNDAY "NEW YORK TIMES," THERE WAS A VERY INTERESTING STORY ABOUT THE FREQUENCY AND THE SORT OF PSYCHOLOGICAL EXPLANATIONS FOR PEOPLE MAKING FALSE CONFESSIONS IN LEGAL SITUATIONS. SO WE NEED TO MAYBE KEEP HALF AN EYE OUT FOR THAT IN THE FUTURE.

**JACK BUCKLEY:** I THINK, TISHA, YOU MENTIONED MARYLAND'S STATUTES FOR CONDUCTING TESTING INVESTIGATIONS, AND, LOU, CERTAINLY YOU WERE

DESCRIBING SOME, SOUNDS TO BE SOME PRETTY DETAILED REGULATIONS. WHAT KIND OF DETAIL IS REALLY IN THERE, AND IS THERE A MODEL THAT WE CAN POINT TO ACROSS STATES OR DISTRICT LEVEL FOR SORT OF CODIFYING THIS?

**TISHA EDWARDS:** I WOULDN'T BE AS BOLD TO SAY THAT MARYLAND SHOULD BE HELD OUT TO NECESSARILY BE THE STANDARD FOR EVERYONE ELSE. I THINK THAT THE PARTNERSHIP THAT OUR SUPERINTENDENT, ANDRES ALONSO, WHO IS A VERY BOLD LEADER, REALLY REQUIRED THAT WE PARTNER WITH THE STATE IN ORDER TO FIGURE OUT HOW TO DO THE RIGHT THING.

WE ARE THE ONLY L.E.A. IN THE STATE WHO HAD EVER HAD A PUBLIC PRESS CONFERENCE AND WHO HAD GONE OUT THERE PUBLICLY TO OWN THE FACT THAT TESTING IRREGULARITIES HAD HAPPENED IN OUR SCHOOL SYSTEM, AND ACTUALLY IN THE STATE, WE HAD ACTUALLY RECEIVED ADVICE AND COUNSEL THAT, EVEN THOUGH THERE WERE OTHER JURISDICTIONS WHERE THERE WERE KNOWN TESTING IRREGULARITIES, THAT THERE HAD BEEN A DECISION MADE THAT THAT INFORMATION WOULD NOT BE GIVEN OUT TO THE PUBLIC.

SO WE WERE SETTING A NEW STANDARD AROUND TRANSPARENCY, AND ONE OF THE OTHER THINGS I WILL SAY IS THAT ALTHOUGH THE STATE HAD REGULATIONS AROUND ANY L.E.A.'S ABILITY TO REVOKE A LICENSE, MOST L.E.A.s RESISTED DOING THAT WHEN THEY COULD NOT ESTABLISH CULPABILITY, EVEN IF THERE WERE SIGNS OF CHEATING, AND ANDRES, WHO

IS ALSO AN ATTORNEY IN HIS FORMER LIFE, WAS VERY BOLD ABOUT THE FACT THAT WE WERE GOING TO USE THE REASONABLE MAN STANDARD, AND HE SAID THE OLD SAYING, WHAT, "IF IT QUACK LIKE A DUCK AND WALKS LIKE A DUCK, IT'S A DUCK, AND WE'RE GOING TO HAVE ACCOUNTABILITY FOR THE DUCKS."

SO HE WAS OUT THERE SAYING TO SCHOOL LEADERS AND TO TEACHERS THAT-- EVEN THOUGH WITHIN THE INVESTIGATIVE PROCESS, HE COULD NOT DETERMINE SPECIFICALLY WHO HAD DONE WHAT--WE WERE ALL CLEAR THAT SOMETHING HAD HAPPENED, AND IT HAS BEEN A HUGE STRETCH FROM A LITIGATION PERSPECTIVE FOR US AND THE STATE TO BE ABLE TO MAINTAIN THE REASONABLE MAN STANDARD WHEN PURSUING SOMEONE'S CERTIFICATE.

SO I DON'T WANT TO GO AS FAR AS TO SAY THAT MARYLAND SHOULD BE ESTABLISHING THE STANDARD, BUT I DO THINK THAT WE ARE TREADING IN GROUNDS THAT WE HAVE NEVER DONE BEFORE, AND BALTIMORE IS KIND OF PUSHING THIS ISSUE OF TRANSPARENCY AND VIGILANCE AROUND THESE TYPES OF ISSUES.

**LOU FABRIZIO:** THE COMMENTS THAT TISHA JUST MADE REMIND ME OF INSTANCES WHERE WE HAVE DONE INVESTIGATIONS AND THEN WHEN WE PRESENT THE SO-CALLED EVIDENCE TO THE ATTORNEYS, INCLUDING ANSWER SHEETS OR TEST BOOKLETS WHERE STUDENTS WITH DISABILITIES HAD MARKED IN THE TEST BOOKLET AND THEN THOSE RESPONSES NEEDED TO BE

TRANSFERRED TO THE ANSWER SHEET, DETERMINING THAT THEY DIDN'T MATCH-UP AND SOMETHING CHANGED SOMEWHERE, AND THEN YOU GET INTO THAT WHOLE ISSUE OF DID ANYBODY SEE THE PERSON DO IT?

AND THAT'S A FRUSTRATION THAT I THINK WE AT THE STATE LEVEL SOMETIMES HAVE, IS THAT WE WILL BE TOLD, YOU KNOW, "WE'VE GOT BIGGER FISH TO FRY, AND WE NEED TO WORRY ABOUT THAT, AND WE'RE NOT GOING TO WIN THIS ONE," AND THAT'S A FRUSTRATION. TO ANSWER YOUR QUESTION, JACK, I'M NOT SAYING WE'VE GOT THE BEST SET OF THINGS, BUT WE'D BE MORE THAN WILLING TO SHARE THEM WITH THE DEPARTMENT. I DID JUST LOOK AT THE LIST OF IRREGULARITIES THAT WE MONITOR, AND THERE ARE 60 OF THEM.

**JACK BUCKLEY:** ALL RIGHT. WELL, AT THAT, I THINK WE CAN THROW IT OPEN TO THE BROADER PANEL IN THE AUDIENCE.

WAYNE?

**WAYNE CAMARA:** I WANT TO COME BACK TO THE ISSUE OF THE DETECTION WHEN STAFF AT A SCHOOL OR A PARENT OR A STUDENT MAY THINK THAT THERE IS SOME IRREGULARITY. THE OTHER PANEL TALKED A LITTLE BIT ABOUT THE TIP HOTLINE, BUT IT ALSO SOUNDS LIKE THERE ARE MANY INSTANCES WHERE A DISTRICT, AN L.E.A., THE PRIME WAY OF REPORTING THESE INSTANCES

IS GOING WITHIN THE SCHOOL BUILDING, AND, I MEAN, I QUESTION WHETHER THAT IS THE BEST VEHICLE.

IF I'M A COLLEAGUE OF SOMEONE WORKING IN A SCHOOL, IF I'M A PARENT--MY WIFE IS A FIRST GRADE TEACHER, AND SHE INSTILLED INTO ME THAT WE'VE HAD KIDS IN SCHOOL, TOO, AND AS A PARENT OR AS A COLLEAGUE, YOU ARE A LITTLE BIT RELUCTANT TO BLOW THE WHISTLE ON A SCHOOL OR A TEACHER. YOU ARE WORRIED ABOUT WHAT TEACHER YOUR CHILDREN ARE GOING TO BE PLACED WITH NEXT YEAR, AND YOU ALSO ARE NOT SURE WHAT HAPPENED, AND SO WE HAVE A LOT OF SUCCESS WITH AN ANONYMOUS TIP HOTLINE.

WE GET A LOT OF FALSE POSITIVES, A LOT OF FALSE POSITIVES, BUT IT ALLOWS ME, AS ANOTHER TEACHER OR AS A PARENT, TO NOT FEEL THAT THERE'S GOING TO BE ANY RETALIATION, AND I CAN REPORT SOMETHING WHEN I AM CLEARLY LESS THAN 100% CERTAINTY, AND I AM WONDERING TO WHAT EXTENT YOU FEEL IT IS IMPORTANT TO GO THROUGH THE SCHOOL BUILDING ROUTE, OR IS IT MORE ADVISABLE TO HAVE SOME ENTITY, WHETHER IT IS THE STATE OR WHETHER IT'S SOMEBODY ELSE. I'M JUST WONDERING, CONTRASTING THE BUILDING SOLUTION VERSUS AN OUTSIDE SOLUTION.

**TISHA EDWARDS:** I WOULD REITERATE WHAT BOB SAID IN HIS PRESENTATION ABOUT YOU NEED TO FOLLOW UP ON GOSSIP. A LOT OF TIMES, PEOPLE DON'T

KNOW THAT THEY ARE REPORTING SOMETHING, AND THEY ACTUALLY ARE, THROUGH CASUAL CONVERSATIONS WITH FOLKS, WHETHER IT IS AT A PTA MEETING OR A PARENT-TEACHER CONFERENCE.

WE HAD LOTS OF SITUATIONS--SHOULDN'T SAY LOTS. THAT'S AN EXAGGERATION. WE HAVE HAD SITUATIONS WHERE PARENTS HAVE BEEN SITTING AT A SPECIAL ED. IEP MEETING AND HAVE SAID, "I DON'T REALLY UNDERSTAND THESE TEST SCORES BECAUSE I KNOW THAT MY SON OR MY DAUGHTER IS NOT PERFORMING AT THESE LEVELS, RIGHT? THIS JUST DOESN'T MAKE SENSE TO ME," AND SAY THAT IN AN OPEN FORMAT WHERE PEOPLE ARE TAKING MINUTES, AND THAT INFORMATION IS FILTERED UP TO THE SCHOOL SYSTEM THROUGH OUR INVESTIGATIVE OFFICE TO SAY THAT THAT'S NOT WHERE A PARENT IS ALLEGING CHEATING PER SE, BUT A PARENT IS SAYING, "THE DATA ABOUT MY CHILD DOESN'T MAKE SENSE TO ME," AND SO I THINK THAT I DO AGREE THAT IT IS NATURAL FOR PEOPLE TO HAVE CONCERNS ABOUT BRINGING INFORMATION FORWARD ABOUT PEOPLE THAT THEY ARE IN A RELATIONSHIP WITH, WHETHER THAT'S A WORK RELATIONSHIP OR SOME OTHER KIND OF PERSONAL RELATIONSHIP.

I THINK WHAT WE WOULD ADVISE SCHOOL SYSTEMS TO DO IS HAVE AS MANY VEHICLES AS POSSIBLE AND LET THE INFORMATION FLOW INTO THE ORGANIZATION THROUGH A VARIETY OF MECHANISMS, BUT MAKE SURE THAT WHEREVER THAT INFORMATION IS COMING IN TO, THAT THAT INFORMATION IS LOGGED, THAT INFORMATION IS TRACKED, AND THAT THERE'S AN

ACCOUNTABILITY SYSTEM FOR FOLLOW-UP, BUT, AS I SAID, THE WOMAN WHO HELPED TO RAISE OUR AWARENESS HAD NO PROBLEMS COMING INTO A SCHOOL BOARD MEETING, AND SHE DIDN'T CARE WHO WAS GOING TO BE TEACHING HER CHILD THE NEXT YEAR. SHE FELT LIKE THIS WAS A GRAVE INJUSTICE THAT NEEDED PUBLIC ACCOUNTABILITY.

SO WE DON'T MAKE A VALUE JUDGMENT ABOUT HOW THE INFORMATION COMES IN. OUR OBLIGATION IS JUST TO FOLLOW UP ON ALL INFORMATION.

**BOB WILSON:** I RECOGNIZE VERY CLEARLY THE DANGERS, AS ANYONE WOULD, TO ANONYMOUS TIPS. YOU CAN HAVE YOUR MORTAL ENEMY CALLING IN JUST TO TRY TO CAUSE YOU PROBLEMS, BUT NOT EVERYBODY IS GOING TO HAVE THE COURAGE THAT IT MIGHT TAKE TO STAND UP AND PUBLICLY PROCLAIM TO THE PRINCIPAL OR THE SUPERINTENDENT OR BOARD OF EDUCATION THAT, "I THINK THIS PROBLEM EXISTS." YET THEIR CONCERNS ARE ACTUALLY VALID, AND THEY NEED TO HAVE AN OUTLET AND A WAY TO EXPRESS THEM. IT MAY BE BECAUSE OF PERSONAL RELATIONSHIPS. "THE TEACHER I WANT TO COMPLAIN ABOUT NOT ONLY TEACHES MY DAUGHTER, BUT LIVES TWO HOUSES DOWN FROM ME, AND MY HUSBAND PLAYS GOLF WITH HER HUSBAND, BUT SOMEBODY NEEDS TO LOOK AT THIS."

SO I THINK IF YOU DON'T HAVE THOSE OUTLETS, THAT DOESN'T MEAN YOU DON'T HAVE A SIFTER AND YOU DON'T TRY TO PROFESSIONALLY ASSESS THIS INFORMATION, BUT I THINK YOU HAVE TO HAVE A WAY THROUGH TIP LINES

AND SO FORTH FOR PEOPLE TO BE ABLE TO CALL AND PROVIDE INFORMATION, EVEN ANONYMOUSLY. I THINK IT'S A MISTAKE IF YOU DON'T.

**TISHA EDWARDS:** I WOULD ALSO ARGUE THAT IT IS GOOD PRACTICE FOR THE ORGANIZATION TO BE TRANSPARENT WITH ITS EMPLOYEES THAT NOT ONLY ARE WE WAITING ON SOMEONE TO GIVE US A TIP, BUT WE'RE DOING STATISTICAL ANALYSIS ON THE BACK END TO FIND THOSE INCONSISTENCIES AND THOSE ANOMALIES.

FOLKS IN OUR SCHOOL SYSTEM KNOW, OF COURSE, WHEN PEOPLE BRING INFORMATION FORWARD, WE WELCOME THAT, BUT THEY ALSO KNOW THAT WE HAVE A SET OF RESOURCES THAT ARE LOOKING FOR DATA, LOOKING AT DATA, AND WILL QUESTION THAT DATA IF THERE ARE THOSE ANOMALIES THAT EXIST, WHICH HAS BEEN THE CASE FOR US.

EVEN WHEN THAT SCHOOL COMMUNITY IS VERY TIGHT-LIPPED AND IT IS IMPOSSIBLE TO GET PEOPLE TO TALK TO YOU ABOUT WHAT'S HAPPENING, WE HAVE PARTICULAR FLAGS OR ALERTS AT THE STATE LEVEL AND AT THE SCHOOL SYSTEM LEVEL WHICH WE USE TO HELP US TO THINK THROUGH WHERE THERE MIGHT BE POSSIBILITIES OF INCONSISTENCIES AND IRREGULARITIES.

**BOB WILSON:** LET ME THROW ONE OTHER THING IN THIS. GO BACK TO THAT STATEMENT THAT'S BEEN MADE A NUMBER OF TIMES, AND I DO BELIEVE IT

STARTS WITH THE LEADERSHIP OF THE SYSTEM, THE YOU MUST PROMOTE A CULTURE OF INTEGRITY IN THAT SYSTEM BECAUSE PEOPLE WON'T EVEN CALL IN WITH ANONYMOUS TIPS IF THEY BELIEVE NOTHING WILL HAPPEN TO IT.

WE HAD A SCHOOL IN OUR A.P.S. SYSTEM THAT THE FOURTH GRADE ONE YEAR WAS 813<sup>TH</sup> IN ITS MATH SCORES OUT OF 1,200 SCHOOLS IN THE STATE. THE NEXT YEAR, IT WAS NUMBER ONE. NOW, SOME OF US ARE OLD ENOUGH TO REMEMBER THE METS' LAST-TO-FIRST, BUT THEY ONLY CLIMBED OVER 7 TEAMS. THIS GROUP WENT FROM 813<sup>TH</sup> TO FIRST, AND A.P.S., WHEN THE NEWSPAPER QUESTIONED IT, SAID, "OH, NO. IT IS ALL DUE TO OUR INCREDIBLE PROGRAMS AND TEACHING." WELL, IF THEY WERE THAT DARN GOOD, I WANT THEM FOR EVERYBODY ELSE.

THAT SAME YEAR, ANOTHER SCHOOL THAT WAS IN THE LOWEST GROUP THE YEAR BEFORE IN THIRD GRADE IN MATH, THOSE SAME STUDENTS NOW IN THE FOURTH GRADE WERE FOURTH IN THE STATE. A.P.S. DEFENDED IT.

THE MESSAGE IS, YOU CAN HIT ME ALL YOU WANT WITH FACTS AND TRUTH, BUT WE ARE DOING IT RIGHT, AND WE ARE GOING TO IGNORE THE OBVIOUS, THAT SOMETHING IS WRONG. PEOPLE WON'T EVEN CALL IN WITH ANONYMOUS CALLS THEN BECAUSE THEY WON'T DO ANYTHING ABOUT IT. IF YOU FOSTER THAT CULTURE AND YOU ACTUALLY DO SOMETHING ABOUT IT, AS WE'RE HEARING ABOUT IN BALTIMORE, I THINK YOU WILL GET PEOPLE MORE WILLING TO TAKE WHATEVER AVENUES ARE OUT THERE--WHETHER IT'S

WALKING INTO THE PRINCIPAL'S OFFICE, SPEAKING AT A BOARD MEETING, WRITING THE SUPERINTENDENT, OR JUST GETTING ENOUGH COURAGE TO CALL THE ANONYMOUS LINE--YOU WILL GET MORE INFORMATION ABOUT YOUR SYSTEM.

**JACK BUCKLEY:** I THINK WE'VE GOT A QUESTION FROM THE ONLINE AUDIENCE.

**JAMES ELIAS:** WE DO. SO THIS WAS ACTUALLY PROMPTED BY STEVE FERRARA'S DESCRIPTION OF GROUPS OF PEOPLE. THIS QUESTIONER WANTS TO KNOW, "IS IT ETHICAL TO WORK COUNTER TO A SYSTEM THAT HAS LOST CREDIBILITY IN TERMS OF ITS GOALS?"

**STEVE FERRARA:** I'M SORRY. COULD YOU REPEAT THAT? I'M GLAD THAT'S YOUR QUESTION.

**JAMES ELIAS:** BEYOND YOUR GROUP OF PEOPLE, IS IT ETHICAL TO WORK COUNTER TO A SYSTEM THAT HAS LOST CREDIBILITY IN TERMS OF ITS GOALS?

**STEVE FERRARA:** YEAH. YOU KNOW, I THINK WE NEED TO WRITE TO THE ETHICIST COLUMN IN THE "NEW YORK TIMES" MAGAZINE. I DON'T THINK THAT IS THE KIND OF QUESTION I SHOULD DISCUSS IN A FORUM LIKE THIS.

**JAMES ELIAS:** SEVERAL PEOPLE FROM STATE DEPARTMENTS OF EDUCATION HAVE SAID THAT IF YOU RECEIVED AN INVESTIGATION RESPONSE OR TIP THAT IS VAGUE OR DOES NOT ADDRESS A VIOLATION, WHAT IS THE APPROPRIATE FOLLOW-UP ACTION?

**JACK BUCKLEY:** I HEARD IT TO BE, FOLKS FROM S.E.A.s ARE WONDERING, IF WE RECEIVE A TIP THAT IS VAGUE AND DOESN'T SORT OF PROVIDE ENOUGH INFORMATION, WHAT IS THE APPROPRIATE FOLLOW-UP ACTION?

SO WHAT DO WE DO WITH A VAGUE TIP?

**LOU FABRIZIO:** THAT I CAN SPEAK TO.

**JACK BUCKLEY:** SURE. PLEASE.

**LOU FABRIZIO:** ANYTIME WE'VE EVER GOTTEN ANY KIND OF ANONYMOUS TIP OR SOME LETTER WITH NO RETURN ADDRESS THAT SAYS SOMETHING IS HAPPENING IN SUCH AND SUCH A SCHOOL SYSTEM, WE IMMEDIATELY CALL THE CENTRAL OFFICE TEST COORDINATOR OF THAT SCHOOL DISTRICT AND SAY, "HEY, LOOK. WE'VE JUST GOTTEN THIS LETTER. WE DON'T KNOW IF IT MEANS ANYTHING, BUT WE NEED FOR YOU TO CHECK AROUND, SEE IF THERE IS SOMETHING GOING ON OR IF YOU ARE HEARING ANYTHING," AND THAT KIND OF THING, AND THEN WE LET THEM GO AND FOLLOW IT.

**STEVE FERRARA:** YEAH. BOB SAID IT. YOU'VE HEARD IT A COUPLE OF TIMES NOW. I THINK YOU HAVE TO FOLLOW UP ON EVERYTHING.

**BOB WILSON:** BUT IF YOUR QUESTION IS, CAN IT BE SO VAGUE THAT YOU CAN'T HARDLY FOLLOW UP, CERTAINLY IT CAN BE. YOU DO THE BEST YOU CAN. THAT IS ALL YOU CAN DO. IF YOU HAVE NO RETURN ADDRESS, NO RETURN PHONE NUMBER, NO NAME, AND IT IS SO AMBIGUOUS SO YOU PASS IT AROUND AMONG A BUNCH OF OTHER FOLKS TO BE SURE YOU ARE NOT THE ONLY IDIOT IN THE ROOM WHO CAN'T UNDERSTAND IT AND EVERYBODY SAYS, "I DON'T KNOW WHAT THAT MEANS. I DON'T KNOW WHAT YOU CAN DO. WELL, WHAT CAN WE DO WITH IT?" THEN I THINK YOU'VE PROBABLY DONE ALL YOU CAN. YOU TRIED, BUT SOMETIMES THERE WILL BE ENOUGH OF A CRACK IN THERE THAT YOU CAN FIGURE OUT, JUST LIKE YOU SAID, IT JUST SAYS SOMETHING IS GOING ON. THAT IS WHAT LOU SAYS.

WELL, OK. YOU AT LEAST NEED TO TALK TO THE DIRECTOR OF THAT DIVISION OR WHATEVER AND SAY, "LOOK. WE GOT THIS LETTER, AND I DON'T KNOW WHAT IS SUPPOSED TO BE GOING ON, BUT DO YOU HAVE SOME REASON?" AND THEY MAY SAY, "WELL, WE HAVE HAD A LITTLE PROBLEM RECENTLY THAT I HAVEN'T TOLD ANYBODY ABOUT."

WELL, MAYBE THAT'S IT. YOU NEVER KNOW. SO YOU JUST DO THE BEST YOU CAN. I THINK THAT'S ALL YOU CAN DO. I DON'T THINK YOU IGNORE IT. YOU JUST DO THE BEST YOU CAN.

**JACK BUCKLEY:** TONY, DO YOU HAVE A QUESTION FROM BEFORE?

**TONY ALPERT:** YEAH. LOU, YOU MIGHT BE ABLE TO TALK A LITTLE BIT ABOUT THIS. THE DIFFERENT STRATEGIES THAT YOU MIGHT NEED TO IMPLEMENT FOR A SMALL, RURAL SCHOOL VERSUS A LARGE URBAN.

**LOU FABRIZIO:** WELL, IF YOU ARE TALKING ABOUT THINGS LIKE HAVING PROCTORS AND--

**TONY ALPERT:** NO, IN TERMS OF INVESTIGATING IMPROPRIETIES IN WHAT IS OTHERWISE A SMALL, CLOSE-KNIT COMMUNITY.

**LOU FABRIZIO:** WELL, I DON'T THINK WE HAVE AS SMALL A SET OF SCHOOLS AS YOU MAY HAVE IN OREGON, BUT WE ALWAYS HAVE THE ASSISTANCE THAT THE STATE WOULD PROVIDE. OUR SMALLEST SCHOOL DISTRICT IN OUR STATE IS, LIKE, 700 STUDENTS, SO--

**TISHA EDWARDS:** BUT I THINK ONE OF THE THINGS THAT WAS ALLUDED TO IS ONE OF THE THINGS YOU MIGHT CONSIDER, IS PARTNERING WITH ANOTHER SCHOOL DISTRICT IN THE STATE AND HAVING SOME TYPE OF RECIPROCAL RELATIONSHIP AROUND. YOU COULD HAVE MULTIPLE SMALL L.E.A.s FORMING A CONSORTIUM OR CLUSTER OF SORTS THAT, WHEN ISSUES ARRIVE, THERE IS

A RELATIONSHIP THAT ALLOWS YOU TO DO INVESTIGATIONS OF ONE ANOTHER'S SCHOOLS, IF NECESSARY.

**JACK BUCKLEY:** MAKES SENSE. JOHN?

**JOHN FREMER:** NOT A QUESTION, BUT AN OBSERVATION THAT GOES BACK TO TIP LINES. SO I HAVE BEEN TALKING TO AGENCIES THAT ARE RESPONSIBLE FOR TESTS--STATES, DISTRICTS, BUT OTHER PLACES, TOO--AND I HAVE PROBABLY HAD 50 DIFFERENT CONVERSATIONS WITH THE PERSONS RESPONSIBLE FOR THE PROGRAM IN CONVERSATIONS, THEIR REVIEWS OF THEIR PROCESSES AND RECOMMENDATIONS, AND WE ALWAYS RECOMMEND A TIP LINE IF THEY DON'T HAVE A TIP LINE, AND THEN WE GET FOLLOW-UP LATER ON, AND IF THEY HAVE A TIP LINE, THEY SHARE THEIR EXPERIENCES, AND THE THING THAT WAYNE CAMARA SAID ABOUT GETTING A LOT OF FALSE POSITIVES, IN GENERAL, THAT IS NOT WHAT IS HAPPENING.

PLACES THAT HAVE HAD TIP LINES ALL ALONG OR HAVE INTRODUCED THEM, AT LEAST THEY ARE NOT GETTING CRANK INQUIRIES. THEY ARE GETTING INQUIRIES THAT, UPON INVESTIGATION, DON'T LEAD TO A CONCLUSION OF CHEATING, BUT ALMOST ALWAYS WHEN SOMEONE TAKES THE TROUBLE TO MAKE A TIP--WHETHER IT'S BY E-MAIL OR PHONE, WHATEVER--THE E-MAIL YOU GIVE YOURSELF AWAY, BUT BY PHONE IT IS WORTH GETTING, AND SO I'M A STRONG ADVOCATE OF TIP LINES. I DON'T THINK YOU HAVE TO WORRY ABOUT WASTING A LOT OF YOUR TIME IF YOU INTRODUCE THEM.

**STEVE FERRARA:** I HAVE A NEW TOPIC. DAVE IN THE EARLIER PANEL MADE THE POINT THAT MUCH OF WHAT WE DO AROUND SECURITY, INTEGRITY, DETECTION, AND SO FORTH IS IN ITS INFANCY, AND I AGREE WITH THAT IN LARGE MEASURE. I HAVE TRIED TO MAKE THE POINT THAT THE DEVELOPMENT OF OUR PROCEDURES IN INVESTIGATIONS, I WOULD SAY, THAT IS VERY MUCH IN ITS INFANCY, OR MAYBE WE HAVE NOT EVEN GOTTEN TO BIRTH YET. THERE IS ANOTHER POINT RELATED TO THIS NOTION OF WE ARE IN OUR INFANCY. SO WE HAVE THESE STATISTICAL DETECTION PROCEDURES FOR MULTIPLE CHOICE ITEMS.

NOW WE HAVE THE COMMON CORE STATE STANDARDS, THE DESIGNS OF PARCC AND SMARTER BALANCED CONSORTIUM, AND WE ARE NOW GOING TO SEE MORE CONSTRUCTED RESPONSE ITEMS AND PERFORMANCE TASKS. I'M SO HAPPY TO SEE ALL THAT STUFF. I WORK IN THE CENTER FOR PERFORMANCE ASSESSMENT. I DID THAT IN THE NINETIES. I'M GLAD THAT IT IS BACK. SO WE REALLY NEED TO START THINKING ABOUT WHAT ARE THE DETECTION PROCEDURES WE ARE GOING TO USE FOR THE VARIOUS FORMS OF ANOMALIES, CHEATING AND OTHER ANOMALIES THAT CAN GO ON IN PERFORMANCE ASSESSMENTS, AND THEN WHAT DO WE DO ABOUT INVESTIGATION.

SO ONE AREA FOR CONSIDERATION IS--SOMEONE REFERRED THIS MORNING TO THESE ONLINE PROGRAMS SO TEACHERS CAN CHECK TO SEE IF KIDS ARE PLAGIARIZING WHEN THEY TURN IN WRITTEN ASSIGNMENTS. THERE MAY BE A

WAY TO ADAPT THOSE PROGRAMS TO LOOK FOR THINGS LIKE THE STORY THAT I TOLD WHERE MULTIPLE KIDS IN THE SAME TESTING GROUP HAVE HIGHLY SIMILAR RESPONSES TO THEIR PERFORMANCE TASKS OR ESSAYS, CONSTRUCTED RESPONSE ITEMS, AND SO FORTH. SO I THINK WE REALLY NEED TO START DEVELOPMENTAL WORK IN THAT DIRECTION.

**BOB WILSON:** STEVE, THE REASON, IN MY OPINION, THAT WE'RE SOMEWHAT IN THE INFANCY IN INVESTIGATIVE TECHNIQUES WITH SCHOOLS, I'M NOT GOING TO SAY TEACHERS DIDN'T CHEAT BEFORE 15 YEARS AGO, BUT UP UNTIL ABOUT 15 YEARS AGO WHEN WE WERE TALKING ABOUT CHEATING IN SCHOOL, WE WERE ALMOST ALWAYS, ALMOST ALWAYS, TALKING ABOUT STUDENTS. IN THE LAST 15 YEARS, THE EQUATION HAS CHANGED BECAUSE THE EQUATION ON THE TEST HAVE CHANGED. UP UNTIL ABOUT 15 YEARS AGO, THE TESTS WERE ALL ABOUT THE STUDENTS.

NOW, I'M NOT GOING TO SIT HERE AND BASH NO CHILD LEFT BEHIND SIMPLY BECAUSE IT SHIFTED THE EQUATION. TEACHERS HAVE TO BE RESPONSIBLE AND HAVE TO BE ACCOUNTABLE, BUT THE SHIFT IS SO GREAT THAT NOW THE TEST, IN THIS INVESTIGATOR'S HUMBLE OPINION, BECAME MORE ABOUT THE TEACHERS AND THE SCHOOLS AND THE PRINCIPALS THAN IT DID ABOUT THE STUDENTS.

SO YOU BROUGHT IN ANOTHER WHOLE UNIT OF FOLKS, AND NOW WE ARE NOT TALKING ABOUT ADULTS TO STUDENTS. WE ARE TALKING ABOUT ADULTS TO

ADULTS. THAT REQUIRES REAL INVESTIGATION, AND THE EQUATION HAS CHANGED.

**JACK BUCKLEY:** GOOD POINT. CORNELIA?

**CORNELIA ORR:** IT HAS BEEN ALLUDED TO THROUGHOUT DISCUSSIONS ALL DAY TODAY ABOUT RESOURCES, AND I THINK IT IS A HUGE ISSUE FOR STATES THAT ARE TRYING TO OPERATE TWO PROGRAMS--A STATE ASSESSMENT, MAYBE A CONSORTIUM-LED ASSESSMENT--TO THINK THEY HAVE ENOUGH RESOURCES, EITHER FISCAL RESOURCES TO CONTRACT OR STAFF RESOURCES, TO DO THIS.

COULD SOME OF YOU ADDRESS WHAT YOU THINK IT TAKES TO BE REASONABLY VIGILANT ON THIS IN TERMS OF RESOURCES?

**BOB WILSON:** PREVENTION IS CHEAPER.

**CORNELIA ORR:** RIGHT. YEAH.

**BOB WILSON:** I CAN TELL YOU THAT BECAUSE I HAVE BEEN ON THE OTHER END OF CAUSING THE BILL TO GO UP. SO THE MORE MONEY YOU CAN PUT INTO PREVENTION AND CREATING THAT CULTURE THAT SAYS, "WE ARE GOING TO PLAY BY THE RULES, AND WE ARE GOING TO DO IT RIGHT," IS A LOT LESS EXPENSIVE THAN AFTER THE FACT. IT IS ALWAYS MORE EXPENSIVE AFTER THE FACT.

**TISHA EDWARDS:** RIGHT. I WOULD JUST REITERATE FROM BALTIMORE CITY'S PERSPECTIVE, IN THE LAST TWO YEARS, WE'VE MADE A SIZABLE INVESTMENT IN PREVENTION AND THIS ISSUE OF THE CULTURE, CHANGING THE CULTURE.

I'M A FORMER PRINCIPAL, AND WHEN MEETING WITH OUR SUPERINTENDENT ABOUT OUR STRATEGY AFTER THIS PARTICULAR INCIDENT--WHICH WAS VERY, VERY HARD FOR THE ORGANIZATION--ONE OF THE THINGS I SAID TO OUR TEAM, OUR LEGAL TEAM, AND TO OUR SUPERINTENDENT IS, "WE SHOULD NOT BE TALKING ABOUT THIS FROM THE PERSPECTIVE OF CATCHING CHEATERS. I THINK THAT THE ORGANIZATION IS NOT GOING TO FIND THAT A PALPABLE MISSION. WE SHOULD BE TALKING ABOUT PROTECTING EDUCATORS AND PROTECTING CHILDREN," AND SO IT SHOULD BE THE FRAME AROUND TESTING INTEGRITY, AND THE CONVERSATION IN OUR DISTRICT HAS SHIFTED FROM, "WE ARE LOOKING FOR THE CHEATER," TO, "WE ARE PROTECTING OUR EDUCATORS, AND WE'RE PROTECTING THE RESULTS OF CHILDREN,"

AND SO I THINK IT MAKES A HUGE DIFFERENCE ABOUT HOW PEOPLE EMBRACE PROCTORS. IT MAKES A HUGE DIFFERENCE IN TERMS OF THE WILLINGNESS OF THE ORGANIZATION TO EVEN MAKE THE INVESTMENT WHEN IT IS NOT ABOUT GOING AFTER THE BAD PEOPLE BUT IT IS ACTUALLY ABOUT PROTECTING PEOPLE WHO COME AND WORK HARD FOR CHILDREN EVERY DAY. SO THAT CULTURAL SHIFT IN MAKING THAT ORGANIZATIONAL KIND OF PARADIGM

CHANGE WAS HUGE FOR OUR SCHOOL SYSTEM, AND SO WE DON'T TALK ABOUT IT.

YES, WE HAVE CHEATING INCIDENTS, AND WE ARE GOING TO GO AFTER THOSE INDIVIDUALS, BUT WE BELIEVE THAT THE MAJORITY OF OUR TEACHERS AND THE MAJORITY OF OUR EDUCATORS ARE TRYING TO DO THE RIGHT THING. THEY MAKE BAD DECISIONS, BUT THEY ARE TRYING TO DO THE RIGHT THING, AND SO OUR TESTING INTEGRITY SYSTEM IS ABOUT PROTECTING THE INTEGRITY OF THE SYSTEM AND PROTECTING PEOPLE SO THAT THERE ARE NOT ALLEGATIONS, SO THAT WE CAN STAND BEHIND THE RESULTS OF CHILDREN.

SO I THINK THAT CULTURAL SHIFT THAT STEVE TALKED ABOUT IN HIS PRESENTATION IS SOMETHING THAT MADE A HUGE DIFFERENCE IN OUR ORGANIZATION.

**STEVE FERRARA:** CORNELIA, I THINK THERE ARE AT LEAST TWO WAYS OF THINKING ABOUT RESOURCES. SORRY ABOUT THAT. ONE IS JUST PERSONNEL RESOURCES. THE OTHER IS EXPERTISE RESOURCES. THAT WAS PART OF MY COMMENTS. WE'VE HEARD TERMS LIKE "LOCAL CONSORTIA," "SHARED RESOURCES," "SHARED RESPONSIBILITY." I THINK, FIRST OF ALL, STATES AND LOCAL SCHOOL SYSTEMS HAVE TO WORK TOGETHER IN THE STATE ASSESSMENT PROGRAMS. IF THERE ARE OTHER WAYS OF SHARING EXPERTISE--IT WAS AN IDEA THAT LOU PROPOSED--THAT IS ANOTHER WAY OF ADDRESSING THE LIMITED-RESOURCES ISSUE, I THINK.

**JACK BUCKLEY:** QUESTION?

**AUDIENCE MEMBER:** GOOD AFTERNOON. MY NAME IS CHRISTINA SAMUELS. I'M A REPORTER WITH "EDUCATION WEEK."

I HAD A QUESTION THAT REFLECTED SOMETHING THAT CHANCELLOR HENDERSON WAS TALKING ABOUT IN THE MORNING. IT SEEMED LIKE ONE OF HER CONCERNS THAT SHE MENTIONED WAS ABOUT, I GUESS, A LACK OF STANDARDIZATION INTO WHAT WOULD TRIGGER A INVESTIGATION AND THEN ALSO HOW AN INVESTIGATION SHOULD BE DONE SO THAT IT DIDN'T FACE PEOPLE FROM THE PRESS OR FROM THE PUBLIC SAYING THAT THE DISTRICT DIDN'T DO ENOUGH OR THAT THE SCHOOL SYSTEM DIDN'T DO ENOUGH TO INVESTIGATE THE ALLEGATIONS, AND I WONDER, AS PEOPLE WHO HAVE BEEN INVOLVED IN INVESTIGATIONS, WHETHER YOU THINK THAT IT IS NECESSARY TO HAVE SOME KIND OF STANDARD SET OF TRIGGERS FOR, "HEY, THIS LOOKS STRANGE," AND THEN ALSO WHAT SHOULD HAPPEN NEXT BECAUSE--WHEN IS IT OK FOR IT TO STAY AT THE SCHOOL? WHEN DOES IT HAVE TO GO OUTSIDE THE SCHOOL, ISSUES LIKE THAT? THANKS.

**STEVE FERRARA:** I HAVE A COMMENT THERE. YOU ASKED, IS IT NECESSARY TO HAVE STANDARDS? I'M NOT SURE WHETHER IT IS NECESSARY OR NOT. I THINK IT WOULD BE EXTREMELY HELPFUL.

THE THING THAT WORRIES ME IS--I'M REALLY PLEASED THAT NCES TOOK ON THIS COMPLICATED AND VERY SENSITIVE ISSUE AND BROUGHT US TOGETHER. WHAT MY WORRY IS, WE ARE GOING TO WALK AWAY FROM HERE WITH A LOT OF NEW WAYS OF THINKING ABOUT THIS, AND I DON'T THINK THERE'S ONE NATURAL, AUTOMATIC CONSTITUENCY THAT IS RESPONSIBLE FOR ALL THE PIECES OF ALL THIS.

SO, FOR EXAMPLE, WE ARE ON THE INVESTIGATION PANEL. I HAVE A LITTLE TROUBLE THINKING ABOUT WHAT IS THE EDUCATIONAL CONSTITUENCY OUT THERE WHO WOULD FEEL COMPELLED PROFESSIONALLY TO DEVELOP OTHER MEETINGS LIKE THIS TO DEVELOP STANDARDS FOR INVESTIGATIONS. MAYBE IT IS THE LEGAL COMMUNITY, BUT, I MEAN, I DON'T KNOW ABOUT THAT. I'M NOT SURE THE RESEARCH COMMUNITY THAT MANY OF US ARE IN IS GOING TO DO THAT, MEMBERS OF NCME, BECAUSE IT IS NOT PART OF RESEARCH.

SO SOME OVERARCHING GROUP--AND I GUESS I WON'T NAME NAMES BECAUSE I DON'T WANT TO PUT PEOPLE ON THE SPOT--BUT THERE ARE ORGANIZATIONS LIKE NCES WHO ARE WORKING ACROSS STATES, ACROSS ASSESSMENT PROGRAMS WHO SHOULD CONTINUE TO PUSH FORWARD, IF IT IS NOT STANDARDS, AT LEAST GUIDELINES THAT WOULD BENEFIT STATES AND LOCAL SCHOOLS AND SCHOOL SYSTEMS AND SO FORTH.

**Lou Fabrizio:** ONE OF THE THOUGHTS THAT COMES TO MY MIND BASED ON HEARING ABOUT REPORTERS, YOU KNOW, WHEN YOU ARE IN YOUR OFFICE AND

YOU GET SOMETHING IN AN E-MAIL AND IT BLURTS OUT SOMETHING ABOUT SOMETHING HAPPENING IN ATLANTA OR SOMETHING HAPPENING IN BALTIMORE, IT REALLY MAKES YOU STOP AND SAY, "YOU KNOW, THAT MAY BE HAPPENING SOMEWHERE IN OUR STATE," AND I KNOW SECRETARY DUNCAN SENT OUT A LETTER TO ALL THE STATE SUPERINTENDENTS TELLING THEM THAT WE NEEDED TO BE THINKING ABOUT TESTING INTEGRITY AND EVERYTHING ELSE.

I IMMEDIATELY DRAFTED SOMETHING FOR OUR STATE SUPERINTENDENT TO SEND TO ALL OF THE LOCAL SCHOOL SUPERINTENDENTS, COPIED IT TO ALL OF THE L.E.A. TEST COORDINATORS, SAYING, "FOLKS, THIS IS WHY WE HAVE THE CODE OF ETHICS. THIS IS WHY WE HAVE THE PROCTORS IN THE CLASSROOM," TRYING TO MAKE SURE THAT WHEN THEY GET CALLS FROM REPORTERS--BECAUSE EVERY TIME ONE OF THOSE STORIES HITS, OUR PHONE STARTS TO RING OFF THE HOOK BECAUSE THEY WANT TO KNOW WHAT IS HAPPENING AND, "WHAT ARE YOU DOING?" AND SO WE USE ALL OF THOSE INSTANCES AS TEACHABLE MOMENTS, AND I LIKE TO ALWAYS START OFF EVERY CONVERSATION WITH A REPORTER WHO IS ASKING ABOUT IT FOR ME TO SAY, "YOU KNOW, LOOK. THERE ARE LAWYERS THAT SOMETIMES DON'T DO THE RIGHT THING. THERE ARE DOCTORS THAT SOMETIMES DON'T DO THE RIGHT THING, AND THERE ARE TEACHERS WHO SOMETIMES MAY NOT DO THE RIGHT THING, BUT IT IS A SMALL NUMBER, AND WE DON'T WANT THAT TO PUT A SHADOW OVER THE GREAT WORK THAT TEACHERS AND EDUCATORS ARE DOING," AND SO I'M ALWAYS TRYING TO REFRAME IT BECAUSE THE STORY IS ALWAYS THESE HORRIBLE THINGS ARE HAPPENING AND IT IS PROBABLY HAPPENING

EVERYWHERE, AND SO YOU JUST HAVE TO DO THAT, OR ELSE YOU END UP GETTING MORE HATE MAIL BECAUSE THEY SAY, "WELL, HOW COME YOU DIDN'T TALK ABOUT THE GOOD TEACHERS?" AND SO I THINK IT IS JUST SOMETHING WE ALL NEED TO BE VERY SENSITIVE ABOUT.

**BOB WILSON:** THE STANDARD OF WHETHER IT IS SOMETHING THAT SHOULD STAY INSIDE THE SYSTEM OR YOU SHOULD GET OUTSIDE COUNSEL OR INVESTIGATORS TO HELP YOU IS A LITTLE TOUGHER LINE TO DRAW, BUT YOU ARE GOING TO HAVE A TEST PROTOCOL. YOU ARE GOING TO HAVE STATE LAWS. YOU ARE GOING TO HAVE CODES OF ETHICS. IF WHAT IS REPORTED OR SUSPECTED MIGHT POSSIBLY BE A VIOLATION OF ANY OF THOSE, THEN IT IS SUBJECT TO INVESTIGATION.

NOW, YOUR CODE OF ETHICS IN ONE STATE MAY BE A LITTLE DIFFERENT THAN SOMEBODY ELSE'S IN ANOTHER STATE. YOUR STATE STATUTES MAY BE DIFFERENT. YOUR TEST PROTOCOLS MAY BE A LITTLE DIFFERENT, BUT IF YOU COULD FIND, WITHIN REASON, THAT WHAT HAS BEEN REPORTED MIGHT VIOLATE ANY OF THOSE, THEN THAT IS SOMETHING TO BE INVESTIGATED,

AND I WILL TELL YOU--I WANT TO GO TO SOMETHING THAT LOU SAID--ONE OF THE SAD THINGS ABOUT ATLANTA, EVEN THOUGH WE SAID THAT IT WAS MAYBE TWO TO 3 TIMES THE NUMBER WE REPORTED, THERE WERE THOUSANDS OF ADMINISTRATORS AND TEACHERS WHO WERE DOING THEIR JOB, DOING IT

RIGHT, DOING IT HONESTLY, AND THEY HAVE SUFFERED GREATLY BECAUSE OF THIS, AND THEY DIDN'T DO A DADBURN THING WRONG.

SO IF THE SYSTEM IS VIGILANT, YOU ARE GOING TO HAVE WHAT LOU SAYS. YOU ARE ALWAYS GOING TO HAVE LAWYERS AND DOCTORS AND FOLKS WHO DO SOME THINGS WRONG, AND THE PUBLIC UNDERSTANDS THAT. WHAT THEY DON'T UNDERSTAND IS WHEN YOU--AT A LEADERSHIP LEVEL, WHETHER YOU ARE THE BOARD OR THE SUPERINTENDENT--ALLOW IT TO GET TO SUCH A MASSIVE NUMBER.

YOU ARE FAILING TO LEAD THE RIGHT WAY SOMEHOW. THEY WILL UNDERSTAND IF YOU HAVE A TEACHER OR TWO MAKE MISTAKES AND CHEAT. YOU GO IN. YOU DO THE BEST YOU CAN WITH IT, AND ROOT IT OUT. THAT IS NOT GOING TO CONDEMN THE WHOLE SYSTEM, BUT WHEN YOU END UP WITH WHAT HAPPENED IN ATLANTA, THE MAJORITY OF THE TEACHERS DID NOTHING WRONG, AND THEY ARE SUFFERING ALONG WITH THAT MINORITY THAT CREATED THE PROBLEM.

SO YOU WANT TO NIP IT IN THE BUD, AND IF IT COULD POSSIBLY BE A VIOLATION OF ANY OF THOSE THINGS, YOU HAVE GOT TO LOOK INTO IT.

**AUDIENCE MEMBER:** HI. MY NAME IS HEATHER GARAM. I ACTUALLY WORK IN CERTIFICATION, BUT I CAME TODAY BECAUSE I THINK THERE ARE A LOT OF PARALLELS YOU CAN DRAW BECAUSE THIS IS ASSESSMENT, WHETHER IT'S IN

SCHOOLS OR CERTIFICATION, AND ALL DAY, WE HAVE BEEN HEARING THIS TALK OF INTEGRITY AND A CULTURE OF HONESTY.

I FEEL LIKE MY QUESTION ACTUALLY GOES TO A WHOLE OTHER SYMPOSIUM, BUT JUST OFF THE TOP OF YOUR HEAD, HOW WOULD RECOMMEND SCHOOLS CREATE THAT CULTURE OF HONESTY AND INTEGRITY?

**STEVE FERRARA:** I THINK TISHA, WHEN SHE DESCRIBED WHAT HAPPENED IN BALTIMORE CITY AFTER A RATHER BLATANT VIOLATION, I THINK WHAT SHE AND ANDRES ALONSO DID IS A GOOD EXAMPLE OF DEVELOPING THAT CULTURE OF INTEGRITY. IT STARTED WITH THE SUPERINTENDENT, AND IT GOES ALL THE WAY DOWN INTO THE SCHOOLS AND THEN BACK UP AGAIN, I HOPE.

**BOB WILSON:** I THINK IT STARTS WITH LEADERSHIP. STARTS WITH YOUR BOARD OF EDUCATION AND WHO THEY HIRE TO ADMINISTER THAT SYSTEM. IT STARTS WITH THE ASSISTANT SUPERINTENDENTS AND THE PRINCIPALS. IF THE PRINCIPALS, THE SUPERINTENDENT, ASSISTANT SUPERINTENDENTS, AND THE BOARD INSIST ON INTEGRITY, SELF-REPORTING, SELF-REPORTING, PLAY THE RULE OF LIFE LIKE GOLF IS REALLY SUPPOSED TO BE PLAYED.

IT IS ONE OF THE GREAT GAMES OF ALL TIME WITH ITS RULES. LET A PLAYER INADVERTENTLY TOUCH THE SAND IN A SAND TRAP BEFORE HE HITS THE BALL WITH HIS CLUB HEAD, HE WILL SELF-REPORT. HE WILL SELF-REPORT, "I GROUNDED THE CLUB," AND IT WILL COST HIM.

**STEVE FERRARA:** MOST OF THEM.

**BOB WILSON:** MOST OF THEM WILL. WELL, THERE SURE WAS ONE A COUPLE OF YEARS AGO ON THE LAST HOLE THAT COST HIMSELF THE TOURNAMENT BY MOVING A TWIG INADVERTENTLY. THE REF THAT WAS ON THE SIDELINE DID NOT SEE IT. HE REPORTED IT. THEY LOOKED AT IT ON TELEVISION, AND YOU COULD BARELY SEE IT MOVE. COST HIM THE TOURNAMENT.

THERE HAS IT BE THIS SENSE OF NOT ONLY ACCOUNTABILITY FOR THE SYSTEM, BUT ACCOUNTABILITY FOR ME, AND IF I'M ABOUT TO READ A STORY TO MY STUDENTS BEFORE THE TEST THAT MIGHT ACTUALLY SHADE THE TEST A LITTLE BIT, MAYBE I NEED TO STEP TO MY PRINCIPAL AND SAY, "CAN I READ THIS?" OR, "SHOULD I HAVE READ THIS?" SO IT IS A CULTURE THAT IS DEVELOPED BY LEADERSHIP.

I ABSOLUTELY AM CONVINCED THAT LEADERS OF ANY ORGANIZATION HAVE A PERMEATING EFFECT THROUGHOUT THAT SYSTEM, AND IT STARTS THERE, AND IT ENDS THERE. THEY HAVE TO DO THAT.

**TISHA EDWARDS:** YEAH. I WOULD REITERATE THAT TWO YEARS AGO, WE ONLY HAD MONITORS IN 13 SCHOOLS, AND THESE WERE SCHOOLS THAT WE HAD ALLEGATIONS OR SUSPICIONS OF CHEATING.

LAST YEAR, WE HAD MONITORS IN EVERY SINGLE SCHOOL IN THE DISTRICT, AND WE DID THE WIDESPREAD TRAINING AND LOTS OF TALKING WITH PEOPLE ABOUT WHY THIS ISSUE WAS SO IMPORTANT IN TERMS OF PROTECTING EDUCATORS AND PROTECTING THE INTEGRITY OF OUR SCHOOL SYSTEM.

I WAS ONE OF THOSE PEOPLE WHO WAS ON CALL THE DAY THAT WE DID TESTING IN OUR SCHOOL SYSTEM LAST YEAR WHEN WE IMPLEMENTED THIS MORE RIGOROUS TEST MONITORING PROCESS, AND WE WERE VERY, VERY EXCITED THAT OUR PHONES STARTED RINGING OFF THE HOOK DURING TESTING ADMINISTRATION WHEN PEOPLE WERE SAYING, "I THINK I HAVE MADE A MISTAKE. I THINK I HAVE DONE SOMETHING WRONG. CAN YOU DEPLOY SOMEONE HERE SO THAT WE CAN INITIATE CORRECTIVE ACTION AROUND HOW TO ADDRESS THIS BEFORE THE END OF THE TESTING ADMINISTRATION?" AND WE SAW THAT AS EVIDENCE OF A HUGE LEAP FOR OUR DISTRICT AROUND THE FACT THAT PEOPLE WANT TO DO THE RIGHT THING.

YOU HAVE TO HAVE PEOPLE IN PLACE TO MONITOR AND TO SEE WHAT IS HAPPENING, BUT THIS WHOLE ISSUE OF PREVENTION ALLOWS YOU TO CREATE THAT CULTURE OF HONESTY, AND IT CAN BE A SITUATION THAT MOVES FROM BEING VERY TRAUMATIC FOR AN ORGANIZATION TO BEING VERY EMPOWERING FOR TEACHERS TO BE ABLE TO SAY, "THERE IS SOMETHING HAPPENING ACROSS THE HALL." THEY WON'T GO AS FAR AS TO SAY IT IS CHEATING, BUT THEY WILL SAY, "IT'S SOMETHING I THINK SOMEONE NEEDS TO CHECK," OR, "SOMEONE NEEDS HELP, OR, "I NEED HELP."

SO WE SAW THAT HAPPEN VERY QUICKLY WITH ONLY ONE YEAR OF MONITORING. SO I AGREE WITH YOU. THERE IS A LOT OF SELF-REPORTING THAT WILL HAPPEN IF YOU HAVE THE PEOPLE AND THE PROCESSES IN PLACE FOR THAT.

**BOB WILSON:** CONSTANTLY TALK ABOUT IT. IT IS NOT TO BE SWEEPED UNDER THE RUG. IT IS NOT TO BE SAID ONE TIME. INTEGRITY IS SOMETHING TO BE DEALT WITH DAY IN AND DAY OUT BY EXAMPLE AND BY DEEDS AND BY WORDS.

**JACK BUCKLEY:** RIGHT. ALL RIGHT. WELL, WITH THAT, WE NEED TO WRAP UP. I'M SORRY. I'M RUNNING OUT OF TIME.

QUICKLY, THEN, TO SUMMARIZE, WHAT DID WE HEAR OF THIS PANEL? WELL, I THINK THERE IS CONSENSUS THAT SAYS INVESTIGATE EVERYTHING, RIGHT? IF YOU GET A TIP, IF YOU GET A CALL, YOU NEED TO DO THE BEST YOU CAN TO INVESTIGATE IT, NO MATTER HOW VAGUE IT IS.

I THINK THERE IS PERHAPS LESS CONSENSUS ABOUT WHAT LEVEL OF DIFFERENT INVESTIGATION REQUIRES OUTSIDE PERSONNEL AND HOW FAR OUTSIDE IS OUTSIDE, AND I THINK THERE IS A SEVERE RESOURCE CONSTRAINT THERE, TOO, DEPENDING ON YOUR DISTRICT AND YOUR STATE, AND THAT IS AN IMPORTANT QUESTION, I THINK, FOR SDAs AND L.E.A.s TO

CONSIDER AS THEY MOVE TO THE NEXT POINT, WHICH IS TO CODIFY THINGS,  
RIGHT?

SO WHEN YOU ARE WRITING YOUR PROCEDURES, WHETHER OR NOT THEY LOOK  
LIKE MARYLAND'S OR NORTH CAROLINA'S OR SOMEONE ELSE'S, YOU NEED TO  
LAY THESE THINGS OUT AND BE VERY CLEAR ABOUT WHAT FOLKS CAN EXPECT  
WITH RESPECT TO WHO WILL CONDUCT AN INVESTIGATION WHEN AND OF WHAT,  
AND THEN, I GUESS IF I HAD TO PULL ONE OTHER PROBABLY FACT THAT I  
HEARD, IS THAT PREVENTION IS CHEAPER.

SO IF WE HAD TO PICK A PANEL TO BE ON, THE FIRST PANEL WOULD  
PROBABLY BE MORE COST-EFFECTIVE THAN LEADING US HERE, WHERE, BY  
THIS POINT, WE PROBABLY SPENT SOME MONEY ON PREVENTION AND NOT  
ENOUGH AND NOW A GREAT DEAL ON RESOURCES ON INVESTIGATION.

WE HAVE GOT ONE PANEL LEFT, BUT WE'RE ACTUALLY GOING TO TAKE OUR  
BREAK. THIS BREAK IS A BIT LONGER THAN THE OTHER BREAKS. IT IS  
ACTUALLY UNTIL 3:15. THIS IS WHAT, IN WASHINGTON, WE CALL SIESTA  
TIME. SO PLEASE JOIN US BACK HERE AT 3:15 PROMPTLY, AND WE WILL  
CONDUCT OUR LAST PANEL.

PANEL IV: TESTING INTEGRITY PRACTICES AND PROCEDURES FOR ONLINE  
AND COMPUTER-BASED ASSESSMENTS

**JACK BUCKLEY:** WELCOME BACK.

OUR FINAL PANEL DISCUSSION OF THE DAY FOCUSES ON TESTING INTEGRITY  
PRACTICES AND PROCEDURES FOR THE FUTURE--OR IN SOME PLACES THE  
PRESENT--FOR ONLINE AND COMPUTER-BASED ASSESSMENTS.

SO MANY ASSESSMENTS, AS WE ALL KNOW, ARE TRANSITIONING FROM  
TRADITIONAL PAPER AND PENCIL TO ONLINE AND COMPUTER-BASED. AND SO  
WE EXPECT THIS TO BE A PARTICULARLY IMPORTANT AND SALIENT  
DISCUSSION, ALL THE MORE SO BECAUSE WE ARE JOINED BY  
REPRESENTATIVES FROM THE TWO RACE TO THE TOP ASSESSMENT CONSORTIA  
WHO CAN REALLY TALK TO US ABOUT THE FUTURE TODAY.

SO AS WE HAVE HAD ALL DAY, WE HAVE A WONDERFUL AND DYNAMIC PANEL  
ASSEMBLED. TODAY WE ARE VERY PLEASED ON THIS FINAL PANEL TO BE  
JOINED FIRST BY WAYNE CAMARA, WHO IS THE VICE PRESIDENT FOR  
RESEARCH AND DEVELOPMENT AT THE COLLEGE BOARD. RESPONSIBLE FOR  
MANAGING RESEARCH AND ASSESSMENT DEVELOPMENT FOR PROGRAMS  
INCLUDING THE S.A.T. AND THE A.P. HE IS THE IMMEDIATE PAST  
PRESIDENT OF THE NATIONAL

COUNCIL ON MEASUREMENT IN EDUCATION, AND PRESIDENT-ELECT OF AERA'S DIVISION D, MEASUREMENT RESEARCH METHODOLOGY. HIS RESEARCH AND PUBLICATIONS FOCUS ON COLLEGE ADMISSIONS, COLLEGE READINESS, LARGE-SCALE ASSESSMENT, ETHICAL AND PROFESSIONAL ISSUES IN ASSESSMENT AND TEST VALIDATION.

WE'RE ALSO PLEASED TO BE JOINED BY JOHN FREMER, THE FOUNDER OF CAVEON TEST SECURITY--A COMPANY THAT HELPS IMPROVE SECURITY IN TEST DEVELOPMENT, TEST ADMINISTRATION, REPORTING AND SCORE USE, WHICH ASSISTS STATES, SCHOOL DISTRICTS, AND OTHER ENTITIES INVESTIGATING TESTING IRREGULARITIES. JOHN IS A PAST PRESIDENT OF THE ASSOCIATION OF TEST PUBLISHERS AS WELL AS NCME AND THE ASSOCIATION FOR ASSESSMENT IN COUNSELING. IN 2007 HE RECEIVED THE ATP AWARD FOR CONTRIBUTIONS TO MEASUREMENT. AND HE SERVED AS THE EDITOR FOR THE NCME JOURNAL, "EDUCATIONAL MEASUREMENT ISSUES AND PRACTICE" AND IS CO-EDITOR OF "THE HANDBOOK OF TEST SECURITY."

WES BRUCE WAS NAMED CHIEF ASSESSMENT OFFICER FOR THE INDIANA DEPARTMENT OF EDUCATION IN JANUARY OF 2009, AND BEGAN WORK WITH THE DEPARTMENT IN 1999 AS THE DIRECTOR OF THE DIVISION OF SCHOOL ASSESSMENT. HE WAS ALSO APPOINTED BY U.S. SECRETARY OF EDUCATION MARGARET SPELLINGS TO THE FIRST NATIONAL TECHNICAL ADVISORY COUNCIL, AND CURRENTLY SERVES ON THE PARCC LEADERSHIP TEAM AND CHAIRS PARCC'S TECHNOLOGY OPERATIONAL WORKING GROUP.

AND TONY ALPERT, WHO SERVES AS THE CHIEF OPERATING OFFICER FOR THE SMARTER BALANCED ASSESSMENT CONSORTIUM, WHERE HE OVERSEES THE FISCAL OPERATIONS OF SMARTER BALANCED, COLLABORATING WITH WASHINGTON AS THE LEAD FISCAL STATE, AND PROVIDES EXPERT GUIDANCE ON ASSESSMENT DESIGN AND TECHNOLOGY. PRIOR TO JOINING SMARTER BALANCED, TONY SERVED AS DIRECTOR OF ASSESSMENT FOR THE OREGON DEPARTMENT OF EDUCATION, WHERE HE DIRECTED THE ADMINISTRATION OF OREGON'S COMPUTER ADAPTIVE ASSESSMENT SYSTEM. AND, LIKE WES, TONY SERVED ON THE U.S. DEPARTMENT'S NATIONAL TECHNICAL ADVISORY COMMITTEE.

WAYNE, COULD YOU PLEASE START US OFF?

**WAYNE CAMARA:** THANK YOU VERY MUCH, JACK AND EUNICE, FOR INVITING ME.

I WANTED TO COMMEND THE NCES AND THE DEPARTMENT FOR HAVING THIS.

I HAVE CERTAINLY LEARNED A LOT TODAY AND HOPE I CAN PROVIDE SOME COMMENTS THAT ARE SOMEWHAT USEFUL. WHAT I WANTED TO DO IS APOLOGIZE IN ADVANCE.

I'M FROM THE BOSTON AREA. I TALK FAIRLY FAST, BUT NOT TO WORRY BECAUSE I USE AN EXCEEDINGLY SMALL FONT ON MANY OF MY SLIDES. SO IF

YOU DON'T GET CATCH THE WORDS, YOU CERTAINLY WON'T CATCH THE WRITING.

I WANTED TO CAST WHAT WE ARE TALKING ABOUT HERE--CHEATING AND INTEGRITY--WITHIN A VALIDITY FRAMEWORK FIRST OF ALL, BEFORE I BEGIN THE REMARKS. I THINK IT IS IMPORTANT TO THINK OF INTEGRITY AND CHEATING OR THE LACK THEREOF AS AN EXTENSION OF VALIDITY ARGUMENT. CERTAINLY, IT RELATES TO THE APPROPRIATE USE OF TESTS AND TO THE EXTENT THE TESTS ARE USED FOR PURPOSES FOR WHICH THEY HAVE NOT BEEN INTENDED OR HAVE NOT DEVELOPED VALIDATION EVIDENCE. THOSE ISSUES OF INTEGRITY AND OPPORTUNITIES TO CHEAT INCREASE.

THE VALIDATION ARGUMENT REQUIRES US TO SUPPORT INTENDED INTERPRETATIONS OF TEST SCORES. AND SO IF WE'RE ATTEMPTING TO USE TEST SCORES TO EVALUATE EDUCATORS OR SCHOOL PERSONNEL, THEN AS MUCH AS THE INTEGRITY AND POTENTIAL THEFT ARE RISKS, GREATER RISKS TO THAT PROGRAM ARE THE LACK OF VALIDATION EVIDENCE. SO, WITHOUT VALIDATION EVIDENCE, I THINK INTEGRITY AND CHEATING ARE MINOR PROBLEMS. AND WHERE INTEGRITY AND CHEATING ARE MAJOR PROBLEMS, THERE IS PROBABLY INSUFFICIENT VALIDATION EVIDENCE BECAUSE THEY THREATEN THE USE OF THE TEST. SO I THINK THAT'S THE FRAMEWORK I WOULD LIKE US TO BEGIN THINKING OF.

IN TERMS OF THE VALIDATION ARGUMENT, I THINK IT IS ALSO IMPORTANT WHEN YOU HAVE CHEATING, MISUSE, LAX PROCEDURES, WHEN YOU HAVE DIFFERENCES IN ADMINISTRATIVE CONDITIONS, EXTREME AMOUNTS OF FLEXIBILITY BY SCHOOL OR BY TEACHER OR BY ADMINISTRATIVE SETTING, THOSE ARE SIGNIFICANT THREATS TO THE VALIDITY. THEY INTRODUCE CONSTRUCT IRRELEVANT VARIANCE TO THE SITUATION. AND I THINK THOSE ARE ISSUES THAT WE NEED TO GUARD AGAINST. IN TERMS OF PAPER VERSUS ONLINE, MOST OF WHAT I'M GONNA TALK ABOUT, CERTAINLY ONLINE TESTING HAS NUMEROUS ADVANTAGES OVER PAPER AND PENCIL, AND THAT EXTENDS TO TEST SECURITY.

BUT WITH ALL ASSESSMENTS, I THINK WHAT WE REALLY NEED TO FOCUS ON IS THE INTENDED PURPOSES AND THE POTENTIAL CONSEQUENCES. AND ONCE WE NAIL THOSE DOWN, THEN WE HAVE TO HONE IN INTO THE SPECIFIC TYPES OF INTEGRITY AND THREATS THAT EXIST. I CAST THEM IN 5 AREAS—ITEM EXPOSURE, CANDIDATE AUTHENTICITY, DATA TRANSMISSION AND STORAGE, PROCTOR AND PERSONNEL INTEGRITY, AND SYSTEMS INTEGRITY. AND I WANT TO REITERATE WHAT I'VE HEARD EARLIER, THAT THIS REALLY SHOULD BE ABOUT PREVENTION AND MORE PREVENTION. AND DETECTION, TO ME, WHILE IT DOES SERVE TO IDENTIFY IRREGULARITIES, THE PRIMARY PURPOSE OF DETECTION IS TO MITIGATE FUTURE RISKS. SO I THINK PREVENTION IS 85% OF THE BATTLE AND DETECTION IS SOMETHING--AND INVESTIGATION IS SOMETHING WE NEED TO DO.

BUT I THINK WHEN YOU FIND SITUATIONS WHERE CHEATING IN INTEGRITY HAVE GONE ON, I THINK WE NEED TO LOOK AT OURSELVES AND WHAT CONDITIONS AND WHAT SYSTEMS HAVE WE CREATED THAT ALLOWED THIS TO HAPPEN.

BECAUSE WE KNOW FROM SOCIAL PSYCHOLOGY LITERATURE THAT THERE ARE A LOT OF PRECURSORS TO THEFT--OR COUNTERPRODUCTIVE BEHAVIOR IN WORKPLACES JUST GENERALLY IN THE LITERATURE. AND THOSE INSTANCES COME ABOUT IN TERMS OF THE SYSTEMS. THE GREATER THE OPPORTUNITY, THE MORE LIKELY YOU WILL TAKE INDIVIDUALS WHO HAVE NO PROCLIVITY TO CHEAT OR TO BE DISHONEST. AND YOU'LL INCREASE THEIR PROBABILITY OF DOING SO.

SO, IT REALLY IS A SYSTEMS ISSUE. I WANTED TO TALK ABOUT CHEATING, WHICH INCREASES WITH THE AGE OF THE STUDENT; WITH BANDWIDTH, WHICH IS JUST SIMPLY THE INFORMATION PER SECOND; AND ALSO THE DISTANCE. WE'RE NOT HERE TALKING ABOUT DISTANCE LEARNING OR PROCTORLESS EXAMS, BUT IF WE WERE, THOSE WOULD PRESENT SIGNIFICANT RISKS.

SIMILARLY, THE SUMMATIVE ASSESSMENTS AND THE DIFFERENT PURPOSES RESULT IN DIFFERENT INTENDED USES OF SCORES AND DIFFERENT THREATS. SCHOOL AND DISTRICT ACCOUNTABILITY PROBABLY PROVIDE LOW THREATS FOR STUDENT CHEATING BUT MODERATE THREATS TO TEACHER. STUDENT

REWARDS ENDORSE DIPLOMAS, COLLEGE CREDIT, SUCH AS AN A.P. EXAMINATION, PROBABLY INCREASE THE STUDENT CHANCES OF CHEATING. PROBABLY HIGH RISK FOR STUDENT BUT LOW RISK FOR TEACHERS AND EDUCATORS.

SIMILARLY, TEACHER AND EDUCATIONAL ACCOUNTABILITY, AGAIN, PROBABLY A LOW RISK FOR—STUDENT BUT HIGH RISK FOR TEACHER, AND SO WITH STUDENT BARRIERS, GRADUATION, RETENTION, DEVELOPMENTAL COURSES, PROBABLY A HIGHER RISK FOR STUDENT CHEATING, LOWER FOR TEACHERS.

SO I THINK THOSE ISSUES WE NEED TO--WHEN WE TALK ABOUT INTEGRITY AND PREVENTION, WE NEED TO REALLY LOOK AT THE INTENDED USE OF THE TEST AND FOCUS ON OUR PREVENTION IN THOSE AREAS.

I'M NOT GOING TO READ ALL OF THIS CHART, BUT I DID WANT TO JUST HIGHLIGHT TWO OF THE AREAS, EXTENDED TESTING WINDOW. AGAIN, THE RISKS TO THE STUDENT ARE A LITTLE BIT DIFFERENT THAN THE TEACHER.

WHEN YOU HAVE AN EXTENDED TESTING WINDOW, WHICH WE TEND TO DO WITH A LOT OF THE COMPUTER-BASED TESTING, I THINK YOU'RE INCREASING THE RISK THAT STUDENTS WILL DISCLOSE ITEMS AND TESTS TO OTHER STUDENTS. CAN YOU IMAGINE A SCENARIO WHERE WE HAVE THE SAME S.A.T. FORM ADMINISTERED IN YOUR SCHOOL FOR A FOUR-WEEK PERIOD?

WE MIGHT SPIRAL THE ITEMS. BUT EVERY ITEM WAS THE SAME. AND WE SPIRALED BETWEEN THREE ESSAY TOPICS. THAT TEST WAS LIVE AND IN COLOR ACROSS THE STATE OF NEW YORK AND MISSOURI FOR THREE WEEKS.

WOULD THAT BE TOLERATED? NO. NO. I THINK WITH THE AGE OF COMPUTER-BASED TESTING--AND ALSO PAPER-BASED TESTING AS RACE TO THE TOP IS REQUIRING TEACHER AND EDUCATOR ACCOUNTABILITY, THE IDEA THAT WE CAN HAVE TEACHERS PROCTOR THEIR OWN EXAM IS LOST. IT HAS ABSOLUTELY NO CREDIBILITY IN THIS DAY AND AGE.

IN TERMS OF REUSE OF ITEMS, I THINK THERE IS A GREATER CHANCE OF INTENTIONAL OR NON-INTENTIONAL DISCLOSURE OF ITEMS IF YOU DO NOT HAVE ADEQUATE POOLS BUT ALSO FOR TEACHERS AND EDUCATORS. IF YOU REUSE THE ITEMS OVER TIME, PRETESTING AND EQUATING COULD BE COMPROMISED, BUT THERE'S GREATER RISK THAT THE TEACHER WILL TEACH TO THE TEST OR DEVELOP TEMPLATES THAT WILL, AS WE HEARD THIS MORNING, WILL SOMEHOW ADVANTAGE ONE GROUP OVER ANOTHER.

LET ME TALK A LITTLE BIT ABOUT THE RECOMMENDATIONS THAT I WANTED--TO LEAVE YOU WITH IN TERMS OF PROCESSES AND POLICIES THAT COULD MITIGATE RISKS TO THE INTEGRITY OF C.B.T. TESTS.

FIRST OF ALL, I THINK WE WANT TO LOOK AT THIS AS A THREE-PHASED APPROACH.

ONE--WE WANT TO ELIMINATE THE OPPORTUNITY TO CHEAT TO THE EXTENT POSSIBLE. WE WANT TO REDUCE THAT OPPORTUNITY.

SECONDLY, WE WANT TO SEND THE MESSAGE THROUGH THE CULTURE, AS WE HAVE TALKED ABOUT, THAT THIS IS NOT TOLERATED.

BUT, THIRD, WE NEED TO EDUCATE. WE NEED TO EDUCATE THE STUDENT AND WE NEED TO EDUCATE THE SCHOOL PERSONNEL, WHAT CONSTITUTES APPROPRIATE AND INAPPROPRIATE ADMINISTRATION CONDITION. WHAT TYPE OF ASSISTANCE IS APPROPRIATE AND INAPPROPRIATE?

AND TO THE EXTENT THAT THERE'S ANY CONFUSION OR DOUBT ABOUT WHAT IS AND WHAT IS NOT APPROPRIATE, YOU ARE INCREASING THE OPPORTUNITY FOR CHEATING. AND SO FROM A FAIRNESS PERSPECTIVE AND FROM A VALIDITY ARGUMENT, I WOULD ARGUE THAT ALL THREE OF THOSE HAVE TO BE DONE.

IN TERMS OF ADMINISTRATION OF SCORING, I WOULD RECOMMEND THAT WE REDUCE THE OPPORTUNITY FOR CHEATING AS MUCH AS POSSIBLE. AS I SAID, I WANT TO REITERATE, CLASSROOM TEACHERS SHOULD NOT BE ADMINISTERING THE TESTS TO THEIR STUDENTS. OTHER CLASSROOM TEACHERS CAN DO THAT. BUT WHEN RESULTS ARE GOING TO HAVE AN IMPLICATION ON THE PERFORMANCE OF AN EDUCATOR OR EVEN THE SCHOOL, I

DON'T THINK WE CAN TOLERATE THAT ANY LONGER. AND IT OPENS THE DOOR.  
PROCTORS SHOULD HAVE NO STAKE IN THE OUTCOME.

THE ENVIRONMENT SHOULD PRECLUDE COPYING RESPONSES FROM OTHER  
STUDENTS SEATED ADJACENTLY. MANDATORY TRAINING IS IMPORTANT. EVEN  
MORE SO WITH COMPUTERS. STUDENTS SHOULD READ AND SIGN A STATEMENT  
LIKE AN HONOR CODE, AND POSSIBLY EDUCATORS AS WELL. YOU NEED TO  
PROHIBIT ALL HAND-HELD ELECTRONIC DEVICES, SMARTPHONE,  
CALCULATORS, YOU NEED TO EMPLOY A VARIETY OF ITEM FORMATS,  
INCLUDING CONSTRUCTED RESPONSE ITEMS. STEVE FERRARA MENTIONED THIS  
EARLIER. BECAUSE IT REDUCES AND CONFUSES THE OPPORTUNITY TO CHEAT.  
YOU NEED TO IMPOSE CONDITIONS ON RETESTING OPPORTUNITIES. YOU  
CANNOT HAVE RETESTING UNLESS YOU HAVE AN UNLIMITED ITEM BANK.

THERE ARE STUDENTS AND THERE ARE NON-STUDENTS WHO WILL ATTEMPT TO  
RETEST AS MUCH AS POSSIBLE TO GAIN AN ADVANTAGE TO GET ACCESS TO  
THE ITEMS.

TECHNOLOGY--PREPARE FOR THE UNEXPECTED. IT WILL OCCUR. HOW ABOUT  
THE STUDENT WHO GOES IN THERE PURPOSELY, UNPLUGS THE COMPUTER, OR  
DOES SOMETHING SO THAT THEIR TEST SETTING WILL BE DISRUPTED, SO  
THEY CAN TEST BACK TOMORROW, EVEN THOUGH THEY HAVE BEEN EXPOSED TO

3/4 OF THE ITEM POOL. THAT IS A VERY COMMON PLOY. AND IT WILL BECOME MORE COMMON WITH COLLEGE AND CAREER READINESS TESTS, IF THESE ARE HIGH STAKES TESTS.

ENSURE STUDENTS CANNOT ACCESS WEB RESOURCES. IF YOU WANT THEM TO USE SIMULATIONS OR INNOVATIVE ITEM POOLS, YOU CAN HAVE SOME OF THOSE WEB RESOURCES IN AN ISOLATED POOL. BUT YOU DO NOT ALLOW THEM TO GET TO THE OUTSIDE ENVIRONMENT. ITEM AND DATA HAVE TO BE ENCRYPTED AND STORED ON A SECURE SERVER.

IF YOU HAVE PAPER FORMS--AND WE ALL WILL FOR STATE ACCOUNTABILITY-- PAPER FORMS USE DIFFERENT ITEM BANKS, AND HAVE THE CHAIN OF CUSTODY ESTABLISHED. YOU NEED TO ALSO AUDIT SOCIAL NETWORKS. YOU NEED TO AUDIT SCHOOL PREPARATION OR TEST PREPARATION ACTIVITIES AND BLOGS TO ENSURE THAT SOME OF THE ITEMS, ESPECIALLY THE CONSTRUCTED RESPONSE ITEMS, ARE NOT BEING EXPOSED.

ENSURE HIGH SYSTEM RELIABILITY SO THAT YOU CANNOT EASILY START AND STOP. GUARD AGAINST SNIFFERS AND ATTEMPTS TO GET THE PROCTOR TO DISCLOSE THEIR PASSWORD. DISABLE NETWORK CAPABILITIES TO THE EXTENT POSSIBLE, USE DETECTION SOFTWARE, AND BACK UP ANY GRADE OR ROSTER THAT KEEPS GRADES OR STUDENT PERFORMANCE.

FINALLY--I WON'T GET INTO ALL THE STATISTICS--BUT THERE'S A LOT WRITTEN IN THIS AREA. I THINK THE NATIONAL COUNCIL ON MEASUREMENT AND EDUCATION, THEIR CONFERENCE. THE ASSOCIATION OF TEST PUBLISHERS ARE PLACES TO LOOK AT. CERTAINLY SOME OF THE JOURNALS.

BUT YOU WANT TO CHECK ON ABERRANCE RATES; YOU WANT TO LOOK AT RETEST AND SCORE VOLATILITY STATISTICS. YOU WANT TO CHECK ON IRREGULAR LATENCIES AND RESPONSE PATTERNS. YOU WANT TO LOOK AT BOTH HIGH AND LOW ABERRANT SCORES, CHEATING INDEXES AND THRESHOLDS, AND--ALSO FOR DISTANCE ASSESSMENTS, WE NEED TO REALLY MONITOR THEM MORE CLOSELY BECAUSE THE OPPORTUNITY TO CHEAT AND THE MOTIVATION DIFFERS IN THOSE SETTINGS. THANK YOU.

**JOHN FREMER:** I'M PRETTY EXCITED TO BE HERE. I'VE BEEN WORKING IN THE FIELD FOR 50 YEARS AND I'VE BEEN TO A LOT OF CONFERENCES, BUT RARELY HAVE I BEEN TO ONE THAT HAS SO MUCH POTENTIAL TO MAKE A DIFFERENCE.

I WORK FULL-TIME IN TEST SECURITY. I HAVE BEEN NOW FOR OVER NINE YEARS. AND I KNOW THAT PEOPLE ARE READY FOR STANDARDS, FOR GUIDELINES, FOR MESSAGES, AND I THINK WE'RE PREPARING THEM, NOT THAT IT FEELS SO COMFORTABLE.

MY METAPHOR OF THE PERFECT TEST SECURITY STORM IS INTENDED TO

PICK UP THE BOOK IN THE MOVIE "THE PERFECT STORM."

WE'RE IN THE MIDDLE OF IT, AND IT'S NOT GOING TO GET BETTER. AND IN MY MIND, IT COMES FROM MANY SOURCES. AND ANY TIME I READ SOMETHING THAT SAYS ALL OF THESE PROBLEMS HAVE TO DO WITH TEACHER BONUSES OR ALL OF THESE PROBLEMS HAVE TO DO WITH ONE THING, I KNOW THAT IT'S NOT TRUE.

THE SITUATION WE'RE IN IS A FUNCTION OF MANY DIFFERENT ISSUES.

IT'S A FUNCTION OF THE KINDS OF REQUIREMENTS THAT CAME WITH NO CHILD LEFT BEHIND. IT'S A FUNCTION OF THE FACT THAT WE ARE ON A TRAIN THAT IS GOING TO USE TEST SCORES AS PART OF TEACHER EVALUATION, AND IT SEEMS LIKE THERE'S NO TURNING THAT TRAIN. AT LEAST EVERYTHING THAT I READ SAYS THAT.

WE HAVE OUR STATE TESTS AT THE HIGH SCHOOL LEVEL USED AS GRADUATION REQUIREMENTS.

WE HAVE INCREASING EVIDENCE THAT CHEATING IS GOING UP. REPORTERS ASK ME--I GET ABOUT ONE INTERVIEW A WEEK FROM THE MEDIA--"IS IT REALLY GOING UP OR IS IT JUST MORE REPORTING?" I THINK IT'S REALLY GOING UP. I THINK THE EVIDENCE IS QUITE CLEAR.

I KNOW IT'S GOING TO STOP AT 100%, BUT IT MAY NOT STOP UNTIL IT'S THERE. I MEAN, IT'S UP TO 2/3, 3/4 NOW FOR YOUR AVERAGE HIGH SCHOOL STUDENTS.

THERE'S ALSO AN INCREASE IN SOPHISTICATION IN CHEATING. JUST LOOK ON THE INTERNET IF YOU HAVE ANY DOUBT ABOUT THAT.

AND THERE'S ALL KINDS OF ADVICE ON HOW TO CHEAT. SO I WON'T GIVE YOU ANY ADVICE OF HOW TO CHEAT. I TRY NOT TO DO THAT IN ANY SETTING.

AND THE CONSORTIA'S PLAN TO USE COMPUTER-BASED TESTING IS GOING TO ADD SOME FASCINATION TO THAT STORM.

SO, AS OTHERS HAVE POINTED OUT, WHAT WILL BE THE EFFECT OF C.B.T.? WELL, IT WILL REDUCE SOME RISKS. OTHERS WILL REMAIN. AND SOME WILL GO UP.

WHAT RISKS WILL BE REDUCED? YOU'RE NOT SO LIKELY NOW TO LOSE A TEST BOOK BECAUSE THERE AREN'T ANY TEST BOOKS. OR THEY'RE NOT SO LIKELY TO BE STOLEN WHEN THE TRUCK DELIVERING THE TESTS STOPS SO THE DRIVER CAN GET A COFFEE OR SOMETHING, WHICH IS SOMETHING THAT HAPPENS NOW. AND COPYING DURING TESTING, IF YOU DO C.B.T. RIGHT AND I THINK OUR CONSORTIUM FOLK WILL--THEY HAVE VERY GOOD PEOPLE

WORKING WITH THEM IN ALL KINDS OF ROLES—BUT THAT STILL LEAVES MOST OF THE OTHER PROBLEMS WE'VE HAD WITH PAPER AND PENCIL.

THEY DON'T GO AWAY. THEY'RE STILL THERE. PRE-KNOWLEDGE IS THERE, ASSISTING DURING THE EXAM IN VARIOUS WAYS, STEALING, MEMORIZING QUESTIONS, COLLUDING. AND BECAUSE OF TECHNOLOGY-ASSISTED CHEATING, THE PEOPLE COLLUDING DON'T EVEN HAVE TO BE IN THE SAME ROOM. THEY JUST NEED TO HAVE A WAY OF COMMUNICATING WITH EACH OTHER.

I HEARD A STORY ABOUT THAT YEARS AGO. IT WAS A WEST AFRICAN COUNTRY, AND THEY WERE DETERMINED TO AVOID CHEATING. AND THEY KEPT HAVING THIS PROBLEM OF ANSWERS BEING TRANSFERRED INTO THE TESTING CENTER FROM OUTSIDE. SO THEY RING THE TEST CENTER WITH ARMED GUARDS SO THERE COULD BE NO POSSIBILITY OF PEOPLE MOVING IN AND OUT. BUT SHORTLY AFTER THE TEST WAS ADMINISTERED, WAS GIVEN AND HANDED OUT, FROM ACROSS THE RIVER, WHICH WAS ANOTHER COUNTRY, THE ANSWERS STARTED TO BE BROADCAST TO THE TESTING SITE. SO THERE ARE WAYS AROUND ALMOST ANYTHING WE WOULD TRY TO DO.

SOME RISKS WILL GO UP.

WE HAVE TO DISTRIBUTE ALL OF THOSE ELECTRONIC TESTS TO THE SITES. THAT'S NOT A PERFECT PROCESS. WE HAVE SOME VERY TECH-SAVVY FOLKS THAT ARE GOING TO TRY TO UNDERMINE IT. AS MANY HAVE MENTIONED,

WE'RE GOING TO HAVE THESE EXTENDED PERIODS. AND THERE'S JUST NO WAY THAT YOU'RE GOING TO HAVE TESTING PERIODS OF FOUR, FIVE, SIX WEEKS WITHOUT KNOWLEDGE GETTING OUT ACROSS THE TESTING AREAS, PARTICULARLY ABOUT ESSAY AND KNOWLEDGE-BASED QUESTIONS. I THINK WE'RE GOING TO HAVE A NEW EXPANDED UNDERGROUND MARKET FOR TESTS BECAUSE OF THE MULTIPLE STATES.

MAYBE IT WASN'T WORTH IT TO STEAL THEM AND TRY TO SELL THEM WHEN THEY WERE JUST FOR ONE STATE, BUT IF THEY'RE FROM 20 DIFFERENT STATES, I THINK THAT MARKET IS GOING TO BUILD. AND WE DO A LOT OF WORK WITH CERTIFICATION PROGRAMS--HUNDREDS OF MILLIONS OF DOLLARS ARE BEING MADE COMPROMISING CERTIFICATION PROGRAMS, NOT SO MUCH YET IN EDUCATION, BUT IT WOULD ONLY HAVE TO BE A FEW MILLIONS FOR IT TO GET TO BE A SERIOUS PROBLEM.

AND I ALSO WORRY THAT C.B.T., IN GENERAL, HAS INCREASED THE COST OF TEST DEVELOPMENT. I THINK YOU CAN MAKE ARGUMENTS THAT IT'S WORTH IT. I WOULD TEND TO GO ALONG WITH THAT, BUT IT DOES COST MORE MONEY. AND THAT MAY MEAN OVER TIME, WE'RE UNWILLING TO INVEST IN SECURITY THE WAY WE NEED TO DO.

FORTUNATELY, THERE ARE SOLUTIONS THAT WE CAN DRAW ON, INCLUDING, WE'RE GETTING BETTER AT USING VARIOUS TYPES OF DATA FORENSICS. SINCE WE WERE STARTING AT A VERY LOW LEVEL, WE STILL HAVE QUITE A

WAYS TO GO, BUT WE'RE GETTING BETTER AT HOW TO ANALYZE PERFORMANCE IN ORDER TO STAVE OFF CHEATING AND COMPUTER-BASED TESTING. AND I THINK DOWN THE ROAD, WE'LL BE ABLE TO DETECT IT WHILE IT'S OCCURRING AND DO SOMETHING TO INTERFERE WITH THE CHEATING ONLINE, WHICH IS WHERE WE WANT TO END UP.

IF WE GET TO HAVE WIDESPREAD USE, WHICH IS WHAT I WOULD PREDICT--SO IT'S ROUTINE THAT YOU HAVE DETECTION BUILT INTO YOUR C.B.T.-- I PREDICT THIS OUTCOME--IT WILL BECOME MORE COST EFFECTIVE, LESS EXPENSIVE. AND BECAUSE WE'LL HAVE A LOT OF EXPERIENCE EXPLAINING IT, WE'LL GET BETTER OUT OF IT. AND ONE OF THE PROBLEMS WE HAVE NOW IS WE RUN DATA FORENSICS ANALYSES, BUT THE PEOPLE THAT WE'RE EXPLAINING IT TO ARE LEARNING IT FOR THE FIRST TIME. THEY KNOW OUR OTHER ANALYTIC METHODS. THEY DON'T KNOW DATA FORENSICS, SO WE HAVE TO GET BETTER EXPLAINING.

IT'S NOT THAT THEY HAVE TO GET BETTER AT UNDERSTANDING. IF YOU EXPLAIN AND THEY DON'T UNDERSTAND, IT'S YOUR FAULT AS THE EXPLAINER. IT'S NOT THE FAULT OF THE PEOPLE RECEIVING THE DATA.

SO WHAT TECHNIQUES DO WE HAVE IN C.B.T.? SOME OF THEM ARE EXACTLY THE SAME TECHNIQUES WE HAD BEFORE. WE CAN DO UNUSUAL GAINS ANALYSES OR COUNTS OF PERFECT OR NEAR-PERFECT SCORES. ONE THING I WANT TO

COMMENT ON THOSE TWO METHODS--THEY CAN BE DONE TYPICALLY BY THE MEDIA OR ANYBODY WHO HAS ACCESS TO A SITE WHERE YOU HAVE THE INFORMATION POSTED. YOU DON'T NEED THE COOPERATION OF THE VENDOR OR THE DEPARTMENT. ALL OF THE OTHER METHODS DO REQUIRE THAT.

YOU CAN'T DO IT ON YOUR OWN, SO YOU CAN'T ANALYZE THE SIMILARITY OF RESPONSES OF INDIVIDUAL TEST-TAKERS UNLESS YOU HAVE ACCESS TO THOSE RESPONSES. YOU CAN'T ANALYZE AN INDIVIDUAL'S RESPONSE PATTERNS, WHAT, IN I.R.T. IS PERSON FIT. IS WHAT THEY'RE RESPONDING LOGICAL? AND THE EASIEST ILLOGICAL EXAMPLE I CAN GIVE IS THEY GET THE HARD QUESTIONS RIGHT AND THEY GET THE EASY QUESTIONS WRONG. THAT DOESN'T HAPPEN IN ANY DEGREE UNLESS THEY'RE PRE-KNOWLEDGE.

WE CAN STILL DO ANSWER CHANGING ANALYSES, THE COMPUTER ANALOG OF ERASURE ANALYSES, EXCEPT NOW WE KNOW EXACTLY WHAT ANSWERS THEY TOOK AND IN WHAT ORDER. AND I THINK WE'RE STILL WORKING ON HOW TO DO THAT. WE HAVE ONE ENTIRELY NEW THING WE CAN DO WHICH IS RESPONSE TIME ANALYSES. HOW MUCH TIME DID THEY SPEND ON INDIVIDUAL QUESTIONS?

WE CAN WORK OUT THEIR PATTERNS. WE CAN LOOK FOR ANOMALIES. AND IN THE THEFT DEPARTMENT, THAT'S A REAL CLUE, IF THEIR TIMING IS VERY DIFFERENT FROM ANYONE ELSE. FAST, FAST, FAST, SLOW, SLOW,

SLOW. IT'S BECAUSE THEY'RE MEMORIZING OR MAKING A RECORD OF THOSE OTHER QUESTIONS. AND ALSO, WE CAN DO WEB MONITORING.

SO I'M GOING TO CLOSE WITH SOME RECOMMENDATIONS. I HOPE THEY SOUND FAMILIAR WITH THINGS YOU'VE HEARD BEFORE.

ACKNOWLEDGE HOW SERIOUS IT IS. EXPECT CHEATING.

USE MULTIPLE DETECTION METHODS.

SOMEHOW KEEP THE WINDOWS AS SMALL AS YOU CAN.

THEY CAN'T BE EVERYBODY ON THE SAME DAY, THE SAME TEST THE WAY IN MANY INSTANCES WE HAVE NOW, BUT MAKE IT AS SHORT AS POSSIBLE.

EVALUATE THE CHAIN OF CUSTODY FOR THE TEST BANKS CAREFULLY.

INCREASE EMPHASIS ON TRAINING, WHICH IS WHAT EVERYONE HAS SAID, BUT WE NEED TO WORK TOGETHER ON THAT.

HAVE ENOUGH RESOURCES FOR TEST SECURITY. WHEN YOU'RE PILOTING EVERYTHING ELSE IN THE CONSORTIA, FOR EXAMPLE, PILOT YOUR CHEATING DETECTION METHODS AT THE SAME TIME SO YOU'RE READY WHEN YOU GO OPERATIONAL.

LEARN FROM EACH OTHER. THE CONSORTIA HAVE FOLKS ASSOCIATED WITH IT, LIKE TONY AND OTHERS, WHO HAVE C.B.T. I GUESS INDIANA IS GOING INTO IT, TOO. AND LEARN FROM OTHERS WHO ARE NOT IN EDUCATION.

AND FINALLY, TRACK EVERYTHING IN THE WAY THAT THE CENTER FOR DISEASE CONTROL TRACKS DISEASES. COLLECT ALL THAT INFORMATION; PUT IT TOGETHER; AND TRY TO FIGURE OUT EXACTLY WHAT'S HAPPENING, BECAUSE I DON'T THINK ANY OF US COMPLETELY UNDERSTAND THE PROCESS YET. AND C.B.T. IS NEW ENOUGH THAT I THINK THERE'S A LOT TO LEARN.

THANK YOU.

**JACK BUCKLEY:** THANK YOU VERY MUCH. WES?

**WES BRUCE:** THANKS. SOME OF THE THINGS I'M GOING TO SAY ARE GOING TO HOPEFULLY EXTEND SOME OF THE COMMENTS FROM BOTH WAYNE AND JOHN, BUT FIRST, I NEED TO REPEAT THE DISCLAIMER YOU HEARD EARLIER.

MY VIEWS DO NOT REPRESENT THOSE OF THE DEPARTMENT OF EDUCATION, NCES, SECRETARY OF EDUCATION, THE PRESIDENT, CONGRESS, SUPREME COURT, OR THE U.N.

AND, OFTEN, THEY DON'T REFLECT MY OWN THOUGHTS.

[LAUGHTER]

BUT I WANT TO TALK TO YOU SOME ABOUT THE TRANSITION, BECAUSE, WHILE IN '14, '15, THE CONSORTIA ANTICIPATE THAT THE TESTS, WITH THE EXCEPTION OF ACCOMMODATIONS, ALL BE ONLINE, MANY OF US IN THE MEANTIME ARE MOVING, ARE MAKING TRANSITIONS, AND TRYING TO THINK ABOUT HOW YOU TRANSITION YOUR SECURITY, YOUR TESTING INTEGRITY PROGRAMS FROM PAPER TO COMPUTER. AND SOME OF THE PROMISES AND CONCERNS THAT IT BRINGS.

SO, FOR MOST OF US, INDIANA IS ONE OF THOSE PLACES WHERE WE'RE TRANSITIONING BY GRADE LEVEL, BY CONTENT AREA, SOMETIMES BY SCHOOL, FROM PAPER TO ONLINE AS CAPACITY IS AVAILABLE. WE DON'T YET HAVE ENOUGH COMPUTERS IN EVERY SCHOOL IN THE STATE OF INDIANA FOR ALL STUDENTS TO TAKE THE TEST ONLINE. SO I'VE GOT SOME PLACES WHERE THE SAME GRADE LEVEL, IN THE SAME SCHOOL DISTRICT, ONE SCHOOL'S TESTING ONLINE, ONE SCHOOL'S TESTING ON PAPER. AND SO YOU NEED TO THINK ABOUT, AS YOU'RE MAKING THOSE TRANSITIONS AND SPENDING LOTS OF TIME THINKING ABOUT HOW YOU TRANSITION YOUR PROGRAM FROM PAPER TO ONLINE, YOU NEED TO THINK ABOUT HOW YOU'RE GOING TO TRANSITION TO SOME OF THESE TOOLS THAT LOTS OF PEOPLE HAVE TALKED ABOUT, BECAUSE YOU WANT TO LEVERAGE THOSE STRATEGIES THAT YOU CAN FOR ONLINE, THAT ARE SPECIFIC TO ONLINE, SOME OF WHICH ARE

COMPLEMENTARY, SOME OF WHICH ARE THE SAME, SOME OF WHICH ARE UNIQUE TO ONLINE.

BUT IN TERMS OF THE FIELD, WHEN I'VE GOT A PRINCIPAL WHO'S GOT HALF OF HIS GRADE LEVELS ONLINE AND HALF OF HIS GRADE LEVELS ON PAPER, THEY DON'T WANT TWICE AS MANY REPORTS ABOUT TESTING INTEGRITY.

YOU HAVE TO THINK ABOUT HOW YOU'RE GOING TO BLEND THOSE THINGS TOGETHER SO THAT THEY CAN ACT ON THAT.

WE'VE TALKED A LOT TODAY ABOUT THE FACT THAT THIS IS A SMALL PERCENTAGE OF FOLKS, WHILE IT MAY BE INCREASING, IT'S STILL, YOU KNOW, KIND OF THE TIP OF THE ICEBERG, ARE THE PEOPLE WHO ARE COMMITTING THESE KINDS OF ACTS. AND WE WANT TO MAKE SURE THAT WE'VE MADE IT AS SIMPLE AS POSSIBLE FOR THOSE WHO NEED TO INVESTIGATE TO BE ABLE TO USE THE INFORMATION.

SO, DO LEVERAGE THE DIFFERENCES THAT COMPUTER-BASED TESTING BRINGS TO YOU.

ONE OF THE PROBLEMS IS GOING TO BE DECIDING WHAT YOU DON'T NEED BECAUSE NOW YOU'VE GOT INFINITELY MORE DATA ON EVERY SINGLE STUDENT. YOU ACTUALLY, IF YOU CHOOSE TO, COULD RECORD EVERY SINGLE KEYSTROKE THAT A STUDENT MAKES. YOU HAVE TO DECIDE HOW MUCH

INFORMATION YOU WANT TO MAINTAIN AND USE, OR HOW MUCH DATA YOU WANT TO KEEP AND TURN INTO INFORMATION AS A PART OF YOUR REGULAR PROCESS, AND WHAT WILL YOU SYSTEMATICALLY KEEP AS YOUR BACKUP, OR KEEP IN YOUR BACK POCKET.

I THINK OF IT AS, YOU KNOW, IF YOU WANT TO USE AN ANALOGY HERE, AN ERASURE ANALYSIS THAT YOU DO ON ALL STUDENTS, BUT WHEN PUSH COMES TO SHOVE, YOU PULL STUDENT DOCUMENTS IN CERTAIN CASES.

YOU DON'T PULL EVERY SINGLE STUDENT DOCUMENT AND LOOK THROUGH THOSE, BUT YOU'VE GOT A SET OF INFORMATION THAT'S GOING TO BE PART OF YOUR REGULAR PROCESS. AND YOU'VE GOT INFORMATION THAT YOU WANT TO HAVE IN YOUR BACK POCKET, BECAUSE YOU'VE GOT, AS JOHN SAID, YOU'VE GOT INFORMATION ON TIME AND HOW MUCH TIME EACH STUDENT SPENT ON EACH ITEM.

YOU ACTUALLY HAVE WHAT CLOCK TIME EACH STUDENT ACTUALLY ANSWERED AN ITEM. AND IF ALL THE STUDENTS ARE ANSWERING THE ITEM WITHIN A COUPLE OF SECONDS, MAYBE YOU HAVE A REAL-TIME WAY TO DETECT SOME POTENTIAL FRAUD.

YOU HAVE THE AMOUNT OF TIME STUDENTS SPEND ON--EACH SESSION AND WHAT WE'RE CONSISTENTLY FINDING IN INDIANA, ON AVERAGE, IS THAT

STUDENTS TAKE LESS TIME WITH THE ONLINE TESTS THAN THEIR REGULAR TIME.

WE CONSISTENTLY HAVE--WHEN THE TEST IS ADMINISTERED ON PAPER, WE GIVE THE SAME TIME LIMITS, WHERE OUR TESTS ARE TIMED. BUT WHEN THEY TAKE IT ONLINE, CONSISTENTLY, ALL THE STUDENTS IN THE CLASSROOM ARE DONE WELL BEFORE THE END OF THE TIME.

BUT THAT'S AN IMPORTANT PIECE, NOT JUST TO JOHN'S POINT, THAT YOU CAN USE AS YOU'RE DEVELOPING TESTS TO MAKE SURE THAT YOU UNDERSTAND THE AMOUNT OF TIME THAT IT REALLY TAKES FOR A TASK. WE'RE EXPLORING SOME NEW ITEM TYPES. AND I THINK WE DO NEED TO LEVERAGE OUR PILOT TESTING THAT WE'RE GOING TO BE DOING AND CAPTURE THAT INFORMATION UP FRONT.

AND WE DON'T REALLY KNOW WHAT THE TIME REQUIREMENTS ARE. WE DO WANT THOSE TIMES--THOSE TESTS TO BE FAIR. BUT YOU'VE GOT, YOU KNOW-- THIS, TO SOME EXTENT, MINUTE LEVEL OF INFORMATION.

YOU HAVE THE ACTUAL ORDER IN WHICH STUDENTS ANSWER TEST ITEMS. ANOTHER THING THAT WE'VE FOUND IS MORE JUMPING AROUND, TO USE A HIGHLY TECHNICAL TERM, IN THE TEST THAN PEOPLE NORMALLY WOULD ANTICIPATE. IF YOU ASK TEACHERS HOW MUCH DO STUDENTS JUMP AROUND ON

THE TEST, THE TYPICAL RESPONSE IS, "WELL, IF THEY CAN'T GET ONE, THEY MIGHT MARK IT TO COME BACK LATER."

WHAT WE SEE IN ONLINE IS ACTUALLY STUDENTS MOVING AROUND MUCH MORE THAN WE WOULD HAVE EXPECTED, WHEN WE START ANALYZING THE ACTUAL PATTERN OF RESPONSES IS VERY DIFFERENT. AND YOU HAVE TO THINK ABOUT HOW YOU WANT TO USE THAT.

AND TO A POINT, YOU ACTUALLY HAVE THE REAL PATTERN OF ANSWER CHOICE CHANGES. AND I'M GOING TO--CIRCLE BACK TO THAT IN AN EXAMPLE.

SO MANY OF THE METRICS--AND JOHN'S TALKED A LITTLE BIT ABOUT THIS-- THAT WE USE WITH PAPER ARE EQUALLY VALID FOR COMPUTER-BASED TESTING--SCORE CHANGE METRICS BOTH AT THE SCHOOL AND THE STUDENT LEVEL; PART TO WHOLE, HOW A STUDENT IS PERFORMING ON INDIVIDUAL ITEMS VERSUS HOW THEY PERFORM ON THE TEST OVERALL, BACK TO THE EASY ITEMS VERSUS DIFFICULT ITEMS, AND LOOKING AT THAT AT BOTH SCHOOL, STUDENT AND CLASS LEVEL; PERFECT SCORE REPORTS, WHICH IS, AS JOHN POINTED OUT, SOMETHING THAT'S EASILY OBTAINABLE.

THAT'S ONE OF THE EARLY INDICATORS THAT WE USE. STILL--CAN NEVER GET OUT OF MY MIND THE OPPORTUNITY I HAD WHEN THERE WERE 24 PERFECT SCORES IN THE STATE OF INDIANA, AND 23 OF THOSE CAME FROM THE SAME HIGH SCHOOL.

THAT WAS A GREAT DAY. AND JUST A SIDE NOTE. ONE OF THE MOST DIFFICULT THINGS THAT THOSE OF US IN THE ASSESSMENT BUSINESS HAVE TO DO IS GO THROUGH THAT PROCESS OF PUNISHING STUDENTS FOR THE SINS OF ADULTS. SO, YOU KNOW, BACK TO THE LITANY THAT YOU'VE HEARD ABOUT PREVENTION BEING REALLY CRITICAL.

SO I WANT TO WALK YOU THROUGH AN ILLUSTRATION ABOUT THE TRANSITION AND TALK ABOUT THE ERASURE ANALYSIS.

SO YOU'VE HEARD A LOT ABOUT IT TODAY. I WANT TO DRILL INTO IT. JACK WAS UPSET THAT WE HADN'T REALLY GOTTEN TO THE WEEDS A LOT TODAY. SO I WANT TO MAKE SURE THAT WE DO GET DEEP DOWN IN.

SO WHEN WE DO AN ERASURE ANALYSIS--TRADITIONALLY WITH PAPER--WHAT WE'RE DOING IS LOOKING AT LIGHTER AND DARKER MARKS ON AN ANSWER DOCUMENT.

AND THE GENERAL LOGIC IS THAT THE LIGHTER RESPONSE IS CONSIDERED TO BE AN ERASURE. AND THE DARKER MARK, IF THERE ARE TWO MARKS FOR ANY PARTICULAR ITEM, IS CONSIDERED TO BE THE FINAL ANSWER, OK?

AND SO THEN WE APPLY LOGIC AND SAY, IF THE LIGHTER ONE WAS THE WRONG ANSWER, WAS A WRONG ANSWER FOR THAT ITEM, AND THE DARKER MARK

IS THE RIGHT ANSWER, THEN WE FLAG THAT ITEM AS A WRONG TO RIGHT,  
OK?

THE PROBLEM IS--AND, OF COURSE, THEN YOU'VE HEARD ALL ABOUT THE  
CRITERIA--HOW MANY STANDARD DEVIATIONS DO YOU USE--CLASSES AND  
SCHOOLS GET FLAGGED FOR EXCEEDING. LET'S SAY, IN INDIANA WE USE A  
FOUR-STANDARD DEVIATION FLAG, AND SOMETIMES THEY GET FLAGGED.

NOW, THE FIRST ISSUE HERE IS THAT, OF COURSE, WHENEVER YOU SET A  
BAR THAT'S BASED ON SOME KIND OF DISTRIBUTION OF SCORES, WHAT ARE  
THE CHANCES THAT IF NO ONE CHEATED ON THE TEST THAT YOU'D HAVE AN  
ERASURE PATTERN AT SOME CLASSROOM AND SOME SCHOOL THAT WAS GREATER  
THAN FOUR STANDARD DEVIATIONS ABOVE EITHER THE MEAN OR THE MEDIAN?  
IT WILL HAPPEN.

SO, YOU KNOW, THERE'S ALWAYS THAT ISSUE ABOUT--STARTING THESE  
INVESTIGATIONS AND DECIDING WHEN, IN FACT, THERE HAS BEEN WHEN  
SOMETHING'S GONE BEYOND A CONCERN, AND REALLY GOTTEN INTO AN ISSUE  
OF CHEATING OR WHERE TEST INTEGRITY HAS REALLY BEEN COMPROMISED.

SO IF YOU TOOK STATISTICS IN COLLEGE, YOU KNOW THAT THERE OFTEN ARE  
PEOPLE BEYOND THE FOURTH STANDARD DEVIATION. SO, THE BIGGEST  
CONCERN THAT I HAVE--AND I USE THESE ANALYSES ALL THE TIME--IS THAT

WE DON'T KNOW WHAT THE ACTUAL PATTERN OF A STUDENT'S RESPONSES WERE ON A PAPER DOCUMENT.

WE HAVE NO IDEA ON EARTH. WE MAKE A HUGE ASSUMPTION ABOUT IT AND INFER FROM THAT BECAUSE WE DON'T KNOW WHETHER OR NOT THE PATTERN WAS WRONG TO RIGHT OR WHETHER IT WAS RIGHT TO WRONG TO RIGHT. NO IDEA.

BUT WITH COMPUTER-BASED TESTING, WE DO KNOW THAT; WE CAN TELL, AND POTENTIALLY REDUCE THE NUMBER OF FALSE POSITIVES BECAUSE WE KNOW THE PATTERN IN WHICH THE STUDENTS ANSWERED. WE KNOW WHEN THEY MADE THEIR FIRST ANSWER, IF THEY CHANGED IT, WHETHER THEY WENT BACK TWO OR THREE OR FOUR TIMES. ALL THAT GETS CAPTURED.

WE CAN FACTOR IN OTHER DIMENSIONS, LIKE WHEN THOSE THINGS WERE CHANGED. BUT IN THE END, THIS IS STILL GOING TO BE A STATISTICAL ANALYSIS, AND WILL BE SUBJECT TO THE SAME KINDS OF LIMITATIONS.

EVEN THOUGH THERE MIGHT BE FEWER FALSE POSITIVES, WE STILL HAVEN'T PROVEN ANYTHING. ALL WE'VE DONE IS CREATE AN INFERENCE. IN INDIANA, WE COMBINE THE TWO ERASURE ANALYSES INTO ONE DOCUMENT FOR SCHOOLS WHO ARE TESTING. WE USE THE PATTERN ANALYSIS FROM THE COMPUTER-BASED VERSION AND COMBINE IT IN A REPORT. WE FLAG CLEARLY WHICH CLASSROOMS TOOK THE TESTS BY COMPUTER AND WHICH BY PAPER,

BUT WE TRY TO PRODUCE ONE REPORT THAT HAS THE SAME KIND OF STATISTICS FOR SCHOOLS TO USE, SO THAT WE USE THE SAME FLAGGING CRITERIA AND HAVE THE SAME EXPECTATIONS IN TERMS OF INVESTIGATING AND REPORTING SO THAT IT'S EASY TO COMMUNICATE.

WE'VE TRIED TO THINK ABOUT THE CONTEXT AND THE COGNITIVE LOAD ON FOLKS--TO A POINT EARLIER ABOUT TRYING TO FIND BETTER WAYS TO COMMUNICATE ABOUT THESE METRICS, AND TO THINK ABOUT HOW YOU MAKE THAT TRANSITION IN THIS WORLD.

SO COMPUTER-BASED SECURITY ISN'T GOING TO BE OMNIPOTENT. IT'S STILL INFERENTIAL. AND WE MAY BE ABLE TO GET STRONGER OR ADDITIONAL EVIDENCE THAT WE COULDN'T HAVE BEFORE. IT STILL PROVES NOTHING. IT TAKES THE INVESTIGATION TO DO THAT. AND AS YOU HEARD FROM BALTIMORE, SOMETIMES EVEN AN INVESTIGATION WON'T PROVE EXACTLY WHAT HAPPENED. SO THE INVESTIGATION IS STILL CRITICAL.

YOU KNOW WHAT HAPPENED; YOU DON'T KNOW WHO OR WHEN OR HOW. THE PRESS LOVES A SCANDAL. AND COMPUTER-BASED TESTING WILL ALLOW YOU TO CREATE AN EVEN BIGGER ONE.

**JACK BUCKLEY:** WES, THANKS VERY MUCH. TONY, I'M AFRAID WES TOOK ALL YOUR TIME. SO...

**TONY ALPERT:** IT'S NOT THE FIRST TIME.

**JACK BUCKLEY:** [LAUGHING] PLEASE GO AHEAD.

**TONY ALPERT:** SO THANK YOU VERY MUCH FOR THE OPPORTUNITY TO TALK WITH YOU. INTERESTING. WHEN WE WERE TOLD WHAT THE PURPOSE OF THE SEMINAR WAS, WE FOCUSED ON HOW TO HELP STATES AND DISTRICTS IMPLEMENT BETTER PROCEDURES. IT DIDN'T OCCUR TO US THAT PEOPLE WOULD BE ASKING THE QUESTION, "WELL, WHAT ARE THE CONSORTIA GOING TO DO TO ADDRESS THESE?"

SO I'M GOING TO GO OFF--I KNOW. WE'VE JUST BEEN BUILDING THE TEST...SO I'M GOING TO GO A LITTLE BIT OFF-SCRIPT.

SOME OF THESE POINTS ARE GOING TO BE DIRECTLY RELATED TO SBAC,

BUT MANY OF THEM ARE CONSISTENT WITH THE APPROACH PARCC IS USING, I'M SURE.

BUT I WILL START OUT TALKING A LITTLE BIT ABOUT SHARED RESPONSIBILITY IN A SLIGHTLY DIFFERENT WAY THAN GREG TALKED ABOUT EARLIER. AND PROBABLY THE MOST CRITICAL IN MY PERSPECTIVE IS THAT WHEN WE EXPECT--WE SHOULD MODEL, AS AN ACTIVITY, RIGOROUS

IMPLEMENTATION AT THE STATE LEVEL. AND I WOULD SAY THE SAME THING AT THE CONSORTIUM LEVEL.

SO WHEN WE PUT OUT AN ASSESSMENT, IT SHOULD BE PERFORMING IN THE MANNER IN WHICH WE DESCRIBED IT, OR SHOULD, SO THAT THERE'S AN EXPECTATION OF PROFESSIONALISM ACROSS THE ENTIRE SYSTEM. AND WHEN SOMETHING GOES WRONG, IT JUST OPENS UP THE DOOR FOR IDIOSYNCRATIC BEHAVIOR. AND SO TO THE EXTENT THAT EVERYTHING WORKS AS IT SHOULD, PEOPLE BEGIN TO EXPECT MORE OF THEMSELVES, AS WELL.

SO, SIMILAR TO THAT, WE SHOULD ESTABLISH A CULTURE OF SECURITY. AND THAT'S BEEN TALKED ABOUT PREVIOUSLY, BUT I THINK IT GOES BEYOND ASSESSMENT SECURITY AND WELL INTO DATA SECURITY BECAUSE THE LINE BETWEEN COMPUTER-BASED ASSESSMENT AND THE DATA IS MUCH

MORE BLURRY THAN IT IS WHEN YOU HAVE A PAPER-BASED ADMINISTERED FORM AND THE DATA COMES BACK MONTHS LATER. WITH COMPUTER-BASED TESTS, IT IS POSSIBLE AND LIKELY THAT THE DATA WILL BE RETURNED IMMEDIATELY, AND THERE'S TRANSFERS OF STUDENT IDENTIFIERS. SO YOU REALLY HAVE TO THINK BROADER ABOUT HOW YOU'RE GOING TO ADDRESS MANY DIFFERENT TYPES OF BREACHES.

AND THESE POLICIES SHOULD INCLUDE THE LARGER NETWORK OF ADULTS THAT NEED TO BE INVOLVED IN ADMINISTERING A COMPUTER-BASED TEST.

SO NO LONGER IS IT A TEACHER IN ISOLATION OR A PROCTOR IN ISOLATION.

YOU HAVE TO GET THE I.T. FOLKS INVOLVED, POSSIBLY THE SECRETARIES INVOLVED IN LOG--IN HELPING THE TICKETING PROCESS FOR THE STUDENTS. SO THERE'S MANY MORE PEOPLE, AND TO THE EXTENT THAT THOSE PEOPLE ARE GOING TO BE INTERACTING WITH--EITHER SECURE DATA OR SECURE MATERIALS IN THE FORM OF THE COMPUTER-BASED TEST, THE POLICIES HAVE TO ADDRESS THAT BREADTH.

THERE ARE ADDITIONAL COMPLEXITIES IN TERMS OF LOGISTICS, WHICH INCLUDES BRINGING STUDENTS DOWN TO A COMPUTER LAB AND GIVING GOOD INSTRUCTIONS ABOUT WHAT SHOULD HAPPEN AND WHAT SHOULDN'T HAPPEN.

THERE'S COLLECTING OF NOTES THAT STUDENTS MIGHT HAVE BEEN TAKING DURING THE SESSION. WHICH, EVEN IN A COMPUTER-BASED TEST, THERE'S PAPER INVOLVED AND THAT COULD VIOLATE THE SECURITY.

THERE'S STUDENTS LOGGING OFF OR OTHERWISE SECURING THE COMPUTER WHEN THE STUDENT GOES TO THE RESTROOM.

ALL THESE PROCEDURES NEED TO BE DELINEATED. OTHERWISE, IT COULD INADVERTENTLY AFFECT THE VALIDITY OF THE ASSESSMENT. AND THEN THERE'S COMPLEXITIES OF NEW ITEM TYPES. SO TO THE EXTENT THAT

STUDENTS ARE LISTENING TO AN AUDIO PORTION OF THE TEST, IF HEADPHONES AREN'T USED OR THE ROOM ISN'T OTHERWISE SECURED, THEN MOST LIKELY THE SECURITY OF THE TEST IS AT-RISK.

AND THEN TO THE EXTENT THAT STUDENTS ARE ACTUALLY SPEAKING, LIKE THEY DO IN OREGON'S ENGLISH LANGUAGE PROFICIENCY ASSESSMENT, THAT CREATES A WHOLE NEW CATEGORY OF POTENTIAL PROBLEMS IN IRREGULARITIES IN ADMINISTRATION. SO THE REASON WHY I TALK ABOUT THOSE, IS AS A CAVEAT HERE, I THINK THE VAST MAJORITY OF IRREGULARITIES ARE MISTAKES THAT ARE OFTEN DUE TO A LACK OF CLEAR DELINEATION BY THE STATE THAT'S PROVIDING THE ASSESSMENT.

THERE ARE CERTAINLY SOME SALACIOUS INSTANCES WHERE PEOPLE INTENTIONALLY CHEAT. BUT, BY FAR, I THINK MOST PEOPLE ARE INVESTED IN ADMINISTERING THE SYSTEM TO THE BEST OF THEIR ABILITY.

AND SO I THINK WE HAVE AN OBLIGATION TO MAKE THAT POSSIBLE BY CREATING GOOD SUPPORTS AND STRUCTURES. WE SHOULD DELINEATE TRAINING REQUIREMENTS BASED ON THOSE ROLES AND RESPONSIBILITIES, SO THAT IF THERE ARE SPECIAL REQUIREMENTS FOR MATH, WE DON'T FORCE THE READING TEACHERS TO GO THROUGH THOSE TRAINING REQUIREMENTS, ALTHOUGH TYPICALLY THEY ARE THE SAME, AND PROVIDE PRACTICE VERSIONS OF THE APPLICATIONS TO MINIMIZE ERRORS IN ADMINISTRATION, BOTH ON THE PARTS OF STUDENTS AND TEACHERS.

WE SHOULD ESTABLISH HELP DESK SUPPORTS THAT ARE CONSISTENT WITH THE LONGER TEST WINDOWS THAT ARE REQUIRED FOR COMPUTER-BASED TESTING SO THAT QUESTIONS CAN BE RESOLVED QUICKLY AND ERRORS AREN'T MADE INADVERTENTLY. AND WE SHOULD CONDUCT USER ACCEPTANCE TESTING IN THE SCHOOLS BECAUSE SOMETIMES WHAT APPEARS TO BE AN IRREGULARITY CAUSED BY THE SCHOOL IS AN IRREGULARITY IN THE WAY THE APPLICATION FUNCTIONS. AND THAT WAS POTENTIALLY DUE TO A LACK OF QUALITY CONTROL.

WE SHOULD BE AWARE OF WHICH ADULTS CAN BE IN THE SECURED TESTS. SO THINKING ABOUT THEM, THE LOCAL SYSTEM AS A PARTNER, BE AWARE OF WHICH ADULTS CAN BE IN THE SECURE TESTING ENVIRONMENT. AND WE SHOULD USE THE OPPORTUNITIES FOR SAMPLE TESTS AND APPLICATIONS WHEN THEY ARE AVAILABLE; PROVIDE CLEAR EXPECTATIONS FOR WHICH INDIVIDUALS MUST ATTEND THOSE TEST TRAININGS; AND PROVIDE A CLEAR PATH FOR IDENTIFICATION AND RESOLUTION OF PROBLEMS.

MOST IMPORTANTLY, BE AWARE THAT COMPUTER-BASED TESTING COULD BE OVERWHELMING FOR NEW TEACHERS AND SUBSTITUTES. BY FAR, IN MY EXPERIENCE, THAT'S WHERE MANY OF THE ERRORS OCCUR. AND THAT'S DUE TO A LACK OF TRAINING.

DON'T EXPOSE STUDENT IDENTIFIERS INADVERTENTLY. AND THEN BECAUSE IT'S SO IMPORTANT, I REPEAT IT AGAIN. PROVIDE A CLEAR PATH FOR IDENTIFICATION AND RESOLUTION OF PROBLEMS.

IN TERMS OF SYSTEM DESIGN--I'LL TAKE THE LAST FEW MINUTES TO TALK ABOUT HOW THE NEXT GENERATION OF ASSESSMENTS ARE GOING TO ATTEND TO SOME OF THESE ISSUES--IN SBAC, WE'RE GOING TO USE A COMPUTER ADAPTIVE TEST WITH LARGE ITEM POOLS BASED ON ITEM TEMPLATES THAT WILL ALLOW US TO RAPIDLY CREATE A WIDE VARIETY OF ITEMS THAT WILL MITIGATE THE RISKS, BUT, AS HAS BEEN STATED BEFORE, WON'T NECESSARILY SOLVE IT.

WE'RE GOING TO MINIMIZE THE NEED FOR ADULT SUPPORTS BY BUILDING ESSENTIAL FUNCTIONS INTO THE TESTING SYSTEM. ONE OF THE BIGGEST PROBLEMS WE HAVE IN MATHEMATICS IS WE ASK TEACHERS TO READ ITEMS ALOUD TO SOME KIDS OFTEN AND THEN WE TELL THEM NOT TO REMEMBER THE ITEMS. AND SO THAT'S REALLY NOT PARTICULARLY FAIR. AND IT DEFINITELY GIVES A MIXED MESSAGE. SO TO THE EXTENT THAT THE COMPUTER CAN TAKE THAT ROLE, IT ALLOWS HUMANS TO BE HUMANS AND ACT ACCORDINGLY BECAUSE ASKING THEM NOT TO REMEMBER SOMETHING THAT THEY READ IS REALLY DIFFICULT TO DO.

WE'RE GOING TO CONTROL ITEM EXPOSURE BASED, IN PART--BASED ON THE ITEM TYPE. SO THOSE ITEMS THAT ARE MORE MEMORABLE ARE PROBABLY

GOING TO BE SEEN BY FEWER KIDS AND ROTATED OUT OF THE ITEM BANK FASTER.

TO GREG'S EARLIER POINT, AS WELL, WE'RE GOING TO BE PROVIDING ASSESSMENTS THAT SERVE SPECIFIC PURPOSES. SO IN THE SBAC THEORY OF ACTION, WE HAVE A FORMATIVE AND AN INTERIM ASSESSMENT, WHICH ARE OPEN, NONSECURE, INTENDED TO HELP TEACHERS UNDERSTAND WHAT THE ASSESSMENT WILL LOOK LIKE, ALLOW THEM TO ASSESS STUDENTS IN ADVANCE SO THAT THEY'RE NOT SURPRISED BY THE RESULTS. AND THEN THE SUMMATIVE CAN REMAIN SECURE AND SERVE THE PURPOSE OF ACCOUNTABILITY, SO THAT THERE ISN'T AS MUCH INCENTIVE FOR CURIOSITY ABOUT WHAT THE QUESTIONS ARE.

WE'RE GOING TO BUILD A SYSTEM ARCHITECTURE THAT HAS A PREMISE OF ENTERPRISE-LEVEL SECURITY. SO IN SBAC'S CASE, WE'VE POSTED OUR FIRST DRAFT OF OUR SYSTEM ARCHITECTURE, WHICH WAS BUILT ON THE ASSUMPTION OF SECURITY AT THE LEVEL THAT WE WOULD EXPECT OF THE BANKING INDUSTRY. AND WE WILL EXPECT THAT ALL OF OUR APPLICATIONS WILL ALSO MAINTAIN THAT PREMISE.

AND THEN AS MENTIONED EARLIER, WE'RE GOING TO PROVIDE SECURE METHODS OF ACCESSING THE TEST TO MINIMIZE THE CHANCES FOR INADVERTENT IRREGULARITIES.

**JACK BUCKLEY:** THANK YOU VERY MUCH. ALL RIGHT. SO, AGAIN...I'M HEARING SOME COMMONALITIES BUT ALSO MAYBE SOME POINTS THAT THE PANEL DOESN'T ENTIRELY AGREE ON. I THINK I WANT TO START WITH SOMETHING THAT'S COME UP AT LEAST IN THREE OF YOUR REMARKS, WHICH IS THE ISSUE OF THE TESTING WINDOW. AND SO ONE THING WE KNOW, CERTAINLY WE ALL SEEM TO BE ASSUMING TO BE TRUE, IS THAT WHAT WE SEE RIGHT NOW WHEN WE MOVE TO COMPUTER-BASED TESTING AND WHAT WE'RE LIKELY TO SEE IN THE FUTURE IN '14, '15 AND A FEW YEARS BEYOND THAT WOULD BE A GREATLY INCREASED TESTING WINDOW CERTAINLY WITH RESPECT TO WHAT JIM WAS TALKING ABOUT IN NEW YORK, WHERE WE COULD ADMINISTER PAPER AND PENCIL ALL IN THE SAME DAY. HOW FEASIBLE IS IT THAT THE TECHNOLOGY CAN CATCH UP IN TIME TO REDUCE THIS PROBLEM? I MEAN, THIS IS A PROBLEM THAT IS NOT INHERENT IN COMPUTER-BASED TESTING NECESSARILY IF YOU HAD SUFFICIENT COMPUTER RESOURCES. OR IS THAT NOT TRUE? OR IS THAT JUST NOT PRACTICAL?

**WES BRUCE:** THE OBVIOUS SOLUTION IS ONE-TO-ONE, YOU KNOW. PRACTICALLY, THOUGH, IN SCHOOLS, ESPECIALLY AS WE RAMP UP TO ONLINE, THERE IS GOING TO BE THAT COMPROMISE THAT WE HAVE TO THINK ABOUT HOW WE MITIGATE AND PROVIDE LOTS OF SECURITY, MAYBE LOTS OF FORMS, MORE FORMS INITIALLY, BECAUSE IN MOST CASES, THE WINDOW IS GOING TO BE LONGER. FOR SOME OF US WHO HAVE WALKED ON THIS PATH ALREADY, WE'VE GIVEN UP ONE OF THE BASIC TENETS OF SECURITY THAT YOU HEARD ABOUT EARLIER, IS THAT IN THE PAPER WORLD, GIVING THE

TEST TO ALL OF THE FOURTH GRADERS IN A SCHOOL AT THE SAME TIME IS VERY, VERY SIMPLE. IF YOU HAVE, HOWEVER, 4 FOURTH GRADES AND ONE COMPUTER LAB, GIVING ALL OF THOSE STUDENTS THE SAME TEST ONLINE AT THE SAME TIME IS IMPOSSIBLE. AND SO YOU HAVE TO THINK ABOUT--AS TONY TALKED ABOUT-- THE LOGISTICS OF THAT CHANGE. AS PEOPLE HAVE MOVED FROM THE PAPER WORLD TO ONLINE WORLD, THEY'VE NEVER HAD TO WORRY ABOUT SCHEDULING MULTIPLE GROUPS OF STUDENTS IN THE SAME GRADE. AND THEY GET CONCERNED ABOUT THE POTENTIAL FOR EXPOSURE. AND SO, YOU KNOW, I THINK FROM A PRACTICAL STANDPOINT, IT IS GOING TO BE AN ISSUE THAT WE HAVE TO CONFRONT UNTIL TECHNOLOGY BECOMES MUCH MORE WIDELY PROLIFERATED, MOST IMPORTANTLY FOR INSTRUCTION. BUT THEN WHEN IT COMES TIME TO ASSESS, THAT WE CAN LEVERAGE THAT TECHNOLOGY FOR ASSESSMENT.

**TONY ALPERT:** AND I AGREE WITH WES. I WOULD ALSO ADD, THOUGH, THAT EVEN IN A ONE-TO-ONE SITUATION, THERE ARE BANDWIDTH LIMITATIONS. SO THAT ISN'T THE SOLE SOLUTION. AND I THINK TO THE EARLIER POINTS, THERE'S ALWAYS A WAY AROUND SECURITY. SO IT'S REALLY CREATING ALL-- TO MITIGATE THE RISKS, WE CAN HOPE TO MITIGATE THE RISK AND CONTINUE TO ADD MORE CONTROLS AND MITIGATE THE RISK MORE, BUT PEOPLE ARE VERY INVENTIVE. SO IF THEY WANT TO GET AROUND THE SYSTEM, THEY'LL FIGURE OUT A WAY.

**WAYNE CAMARA:** I WOULD SAY THAT-- I PERSONALLY BELIEVE THAT, ASSUMING YOU RUN INTO A REALLY GOOD TESTING PROGRAM AND YOU'VE GOT THE SECURITY AND YOU'VE DONE THE RIGHT THINGS, THAT THE ISSUE OF THE TESTING WINDOW COULD PRESENT THE SINGLE LARGEST RISK TO C.B.T., ESPECIALLY TO HIGH SCHOOL. AND IF THOSE STAKES ARE VERY HIGH TO A STUDENT, MEANING THAT YOU DO NOT NEED TO TAKE ANY DEVELOPMENTAL COURSE OR YOUR HIGH SCHOOL DIPLOMA DEMANDS YOU PASS THE EXAM, I THINK IT'S AN UNACCEPTABLY HIGH RISK. I THINK YOU DO NOT NEED A ONE-TO-ONE, BUT AT LEAST A ONE-TO-FOUR, MEANING THAT IN A HIGH SCHOOL, YOU NEED ENOUGH COMPUTERS TO OUTFIT THE NINTH-GRADE CLASS. YOU DON'T NEED ONE COMPUTER FOR EVERY STUDENT, BUT YOU NEED ONE FOR FOUR, ONE FOR THREE. AND I THINK UNTIL YOU HAVE THAT, YOU'LL BEGIN COMPROMISING OTHER ASPECTS OF THE TEST. SO FOR EXAMPLE, JUST THOUGHT LOGIC, WE HAVE GOT THE ADVANCED PLACEMENT EXAM THAT WE ADMINISTER NATIONALLY IN HIGH SCHOOLS EVERY YEAR. AND MONDAY, MAY 3RD, IN THE MORNING, MAYBE CALCULUS A-B, AND THAT SAME DAY, IN THE AFTERNOON IS BIOLOGY. AND IF WE WERE TO BE REQUIRED TO OFFER THAT TEST ON THREE OR FOUR DAYS, NOW WE MUST USE DIFFERENT CONSTRUCTED RESPONSE TASKS. WE ALSO KNOW THAT THEY MAY NOT BE PARALLEL. SO WE MAY BE GIVING YOU A FORM THAT THE MULTIPLE-CHOICE ITEMS ARE EQUATED QUITE WELL BUT THE CONSTRUCTED RESPONSE TASKS ARE A LITTLE BIT EASIER ON FORM "B" THAN THEY ARE ON FORM "A." WE COULD ALWAYS EQUATE THE TASKS, BUT THEN WE'D BE EXPOSING THEM. SO, IN ORDER TO MAINTAIN SECURITY, YOU BEGIN TO DIMINISH THE PSYCHOMETRIC

QUALITIES OF THE TESTING PROGRAM TO SUCH A LEVEL THAT THE PROGRAM REALLY DOES NOT HAVE THE RIGOR OF A GOOD PAPER PROGRAM. AND I THINK THAT'S SOMETHING WE WANT TO GUARD AGAINST.

**JOHN FREMER:** WE'RE SUPPOSED TO BE SOLVING PROBLEMS, NOT SIMPLY AMPLIFYING THEM, BUT I SPENT LAST WEEK DOING A SECURITY AUDIT WITH A COLLEAGUE WHO USED TO BE EXECUTIVE DIRECTOR FOR COMPUTER-BASED TESTING AT E.T.S. AND I WAS DESCRIBING THIS SYMPOSIUM AND C.B.T. AND THE CONSORTIA. AND HE SAID, "WELL, THERE ARE THREE PROBLEMS THAT THEY HAVE TO SOLVE," TWO OF WHICH HAVE ALREADY BEEN MENTIONED--COMPUTERS AND BANDWIDTH. HE SAID, "YOU REALIZE THAT IN A LOT OF SCHOOLS AROUND THE COUNTRY, THERE JUST ISN'T ENOUGH ELECTRICAL HOOKUPS TO DO IT." SO NOT THAT WE--I HAVE A FEELING THAT MIGHT BE THE EASIEST OF THE THREE TO SOLVE. BUT RIGHT NOW WITH HAVING TO REACH ALL THOSE STATES AND ALL THOSE DISTRICTS WITHIN STATES, SOME OF WHICH HAVEN'T HAD MONEY TO BUY EQUIPMENT, I UNDERSTAND, FOR YEARS AND YEARS, IT'S A CHALLENGE.

**JACK BUCKLEY:** TONY, YOU HAD A RESPONSE TO WAYNE'S COMMENTS EARLIER?

**TONY ALPERT:** SO I THINK THERE ARE RISKS. IN MY EXPERIENCE, THOUGH, WE ARE ABLE TO EQUATE CONSTRUCTIVE RESPONSE IN SUCH A WAY USING A COMPUTER ADAPTIVE TEST AT LEAST AND PRE-EQUATING. AS LONG AS THE FIELD TEST IS COMPREHENSIVE ENOUGH AND WIDESPREAD ENOUGH, I THINK

THE EXPOSURE RISKS ARE MITIGATED SIGNIFICANTLY. AND I HAVE NEVER EXPERIENCED A PAPER TEST THAT PROVIDED THE ADVANTAGES OF COMPUTER-BASED TESTS. SO IN MY EXPERIENCE, A BAD COMPUTER TEST IS STILL BETTER THAN A GOOD PAPER TEST, ESPECIALLY A COMPUTER-ADAPTIVE TEST IN VIRTUALLY EVERY CIRCUMSTANCE.

**JACK BUCKLEY:** SO MOVING ON TO A SECOND ISSUE THAT CAME OUT OF THE DISCUSSION. SO LET'S ASSUME THAT THE TESTING WINDOW PROBLEM IS MINIMIZABLE. WE LISTEN TO YOU, WAYNE, AND WE SAY, "YOU KNOW WHAT? THIS IS THE BIGGEST PROBLEM," SO WE THROW THE RESOURCES AT IT AND WE ACTUALLY GET ELECTRICAL POWER WHERE WE NEED IT AND GET THE MACHINES THERE. WHAT ABOUT THE INNOVATIVE ITEM FORMAT? SO WHEN WE'RE LOOKING AT THIS NEXT GENERATION OF ASSESSMENT, ONE OF THE PROBLEMS THAT I THINK SEVERAL OF YOU, ESPECIALLY, JOHN, YOU MENTIONED, WE'RE NOW MOVING BEYOND SORT OF SIMPLE, AUTOMATED--YOU COULD IMAGINE--A.I.-DEVELOPED MULTIPLE-CHOICE ITEMS TO SOME REALLY COMPLEX HANDS-ON COMPUTER TASKS OR REALLY INTERESTING SIMULATIONS. AND THERE ARE SORT OF TWO OBVIOUS PROBLEMS THERE. THE FIRST IS THAT THEY'RE MEMORABLE, AND THE SECOND IS THAT THEY'RE REAL EXPENSIVE TO DEVELOP. SO YOU PROBABLY DON'T HAVE A DEEP POOL OF THEM. WHAT DO WE DO? HOW DO WE FIX THAT PROBLEM? BECAUSE, I MEAN, THE DEPARTMENT, CERTAINLY, WE ARE INTERESTED IN ENSURING THAT WE CONTINUE TO DRIVE INNOVATION, AND THAT INCLUDES IN NAEP. AND SO WE HAVE THIS PROBLEM, AS WELL.

**TONY ALPERT:** SO, ONE OF THE ADVANTAGES OF THOSE INTERACTIVE ITEM TYPES IS THAT THE CONSTRUCT THAT YOU'RE PROBABLY MEASURING IS PROCESS. SO, YOU KNOW, AT THE LEAST, IF KIDS ARE LEARNING THAT AND THEY'RE ABLE TO RESPOND TO THAT, THEY'VE PROBABLY LEARNED IT, WHICH ISN'T NECESSARILY A BAD THING. YOU CAN VARY, IF YOU ORGANIZE THESE MEMORABLE TESTS WELL AND YOU VARY SOME OF THE CONTENT WITHIN THE TEMPLATE, SUCH THAT THERE ARE SOME UNIQUE VARIABLES THAT ARE INCLUDED THAT ARE ADJUSTED ENOUGH SO THEY CAN'T MEMORIZE THE WHOLE TASK, THEN THERE STILL IS—YOU STILL ARE ABLE TO GENERALIZE THE SKILL TO THE LARGER BODY OF KNOWLEDGE. AND THAT ITEM STILL IS ABLE TO EVALUATE THE DEGREE TO WHICH THE STUDENT HAS MASTERED THAT CONTENT. BUT, CERTAINLY, IT IS A RISK. CERTAINLY, WE WILL HAVE TO CONTROL THE EXPOSURE OF THOSE ITEM TYPES A LITTLE BIT MORE CAREFULLY. AND THE RISK IS GREATER EARLY ON IN THE ASSESSMENT PROGRAM. IT WILL BE MITIGATED LATER AS WE ADD MORE AND MORE ITEMS TO THE POOL.

**WAYNE CAMARA:** I WOULD JUST SAY THAT THE--IT DEPENDS ON THE TYPE OF ITEM. I THINK THERE ARE A LOT OF REALLY INNOVATIVE ITEMS THAT THE COMPUTER CAN DELIVER THAT ARE OBJECTIVE ITEMS AND THAT ARE VERY SHORT, CONSTRUCTED RESPONSE ITEMS. AND I THINK IN THOSE INSTANCES, YOU DON'T HAVE TO SACRIFICE HAVING MULTIPLE ITEMS. YOU CAN DEVELOP A LOT. I THINK WHEN YOU'RE TALKING ABOUT SCENARIOS, WHEN YOU'RE

VIDEOTAPING, FOR EXAMPLE, A SIMULATION--WE DO A LOT OF THAT IN CERTIFICATION, THE F.A.A. TESTS. SOME REALLY GOOD SIMULATIONS. I DON'T THINK YOU CAN AFFORD TO DEVELOP A LOT OF THESE. AND SO ULTIMATELY, I WOULD SAY, IF YOU ARE SOLVING THE BANDWIDTH SOLUTION AND THE COMPUTER SOLUTION, I REALLY AM A FAN OF ONE AND DONE...BECAUSE IF YOU'RE GOING TO BE USING EXTENDED CONSTRUCTED RESPONSE TASKS THAT REQUIRE ME TO DO WHAT I DO IN COLLEGE, WHICH IS TO GO ON A LIMITED WEBSITE; TO LOOK AT ORIGINAL SOURCE DOCUMENTS; TO BE EVALUATING THEM AND WRITING AN ESSAY AND YOU'RE GIVING ME 60 TO 90 MINUTES TO DO THAT TASK, IT'S REALISTIC; IT HAS HIGH FIDELITY; IT'S WHAT I'M GOING TO DO IN COLLEGE. AND WE CANNOT AFFORD TO DEVELOP 50 OF THOSE TASKS. AND IF WE ARE, I DON'T BELIEVE THEY WILL BE STATISTICALLY PARALLEL NO MATTER HOW MUCH EQUATING AND PRE-EQUATING WE DO. BUT I DO BELIEVE WITH ONE AND DONE, YOU DON'T HAVE TO WORRY ABOUT THAT. YOU HAVE TO HAVE A MAIN FORM, MAYBE TWO MAIN FORMS AND A MAKEUP. SO, AGAIN, I PUT IT BACK TO, FOR THE REAL HIGH FIDELITY AND INNOVATIVE ITEMS THAT REQUIRE A LOT OF CONSTRUCTED RESPONSE, I DO BELIEVE THAT IF WE CAN MINIMIZE THAT WINDOW, WE WILL HAVE MITIGATED THESE OTHER ISSUES DRAMATICALLY, AND IT WON'T BE A CONCERN.

**JOHN FREMER:** IT STRIKES ME AS POSSIBLE-- AND I HAVEN'T HAD TO WRESTLE WITH THESE PROBLEMS THE WAY MY COLLEAGUES HAVE--THAT THIS ISSUE OF THE LARGE WINDOW, THERE MAY BE WAYS TO HAVE A QUITE SHRUNKEN WINDOW

FOR CERTAIN TYPES OF THINGS AND A LONGER WINDOW FOR OTHERS, MAYBE SOME COMBINATION OF THE KINDS OF THINGS WAYNE WAS TALKING ABOUT-- DIFFERENT SIZE WINDOWS. MY SENSE IS THAT THIS IS A SOLVABLE PROBLEM. IT'S JUST I WISH...IT DIDN'T HAVE TO BE SOLVED SO SOON. IF WE HAD, YOU KNOW, FIVE OR SIX YEARS--I'M SURE THEY THINK THAT ALL THE TIME, BUT I'M ALLOWED TO THINK IT BECAUSE I DON'T HAVE ANY CONTRACT THAT HAS ANYTHING TO DO WITH IT.

**JACK BUCKLEY:** LET ME OPEN UP NOW AGAIN TO THE PANEL AND THE BROADER AUDIENCE. STEVE?

**STEVE FERRARA:** SORRY, FOLKS. I HAVE A QUESTION ON THIS. WHAT ABOUT OTHER DEVICES? I THINK IT'S A POSSIBLE SOLUTION FOR ONE OF THE PROBLEMS. YOU DON'T HAVE ENOUGH DEVICES, YOU DON'T HAVE ENOUGH ELECTRICAL OUTLETS, CORNELIA ORR HAS A GREAT STORY ABOUT ELECTRICAL OUTLETS IN FLORIDA SCHOOLS THAT SHE TOLD ME A COUPLE YEARS AGO. ANYWAY, YOU DON'T HAVE ENOUGH ELECTRICAL OUTLETS. I KNOW BANDWIDTH IS AN ISSUE. TONY MENTIONED THAT. BUT YOU CAN SOLVE THE DEVICES ISSUES MUCH MORE CHEAPLY THAN DESKTOPS. SO WHAT DO YOU WANT TO SAY ABOUT THE USAGE, OTHER DEVICES, INCLUDING STUDENT-OWNED DEVICES?

**WES BRUCE:** THEY'RE EVIL.

[LAUGHTER]

YEAH...I THINK... ONE OF THE ISSUES THAT BOTH CONSORTIA ARE WRESTLING WITH AND NOT JUST THE CONSORTIA, BUT STATES ARE WRESTLING WITH IS THE EXPLOSION OF THE ADOPTION OF DEVICES THAT DIDN'T EXIST, WHAT, 24 MONTHS AGO. YOU KNOW, THE WHOLE REALM OF TABLET DEVICES AND SMART PHONES. BUT I THINK TABLET DEVICES ARE PUSHING US MORE. AND WHILE IT'S POSSIBLE, ONE OF THE THINGS THAT WE'RE MOST CONCERNED ABOUT IS THE VALIDITY IN ANSWERING THE COMPARABILITY QUESTIONS.

A TABLET DEVICE OR A DEVICE THAT DOESN'T HAVE A KEYBOARD BRINGS UP A SET OF COMPARABILITY ISSUES THAT HAS TO BE ANSWERED. I DON'T KNOW IF IT HAS TO BE ANSWERED A COUPLE TIMES, BUT IT HAS TO BE ANSWERED AT LEAST ONCE--ABOUT WHETHER OR NOT INPUTTING ON AN ON-SCREEN KEYBOARD PUTS A STUDENT AT AN ADVANTAGE OR DISADVANTAGE TO ALL THOSE OTHER KIDS WHO ARE USING FOUR- AND FIVE-YEAR-OLD TECHNOLOGY WHERE THEY HAVE GOT A PHYSICAL KEYBOARD. I'VE SEEN SOME RESEARCH THAT TALKS ABOUT THAT THERE IS AN IMPACT CURRENTLY FOR STUDENT PERFORMANCE BASED ON HOW MUCH THEY HAVE TO SCROLL. NOW, I DON'T BELIEVE THAT THAT IS NECESSARILY GOING TO BE THE CASE THREE OR FOUR OR FIVE YEARS FROM NOW. BUT AS LONG AS I'VE GOT EVIDENCE AS SOMEONE WHO'S HELPING DESIGN A TEST PROGRAM THAT IF YOU SCROLL MORE, YOUR SCORE IS LOWER, I'M GOING TO HAVE TO SAY, "I WANT A STANDARD SIZE

OF A DEVICE. AND I WANT TO BE ABLE TO DISPLAY EACH ITEM IN A SIMILAR WAY ACROSS THOSE DEVICES. YOU MAY HAVE A LOT OF EXTRA REAL ESTATE IF YOU'RE LOOKING AT A SCREEN LIKE ONE OF THESE AND YOU'RE TAKING YOUR TEST. BUT I THINK, YOU KNOW, WE'RE LOOKING AT SETTING A MINIMUM SCREEN SIZE AT LEAST INITIALLY. I CAN'T PREDICT WHERE WE WILL BE IN EIGHT OR 10 YEARS. AS JOHN'S POINTED OUT, WE HAVE GOT ALL THIS TIME BETWEEN NOW AND '14 AND '15 TO SOLVE THESE ISSUES, BUT I THINK WE ALSO CAN'T HIDE OUR HEADS IN THE SAND, THAT WE HAVE TO THINK ABOUT THE USE OF THOSE DEVICES. AND I KNOW WE'RE BOTH ACTIVELY-- BOTH CONSORTIA ARE ACTIVELY LOOKING AT, HOW ARE WE GOING TO DELIVER ON TABLETS BECAUSE CURRENTLY THERE AREN'T MANY PROGRAMS THAT ARE ACTUALLY USING THE TABLETS? MY TESTING PROGRAM RIGHT NOW, NO TABLET IS SUPPORTED ON IT.

WE ARE LOOKING AT HOW TO DO THAT. BUT AS YOU POINTED OUT, THERE ARE A HOST OF ISSUES, NOT ONLY THE SECURITY ISSUE YOU POINTED OUT, BUT ALSO, HOW MUCH OF THE REAL ESTATE YOU CAN ACTUALLY ACCESS AND CONTROL VERSUS THOSE KIND OF IN-BUILT NATIVE CONTROLS THAT ARE THERE.

**TONY ALPERT:** JUST TO ADD ON TO WHAT WES WAS SAYING, ONE OF OUR CHARGES AS CONSORTIA IS TO ASSES THE FULL BREADTH AND DEPTH OF THE COMMON CORE AND GET AT DEEP KNOWLEDGE AND SKILLS ASSOCIATED WITH CAREER AND COLLEGE READINESS. AND TO THE EXTENT THAT CAN BE

ACCOMPLISHED ON A SMALLER SCREEN SIZE, WE'RE INTERESTED IN INVESTIGATING THAT, BUT VALIDITY IS PARAMOUNT. AND IT GOES BEYOND JUST EVEN TYPING. SO AN EXAMPLE FROM A SCHOOL DISTRICT IN OREGON, THEY TRIED IMPLEMENTING NET BOOKS. AND IT WAS RELATIVELY SUCCESSFUL FOR THE THIRD-GRADERS. BUT HIGH SCHOOL FOOTBALL PLAYERS AND THEIR BIGGER FINGERS DIDN'T REALLY WORK WITH THE KEYBOARD, SO THERE'S AN INTERACTION BETWEEN THE STUDENT AND THE DEVICE THAT HAS TO BE CONSIDERED. AS PART OF THE APPROACH TO EVALUATING MULTIPLE DEVICES, WE'RE GOING TO BE CONDUCTING COGNITIVE LABS TO INVESTIGATE, WHAT ARE THE IMPLICATIONS FOR DIFFERENT ITEM TYPES AND THE MEDIUM THAT STUDENTS WILL USE TO INTERACT WITH THOSE, AND FIND OUT FROM THE STUDENTS HOW WELL IS IT WORKING. ARE WE REALLY GETTING THE ANSWER FROM YOU THAT YOU TRIED TO GIVE US, WHICH IS AN IMPORTANT PART OF THE DESIGN PROCESS.

**JACK BUCKLEY:** OTHER QUESTIONS? SIR.

**AUDIENCE MEMBER:** WITH PAPER AND PENCIL TESTS, AS THEY ROLL IN, SORT OF, TO THE VENDOR, YOU HAVE EITHER DOING EQUATING OR RUN STAT ANALYSIS, IS REALLY SCORING THE ACTUAL FORMS, YOU HAVE THE OPPORTUNITY TO INVESTIGATE THE INDICES FOR POSSIBLE COLLABORATIVE EFFORTS OR ANSWER COPYING. AND I'M SURE DR. WHITEHEAD CAN ATTEST TO THIS, THAT IT GIVES YOU SORT OF THIS WINDOW OF OPPORTUNITY WHERE YOU CAN REALLY INVESTIGATE SOME OF THESE BREACHES IN TEST SECURITY.

WHEN YOU'RE GOING TO COMPUTERIZED-BASED TESTING OR C.A.T., THE PRESSURE IS PROBABLY GOING TO EXIST TO HAVE SCORES IMMEDIATELY REPORTED TO THE USER EITHER FOR REMEDIATION OR FOR WHATEVER PURPOSES. SO YOU'RE CUTTING OUT THAT POTENTIAL PREPROCESSING WINDOW IN WHICH YOU COULD HOLD SCORES AND NOT RELEASE THEM NECESSARILY TO THE EXAMINEE, SO TO THE STATE. AND SO NOW YOU'RE LEFT WITH POTENTIALLY HAVING TO CANCEL SCORES WHEN YOU'RE THEN DOING-- THE ANALYSES LATER RATHER THAN HOLDING THEM FOR RELEASE. SO I WAS JUST WONDERING ABOUT YOUR OPINIONS--HOW YOU ENVISION THIS-- WHETHER YOUR ADVICE WOULD BE TO NOT EVER RECOMMEND FOR RELEASING SCORES RIGHT AWAY. I DON'T KNOW HOW YOU'RE SEEING, SORT OF, TEST SECURITY IN TERMS OF PREPROCESSING RATHER THAN THEN HAVING TO PAY FOR REISSUING SCORES.

**WAYNE CAMARA:** IT'S A LITTLE BIT DIFFERENT. WHAT WE KNOW IN SOCIAL PSYCHOLOGY IN TERMS OF THE IMPLICIT PRESSURE. AND FOR A STATE DEPARTMENT OR FOR A PRINCIPAL OR SUPERINTENDENT, I THINK THERE IS A DIFFERENCE IN TERMS OF PROCESSING A COMPLAINT OR A SUSPICIOUS--OR A CHARGE THAT COMES TO ME AFTER THE TEST WAS ADMINISTERED TWO DAYS AGO AND TWO MONTHS BEFORE THE SCORES ARE GOING TO BE RELEASED. AND WHAT WILL I DO WITH THAT? WHAT IS MY INTENTION TO DO VERSUS AFTER THE FACT WHEN I'VE REPORTED THE DATA TO MY SCHOOL BOARD WHEN THE SCORES HAVE GONE OUT AND ACTIONS HAVE TAKEN? AND, UNINTENTIONALLY--WE'RE HUMAN BEINGS AND WE PROBABLY HAVE DIFFERENT MOTIVATIONS

BASED ON THOSE TWO SCENARIOS-- SO MY CONCERN WOULD BE NOT WITH  
RELEASING SCORES BECAUSE I THINK THERE'S A WAY TO RELEASE THEM AS  
DRAFT SCORES OR TENTATIVE SCORES--BUT MY CONCERN WOULD BE, THEN,  
WHAT WOULD THAT SAY ABOUT OUR PROCLIVITY OR OUR WILLINGNESS TO  
INVESTIGATE CERTAIN ACTIVITIES? I THINK IT WOULD BE DIFFERENTIAL.  
AND THAT'S WHY I THINK IT WOULD BE MUCH MORE IMPORTANT IN THAT  
SCENARIO TO HAVE INDEPENDENT GROUPS APART FROM THE SCHOOL  
ADMINISTRATORS AND CERTAINLY APART FROM THE ASSESSMENT FOLKS IN A  
STATE DEPARTMENT OF EDUCATION, AS WE'VE SEEN TODAY. BUT SOME  
INDEPENDENT GROUP WHO PROCESSES THOSE COMPLAINTS OR  
INVESTIGATIONS, BECAUSE OTHERWISE I'M WEARING TWO HATS. I'M  
WORRIED ABOUT THE INTEGRITY OF MY PROGRAM OR MY SCHOOL, AND NOW I'M  
HEARING THESE RUMORS. THEY SOUND UNFOUNDED. AND I WANT THEM TO GO  
AWAY. AND I WANT THEM TO GO AWAY MUCH MORE AFTER THE SCORES HAVE  
BEEN REPORTED IN THE MEDIA AND TO PARENTS THAN I WOULD IF THEY WERE  
NOT RELEASED.

**JOHN FREMER:** ...MY PERSPECTIVE, IT'S BETTER NOT TO REPORT SCORES AS  
OF THE TIME OF TAKING THE TEST. AND IF YOUR MODEL IS, WELL, WE'RE  
TRANSITIONING FROM A PAPER AND PENCIL STATE ASSESSMENT PROGRAM,  
THEY'RE NOT USED TO GETTING THEIR SCORES RIGHT AWAY. THEY'RE USED  
TO WAITING. THE PROBLEM OCCURS IF YOU HAVE A PROGRAM AND IT ALREADY  
IS REPORTING RIGHT AWAY. THEN YOU THINK THAT'S THE ONLY WAY IT CAN  
BE. WE DO A LOT OF WORK AT CAVEON WITH THE I.T. INDUSTRY. AND THEY

REPORT RIGHT AWAY. AND THEY PROBABLY HAVE THE MOST TROUBLE OF ANY INDUSTRY WITH PROBLEMS WITH THE CORRUPTION OF SCORES. WE DON'T HAVE THAT MUCH OF A TROUBLE IN EDUCATION. I ESTIMATE IT'S ONLY MAYBE 1% OR 2%. SOMETIMES YOU SEE LARGER ESTIMATES, BUT YOU DON'T HEAR, OTHER THAN AUDREY'S SUMMARY, YOU DON'T HEAR 10% OR 20%. IN THE I.T. INDUSTRY, SOME OF THEM, 50% OF THEIR SCORES SHOW SOME EVIDENCE OF COMPROMISE. LET'S NOT IMITATE THEM AND THEIR IMMEDIATE SCORE REPORTING.

**DAVID FOSTER:** BUT I JUST WANT TO ADD I DON'T THINK REPORTING SCORES IMMEDIATELY IN THAT PARTICULAR EXAMPLE CONTRIBUTED TO THE PROBLEMS THAT THAT INDUSTRY HAS. IT WAS REALLY POOR TEST ADMINISTRATION PROCEDURES. WHEN YOU HAVE COMPUTERIZED TESTS, PEOPLE ARE GOING TO EXPECT YOU TO BE A LITTLE FAST WITH YOUR RESPONSES. THEY'RE NOT GOING TO, "WELL, WE'RE GOING TO WAIT THREE WEEKS" OR SIX WEEKS OR WHATEVER. NAH. YOU KNOW, IT'S NOT GOING TO FLY. THIS IS AN IMMEDIATE WORLD WE LIVE IN. SO I THINK WE NEED TO--I LIKE THE IDEA OF PROVISIONAL SCORES, WHICH ALLOWS YOU IN A CERTAIN PERIOD OF TIME--PROBABLY DAYS-- TO MAKE A DECISION THAT THAT TEST IS VERIFIED OR NOT AND MAKE A CHANGE, BUT IT'S GOING TO HAVE TO BE DONE FAIRLY QUICKLY.

**JACK BUCKLEY:** WELL, I'LL ASK ONE MORE, WHICH ACTUALLY GETS BACK TO A POINT THAT, JOHN, YOU MADE, WHICH IS THE ISSUE ABOUT, AGAIN, NOW

LOOKING FOR IT--I GUESS IT'S FAIR TO SAY, SO, REALLY, WE'RE TALKING ABOUT TWO THINGS HERE. AND WE'VE KIND OF CONFLATED THEM, OR AT LEAST I HAVE. I'M GUILTY OF THIS. COMPUTER-BASED TESTING OR ASSESSMENT BUT ALSO THE RACE TO THE TOP ASSESSMENTS OR THE COMMON CORE OR STATE STANDARDS ASSESSMENTS. AND, YES, BECAUSE THEY'RE INTERTWINED, IT'S EASY TO SORT OF CONFLATE THE ISSUES. JOHN, YOU RAISED THE QUESTION ABOUT PILOTING THE CHEATING DETECTION OR WHATEVER GOES INTO THE TECHNOLOGY IN THAT ARCHITECTURE AT THE SAME TIME AS ACTUALLY PILOTING THE ASSESSMENTS. AND I JUST WANTED TO ASK WES AND TONY IF YOU THOUGHT THAT WAS FEASIBLE BECAUSE IT SOUNDS LIKE A GOOD IDEA FROM WHERE I SIT.

**TONY ALPERT:** SO THE SMARTER BALANCED APPROACH, WE'RE GOING TO HAVE A PILOT WHICH IS RELATIVELY SMALL SCALE. IT'LL ONLY BE ABOUT 30,000 STUDENTS. I KNOW. THAT'S A WHOLE DIFFERENT WORLD. AND THEN THERE WILL BE A FIELD TEST. THE PILOT, BY ITS NATURE, SINCE THE ITEMS WON'T BE SCALED, IT WOULD BE HARD TO DETECT IRREGULARITIES. THE FIELD TEST, WE'RE PROBABLY GOING TO HAVE SOME PRELIMINARY SCALING THAT MIGHT BE USED AS A BASIS FOR EVALUATING CHANGES IN SCORES, BUT WE'RE NOT GOING TO HAVE THE BREADTH OF DATA THAT WE MIGHT OTHERWISE EXPECT IN AN OPERATIONAL ADMINISTRATION. SO I DON'T THINK WE COULD PILOT THE FULL BREADTH OF SECURITY MEASURES, BUT IT'S POSSIBLE TO DO SOME OF THEM.

**WES BRUCE:** YEAH. I THINK I WOULD AGREE THAT, YOU KNOW, PART OF THIS IS TO TRY TO START THINKING IN DIFFERENT WAYS THAN WE DO WITH PAPER, TO THINK ABOUT WHAT DATA IS IT POSSIBLE TO COLLECT IN "REAL TIME," WHETHER USING PROXY SERVERS OR WHETHER YOU'RE ACTUALLY DELIVERING THE TEST OVER THE INTERNET. CAN YOU COLLECT DATA THAT SHOWS EVIDENCE OF SOME KIND OF COLLUSION OR SOME KIND OF ADULT INTERVENTION WHILE THE TEST IS GOING ON. BECAUSE YOU'RE CAPTURING EVERY ITEM, YOU'RE DOING SOMETHING WITH IT. AND AT THE VERY LEAST, YOU'RE ENCRYPTING IT AND STORING IT. BUT YOU DO HAVE--YOU KNOW, YOU DON'T HAVE THAT BUILT-IN DELAY WHERE YOU'RE WAITING FOR EVERYTHING TO RUN THROUGH A SCANNER. SO I THINK TO TONY'S POINT, I DON'T THINK WE CAN DO THE FULL RANGE, BUT CERTAINLY SOME OF THOSE THINGS THAT WE HAVE NOT BEEN DOING, WE CAN START THINKING ABOUT AND DEVELOPING. AND IT ONLY MAKES SENSE TO TEST THOSE OUT IN THAT FIELD TESTING TIME TO SEE WHETHER OR NOT IT LOOKS LIKE THOSE ARE GOING TO WORK AT SCALE. OR SUDDENLY YOU FIND OUT THAT THOSE ARE HOGGING UP A BUNCH OF YOUR BANDWIDTH AND THEY'RE NOT GOING TO BE PRACTICAL TO DO IN REAL TIME. YOU MAY HAVE TO DO THEM OFF-LINE. BUT I THINK JOHN'S SUGGESTION IS A GOOD ONE AND CERTAINLY ONE THAT WE'LL WANT TO CONSIDER.

**JACK BUCKLEY:** ALL RIGHT. WELL, IF THERE ARE NO OTHER QUESTIONS, I GUESS WE WILL WRAP THIS PANEL UP. I THINK, AGAIN, LOOKING BACK AT WHAT I'VE HEARD, CERTAINLY THERE'S CONCERNS IN COMPUTER-BASED

TESTING ABOUT NOT ONLY WILL SOME THREATS TO TESTING INTEGRITY BE MITIGATED BUT OTHERS WILL BE CONSTANT OR INCREASED. AND I THINK THAT SEEMS TO BE WHAT'S EMERGED AS THE LARGEST AREA OF CONCERN IS THE TESTING WINDOW OR THE SIMPLE SIZE OF THE TESTING WINDOW, WHICH IS LARGELY DRIVEN BY LIMITATIONS OF TECHNOLOGY AND AVAILABILITY. AND TO THE POINT OF WHETHER OR NOT WE WOULD BROADEN THE POOL OF ACCEPTED TECHNOLOGY BY ALLOWING DIFFERENT DEVICES, NON-STANDARD DEVICES, OR STUDENTS' OWN DEVICES, WE WERE TOLD THAT WAS EVIL.

[LAUGHTER] BEYOND JUST THE TESTING WINDOW, CERTAINLY, THERE'S POTENTIAL PROBLEMS WITH INNOVATIVE ITEM TYPES, ALTHOUGH I THINK TONY HAD A VERY GOOD EXPLANATION FOR HOW WE COULD MITIGATE A PROBLEM THERE. AND, ALSO, SOMETHING ELSE THAT JOHN BROUGHT UP THAT WE DIDN'T REALLY GET BACK TO, BUT JUST THE SIMPLE THREAT WITH NOT JUST COMPUTER-BASED TESTING BUT WITH THESE PARTICULAR COMMON CORE STATE STANDARDS ASSESSMENTS, THAT THE MULTISTATE NATURE OF THEM IS LIKELY TO LEAD TO MUCH GREATER THEFT INCENTIVE. AND I THINK THAT'S A VERY REAL RISK BOTH ON THE PART OF STUDENTS SHARING ITEMS THROUGH SOCIAL NETWORKING AND ALSO ON THE PART OF ADULTS WHERE THERE COULD BE A REAL FINANCIAL GAIN. AND THAT'S NOT JUST FOLKS IN SCHOOLS. WE'RE TALKING ABOUT EVERYWHERE-- EVERYWHERE WITH ACCESS--WHICH MEANS CHAIN OF CUSTODY ISSUES WILL ALSO BE VERY IMPORTANT AND HARDER TO TRACK. I ALSO RECALL THAT WE TOUCHED ON SOME OTHER VERY INTERESTING POINTS, WHICH INCLUDE THE FACT THAT THESE NEW METHODS

OF TEST ADMINISTRATION ALSO PROVIDE POTENTIALLY NEW STATISTICS THAT WILL HELP US MONITOR AND POTENTIALLY EVEN IN REAL TIME, BUT CERTAINLY LATER, TEST TAKING AND POTENTIALLY ALLOW US NOT ONLY TO FIND--MORE USEFUL INFORMATION BUT ALSO FOR CHECKING THE INTEGRITY OF THE TESTS. AND, WES, THANK YOU FOR THAT EXAMPLE, LOOKING AT SORT OF HOW WE WOULD MOVE BEYOND WRONG-TO-RIGHT ERASURE ANALYSIS, WHICH WE SPENT A LOT OF TIME ON EARLIER, TO REALLY GETTING A LOT MORE INFORMATION OUT OF THAT.

I THINK AT THIS POINT WHAT I'D LIKE TO DO IS JUST SUMMARIZE--WELL, ACTUALLY, JOANNE, IF I COULD ASK YOU TO PROVIDE SOME CLOSING REMARKS, I THINK AT THIS POINT THAT WOULD BE A GOOD TIME TO DO IT.

SO, FIRST OF ALL, I'M VERY PLEASED, THEN, TO WELCOME OUR LAST SPEAKER OF THE DAY, JOANNE WEISS, THE CHIEF OF STAFF TO THE UNITED STATES SECRETARY OF EDUCATION, ARNE DUNCAN.

JOANNE JOINED THE DEPARTMENT IN 2009 TO SERVE AS SENIOR ADVISER TO THE SECRETARY AND THE DIRECTOR OF THE RACE TO THE TOP FUND. AND IN THIS CAPACITY, SHE LED THE DEPARTMENT'S RACE TO THE TOP PROGRAM, WHICH WAS DESIGNED TO ENCOURAGE, AS YOU KNOW, AND REWARD STATES MAKING SYSTEM-WIDE, COMPREHENSIVE, AND COHERENT EDUCATION REFORM.

PRIOR TO JOINING THE ADMINISTRATION, JOANNE WAS THE PARTNER AND CHIEF OPERATING OFFICER AT THE NEW SCHOOLS VENTURE FUND, A VENTURE PHILANTHROPY FIRM WORKING TO TRANSFORM PUBLIC EDUCATION BY SUPPORTING EDUCATION ENTREPRENEURS AND CONNECTING THEIR WORK TO SYSTEMIC CHANGE.

WE'RE VERY PLEASED THAT JOANNE IS ABLE TO JOIN US TODAY. AND I'D LIKE TO THANK HER FOR PARTICIPATING. AND I TURN THINGS OVER TO JOANNE.

**JOANNE WEISS:** THANKS, JACK. UM, CAN YOU HEAR ME? AM I TURNED ON UP HERE? YES, NO? IT'S NOT, RIGHT? OK, NOW? NO.

A LITTLE BATTERY SWITCHING GOING ON NOW.

AH, GREAT. THANK YOU.

SO, THANK YOU, JACK, AND THANK YOU, JOHN, FOR HOSTING THIS EVENT TODAY ON BEHALF OF THE SECRETARY. THANK YOU FOR MODERATING THESE PANELS, JACK. YOUR SUMMARIES WERE TERRIFIC RIGHT UP UNTIL THE LAST ONE. I'M IMPRESSED.

I'M GOING TO KEEP MY REMARKS VERY BRIEF BECAUSE I REALIZE THAT I'M ALL THAT'S STANDING BETWEEN MANY OF YOU AND THE AIRPORT. BUT I DO

WANT TO REALLY THANK ALL OF YOU FOR PARTICIPATING IN TODAY'S TESTING INTEGRITY SYMPOSIUM. WE'RE GOING TO SHARE WHAT WE LEARNED TODAY WITH THE SECRETARY, OF COURSE, BUT MUCH MORE IMPORTANTLY, WE'RE GOING TO SHARE IT WITH THE PUBLIC.

WE'VE ALREADY PUBLISHED ONLINE ALL OF THE RESPONSES THAT WE'VE RECEIVED TO THE REQUEST FOR INFORMATION THAT WE'VE PUT OUT.

ALL THE WEBINAR TODAY, INCLUDING EVERYBODY'S POWERPOINTS AND REMARKS, ARE ONLINE AND WILL REMAIN UP THERE SO THAT PEOPLE CAN LOOK AT THIS AND LEARN FROM IT. AND WE'RE GOING TO BE DRAFTING A SUMMARY OF THE FINDINGS SO THAT WE CAN HELP STATES AND SCHOOL DISTRICTS IDENTIFY AND IMPLEMENT BEST PRACTICES IN THIS AREA.

THE GENESIS OF THIS EFFORT, AS I THINK YOU HEARD THIS MORNING, CAME FROM A CONVERSATION THAT WE HAD WITH D.C. PUBLIC SCHOOLS CHANCELLOR KAYA HENDERSON. LAST SUMMER WHEN SHE TOLD ME THAT AS D.C. WAS PURSUING ITS INVESTIGATIONS INTO ALLEGATIONS OF CHEATING, THERE WAS VIRTUALLY NO LIBRARY OF BEST PRACTICES TO RELY ON AND NO STANDARDS OF TESTING INTEGRITY FOR THEM TO RELY ON. SO THEY WERE REALLY FLYING BLIND. AND SHE ASKED WHETHER THE DEPARTMENT WOULD STEP IN TO HELP STATES AND DISTRICTS BY PULLING TOGETHER A GROUP OF EXPERTS, POOLING THEIR KNOWLEDGE AND PUBLISHING THEIR

RECOMMENDATIONS. AND TODAY IS THE CULMINATING EVENT OF THAT, SHORT, OF COURSE, OF US PUBLISHING ALL OF THE INFORMATION THAT WE'VE LEARNED.

BUT THE GOAL, REALLY, OF THIS WHOLE ENDEAVOR HAS BEEN TO HELP STATES AND DISTRICTS AND, OF COURSE, THE DEPARTMENT, AS WE ALSO THROUGH NAEP AND OTHER VEHICLES ARE LARGE-SCALE ASSESSMENT PROVIDERS, BUT REALLY TO HELP STATES AND DISTRICTS HAVE THE BENEFIT OF THE COLLECTIVE WISDOM OF THE EXPERTS IN THIS FIELD, WHO HAVE BEEN SO HARD TO TAP INTO FOR THE 14,000 SCHOOL DISTRICTS ACROSS AMERICA THAT ARE ADMINISTERING THESE KINDS OF ASSESSMENTS EVERY YEAR.

THE CHANCELLOR TOLD ME WHEN SHE CALLED ME LAST SUMMER THAT THERE WERE THREE THINGS THAT SHE NEEDED TO BETTER UNDERSTAND.

FIRST, SHE ASKED US TO HELP UNDERSTAND HOW TO PREVENT CHEATING FROM HAPPENING IN THE FIRST PLACE.

SECOND, SHE WANTED MORE INFORMATION ON WHAT BEST PRACTICES WERE FOR IDENTIFYING OR FLAGGING THE PLACES WHERE CHEATING MIGHT HAVE OCCURRED, TO REALLY MAKE THE AUDITING PROCESS MORE EFFECTIVE AND MORE COST-EFFECTIVE.

AND THE THIRD THING WAS WHAT BEST PRACTICES LOOKED LIKE FOR INVESTIGATING ALLEGATIONS OF CHEATING, WHAT THE PROPER PROCEDURES WOULD BE FOR--CONDUCTING SUCH INVESTIGATIONS.

SO--I WANT TO ALSO JUST REALLY THANK KAYA FOR RAISING THIS ISSUE TO US, AND SPEAKING PERSONALLY, I HOPE, ON BEHALF OF MANY OF YOU, AS WELL, HERE, THOUGH, TODAY WAS REALLY A VALUABLE LEARNING EXPERIENCE.

I WANT TO JUST THANK ALL OF YOU FOR SHARING SO MUCH WISDOM WITH US TODAY. THERE WAS A LOT TO DIGEST. AND I THINK THAT WE'LL BE KIND OF GOING BACK OVER THESE TRANSCRIPTS AND THE SUMMARY THAT WE'LL PUBLISH FOR A LONG TIME JUST TO RETURN TO IT AND REMIND OURSELVES OF THE KIND OF PRACTICES AND PROCEDURES THAT ARE GOING TO HELP US THROUGH THIS.

WE REALLY, AS A COUNTRY, RELY ON ACCURATE, RELIABLE, AND TIMELY INFORMATION TO UNDERSTAND WHAT STUDENTS KNOW AND CAN DO, AND WE HAVE TO HAVE INFORMATION THAT'S ROBUST AND RELIABLE SO THAT PARENTS, SO THAT TEACHERS, SO THAT COMMUNITIES, SO THAT THE PRESS HAVE INFORMATION THAT THEY CAN TRUST.

HIGH-QUALITY DATA IS CRITICAL TO MAKING SURE THAT WE CAN EVALUATE THE EFFECTIVENESS OF OUR POLICIES AT THE FEDERAL, STATE, AND LOCAL

LEVELS. WE NEED TO USE IT TO EVALUATE THE EFFECTIVENESS OF TEACHING IN OUR CLASSROOMS, OF THE CURRICULUM MATERIALS THAT ARE PROVIDED TO OUR KIDS, AND OF THE SCHOOLS AND THE DISTRICTS THAT ARE PROVIDING SERVICES TO OUR STUDENTS.

IN SHORT, IT REALLY OFFERS AN INVALUABLE LENS ON HOW WE CAN DO A BETTER JOB OF EDUCATING EACH CHILD.

SO WE NEED TO DO EVERYTHING THAT WE CAN TO DEFEND AGAINST SECURITY BREACHES AND THREATS TO DATA QUALITY. WE REALLY HAVE TO BE ABLE TO TRUST THE INFORMATION THAT WE RECEIVE.

YOU KNOW, THERE'S BEEN A LOT OF TALK TODAY ABOUT SORT OF TESTING VERSUS TEACHING. AND TESTING, AS WE ALL KNOW, IS REALLY AN INTEGRAL PART OF EDUCATION. TESTING AND TEACHING ARE NOT AT ODDS. THEY'VE GONE HAND-IN-HAND SINCE FORMAL EDUCATION BEGAN.

THE ANSWER TO THE EXISTENCE OF TEACHING, AS MANY OF YOU SAID TODAY, IS NOT TO REJECT TESTING, IT'S TO DEAL WITH THE CHEATING.

SECRETARY DUNCAN SAID NOT TOO LONG AGO, "THE EXISTENCE OF CHEATING SAYS NOTHING ABOUT THE MERITS OF TESTING. INSTEAD, CHEATING REFLECTS A WILLINGNESS TO LIE AT CHILDREN'S EXPENSE TO AVOID ACCOUNTABILITY, AN APPROACH THAT I REJECT ENTIRELY."

ARNE IS NOT ALONE.

IT'S ALSO AN APPROACH THAT'S REJECTED BY THE VAST MAJORITY OF EDUCATORS, AS MANY OF YOU TODAY POINTED OUT--PEOPLE WHO WOULD NEVER PARTICIPATE IN OR EXCUSE TEACHING. CURRENT DATA SUGGESTS THAT TEACHING IS AN ISSUE IN ABOUT 4% TO 5% OF ELEMENTARY SCHOOL CLASSROOMS. THE VAST MAJORITY OF EDUCATORS ARE BEHAVING IN ETHICAL WAYS. THEY'RE ACTING WITH INTEGRITY. AND, MOST IMPORTANTLY, THEY'RE SERVING AS STRONG ROLE MODELS FOR OUR STUDENTS. BUT PUTTING OUR HEADS IN THE SAND AND PRETENDING THAT CHEATING DOESN'T EXIST CLEARLY DOESN'T SERVE OUR CHILDREN WELL.

CHEATING DOES OCCUR IN OUR SCHOOLS. AND WE NEED TO BE MORE VIGILANT AND MORE PREPARED. AND TODAY'S DISCUSSION HAS HELPED US START TO BUILD A ROAD MAP FOR WHAT TO DO AND HOW TO ADDRESS THIS. SO WHILE WE MAY NOT BE ABLE TO ELIMINATE CHEATING ENTIRELY, BY EMPLOYING A BUNCH OF THE STRATEGIES THAT WE HEARD TODAY, WE CAN HELP TO MAKE SURE THAT IT DOESN'T OCCUR ON A SYSTEMIC LEVEL.

SO I WANT TO JUST THANK YOU AGAIN FOR SHARING YOUR EXPERTISE WITH US TODAY AND FOR YOUR LEADERSHIP IN THIS AREA. WE ALL LEARNED A LOT FROM YOU. AND WE AT THE DEPARTMENT PLEDGE OVER THE COMING WEEKS

TO PACKAGE THE EXPERTISE YOU'VE SHARED WITH US IN WAYS THAT ARE USEFUL TO PRACTITIONERS AND ACCESSIBLE TO PRACTITIONERS SO THAT STATES AND DISTRICTS HAVE THE BENEFIT OF THIS INFORMATION GOING FORWARD.

I'M NOT SURE IF JACK'S GOING TO WRAP UP, BUT LET ME WISH YOU SAFE TRAVELS HOME. AND THANK YOU, AGAIN, FOR COMING, AND SHARING YOUR EXPERTISE WITH US TODAY.

WE GREATLY APPRECIATE IT.

**JACK BUCKLEY:** THANKS VERY MUCH, JOANNE.

[APPLAUSE]

AT THIS POINT, I WILL NOT INFLICT ANOTHER SUMMARY ON YOU. SO JUST, AGAIN, LET ME ECHO JOANNE'S REMARKS THAT ON BEHALF OF THE DEPARTMENT, THE INSTITUTE OF EDUCATION SCIENCES, AND THE NATIONAL CENTER FOR EDUCATION STATISTICS, THANK YOU VERY MUCH, PANELISTS, FOR YOUR PARTICIPATION AND YOUR OUTSTANDING REMARKS AND ASSISTANCE, AND AUDIENCE FOR YOUR THOUGHTFUL ATTENTION AND WONDERFUL QUESTIONS.

AND HAVE A GREAT DAY.

[APPLAUSE]