

UNREPORTED

IN THE COURT OF SPECIAL APPEALS

OF MARYLAND

No. 1557

SEPTEMBER TERM, 2014

SHAYLIN TODD

v.

BALTIMORE CITY BOARD OF SCHOOL
COMMISSIONERS ET AL.

Eyler, Deborah, S.,
Graeff,
Hotten,

JJ.

Opinion by Eyler, Deborah, S., J.

Filed: October 19, 2015

*This is an unreported opinion, and it may not be cited in any paper, brief, motion, or other document filed in this Court or any other Maryland court as either precedent within the rule of *stare decisis* or as persuasive authority. Md. Rule 1-104.

In this administrative agency appeal, Shaylin Todd, the appellant, a former principal at Fort Worthington Elementary School, challenges the decision of Lillian M. Lowery, Ed.D., the Maryland State Superintendent of Schools (“the Superintendent”), to revoke her Maryland State Department of Education (“MSDE”) certifications based on a finding that Todd violated test security and data reporting policies with respect to the administration of standardized testing at her school. The Baltimore City Board of School Commissioners (“the Local Board”) is the appellee.¹ The Superintendent’s decision was affirmed on judicial review by the Circuit Court for Baltimore City.

Todd presents three questions for review,² which we have condensed and rephrased as one: Was the Superintendent’s decision to revoke Todd’s certifications legally correct and supported by substantial evidence? We conclude that it was, and therefore shall affirm the judgment of circuit court.

¹ The MSDE and the Superintendent were not parties before the Office of Administrative Hearings or the circuit court and filed a notice of intent not to participate in the instant appeal. Thus, the Local Board is the only appellee.

² The questions as posed by Todd are:

1. Was the [ALJ’s] Proposed Decision on this case, which was adopted by the Superintendent, fatally tainted with bias?
2. Does the admission of, and reliance upon, anonymous and unsworn statements not subject to cross-examination warrant reversal of the Superintendent’s decision?
3. Was there substantial evidence to support the Superintendent’s decision that Shaylin Todd willfully and knowingly violated COMAR 13A.12.05.02C(2)(c)?

FACTS AND PROCEEDINGS

A. Test Security and Data Reporting Regulations

Before recounting the facts, we shall review the relevant regulations. Test security and data reporting policies and procedures are set forth in COMAR 13A.03.04.05 & .06.

Pursuant to COMAR 13A.03.04.05A,

[i]t is a violation of test security for an individual to fail to follow test administration procedures promulgated by the local board of education or the State Board of Education and published in test administration manuals and related materials for mandatory tests administered by or through the State Board of Education to students or educators.

And, pursuant to COMAR 13A.03.04.05B,

[i]t is a violation of test security for an individual to:

(4) Provide answer keys or answers orally, in writing, or by any other means, to examinees;

(5) Coach examinees during testing by giving them answers to test questions or otherwise directing or guiding their responses or altering or interfering with examinees' responses in any way;

(6) Fail to follow security regulations and procedures for distribution and return of secure test materials . . .;

(7) Fail to properly monitor test administration, including permitting inappropriate collaboration between or among individuals;

(9) Participate in, direct, aid, counsel, assist, encourage, or fail to report any of the acts prohibited in this chapter[.]

As relevant here, pursuant to COMAR 13A.03.04.06,

[i]t is a violation of data collection and reporting for an individual, school, or school system to:

(1) Fail to report test scores, numbers of students tested, and other indicators of test performance on mandatory tests administered by or through the State Board of Education, as well as all other data elements reported to the Department; [or]

(2) Report incorrect or otherwise inaccurate test scores, numbers of students tested, other indicators of test performance, and participation on mandatory tests administered by or through the State Board of Education, as well as all other data elements reported to the Department[.]

The Superintendent “shall” revoke or suspend a teaching certificate of an individual who “[w]illfully and knowingly . . . [c]ommits a violation of the test security and data reporting policy and procedures set forth in COMAR 13.A.03.04.” COMAR 13A.12.05.02C(2).

B. Todd’s Background

In 1999, Todd received her Bachelor’s Degree from Virginia State University. She was accepted into Teach for America (“TFA”), a national non-profit organization that trains and places recent college graduates in underserved school districts. In the fall of 2000, TFA placed Todd at Belmont Elementary School, in the Baltimore City Public School (“BCPS”) system.

At Belmont, Todd taught first grade in 2000/2001 and second grade in 2001/2002. During that time, she obtained her Master’s Degree in Education from Johns Hopkins University. For the next three years, Todd taught second grade at Waverly Elementary School.

In the fall of 2005, Todd was accepted into the New Leaders for New Schools program (“New Leaders”). New Leaders is a non-profit organization that recruits and trains young teachers to become school administrators. Under the program, a recruit spends a year as a “resident principal” shadowing a “mentor principal” at a school. New Leaders provides a coach to support its recruits as well. New Leaders placed Todd as a “resident principal” at Baybrook Elementary School, under the mentorship of principal Lydia Lemon.

In the fall of 2006, Todd became principal at Fort Worthington Elementary (“Fort Worthington”), in southeast Baltimore. The school serves an impoverished population, with ninety percent of its students qualifying for free or reduced meals. It also was a historically failing school. Fort Worthington had failed to make adequate yearly progress (“AYP”), a benchmark created under the No Child Left Behind Act (“NCLB Act”), for three years.

The NCLB Act set a deadline requiring that by 2014, all students, including special education students, would be proficient or advanced in reading and math as measured by state tests meeting certain criteria. In Maryland, the Maryland School Assessment (“MSA”) was the standardized test that measures students’ proficiency in reading and math under the NCLB Act.³ It was administered in March of each year to students in grades 3 through 8. The MSA included multiple-choice questions, which

³ The MSA has since been replaced by a new testing regimen.

students answer by selecting a response and filling in a bubble in pencil next to it on an answer form. It also included brief constructed response questions (“BCRs”) for which students write out substantive responses.

Between 2004 and 2006, over 44 percent of the students at Fort Worthington did not score at the proficient or advanced level on the MSA. For special education students at Fort Worthington, less than a third scored proficient or advanced during that time period.

After Todd took over as the principal in 2006, she brought in new staff and implemented numerous new programs. She offered teachers professional development opportunities, including Saturday sessions. She assigned an instructional support teacher (“IST”) to each grade level. She brought in reading and math programs that had shown success in other failing schools.

To track student progress, Todd held regularly scheduled assessment conferences with teachers to discuss their students’ performance on three “benchmark assessments” performed during the school year. The benchmark assessments were lettered A, B, and C, with assessment C occurring in the winter, shortly before the MSA. Todd kept a data card detailing each individual student’s performance on these benchmark assessments on an “assessment wall.”

In March of 2007, the MSA was administered at Fort Worthington. Fifty-eight percent of total students and nearly twenty-three percent of special education students scored proficient or advanced in reading. Sixty-three percent of total students and thirty-

nine percent of special education students scored proficient or advanced in math. These scores represented a ten percent increase in reading proficiency and a fifteen percent increase in math proficiency for Fort Worthington overall, and a nineteen percent increase in math proficiency for the special education students. The special education students' proficiency in reading remained the same as in the prior year.

In the 2007/2008 and 2008/2009 school years, Fort Worthington's MSA scores skyrocketed. In 2008, 89.8 percent of students tested proficient in reading, 93.3 percent of special education students tested proficient in reading, 89.8 percent of students tested proficient in math, and 83.7 percent of special education students tested proficient in math. In 2009, 83.7 percent of students tested proficient in reading, 86.8 percent of special education students tested proficient in reading, 86.7 percent of students tested proficient in math, and 78.9 percent of special education students tested proficient in math. Fort Worthington and Todd were celebrated for these achievements.

C. The BCPS and MSDE Investigation

In the 2009/2010 school year, the MSA was administered between March 9 and March 23, 2010, to 225 students. Two days after testing closed, a BCPS employee called BCPS's fraud hotline anonymously, complaining that Todd had violated MSA testing procedures. A second complaint, also from a BCPS employee, came in the next day, alleging that Todd had falsified MSA test scores. On April 4, 2010, a teacher at Fort Worthington made a complaint to the fraud hotline, alleging that Todd had instructed teachers to prompt students to choose the correct answers on the MSA. On April 8, 2010,

a parent of three Fort Worthington students called the BCPS fraud hotline and reported that Todd had told teachers and other school personnel to cheat on the MSA.

Benjamin Feldman, Ph.D, the Chief of the Achievement and Accountability Office (“AAO”) for BCPS, reported the alleged testing irregularities to Leslie Wilson, Ph.D, the Assistant Superintendent. Dr. Feldman asked that Tamara Lewis, Ph.D, the State Test Security Officer, open a test security case at Fort Worthington.

In response, Dr. Lewis prepared a report for the State Test Security Committee and directed the BCPS to conduct a full investigation. The BCPS assigned the investigation to James Gresham, Ed.D., the Coordinator of Assessments and Security Investigations for Testing at the AAO.

On May 10, 2010, a former Fort Worthington teacher made an anonymous phone call to the BCPS fraud hotline alleging that Todd, her ISTs, and other staff members had changed students’ answers on the 2008 MSA. The BCPS also received a letter dated May 10, 2010 from a former teacher at Fort Worthington. The teacher had taught third grade in 2006/2007 and 2007/2008, the first two years Todd was principal. The teacher alleged that during administration of the MSA in 2008, the IST for the third grade removed students from her classroom to be tested separately; students were given extra time to complete the MSA; she noticed an unusual number of erasures in students’ test booklets; an IST told her that one of her special education students, who was unable to read or write, had performed well on the MSA; and ISTs were locked in a room with the blinds drawn for some period of time after the test was administered. The author of the

letter also complained that Todd had failed to follow an Individual Education Plan (“IEP”) for a student in her class and that Todd had been dismissive when the teacher brought that issue to her attention.

On August 6, 2010, while Dr. Gresham’s investigation was in progress, Todd resigned from her position as principal at Fort Worthington.

On March 15, 2011, Dr. Gresham issued a report on the testing investigation at Fort Worthington. In it, Dr. Gresham explained that he and other AAO staff had interviewed Todd; Assistant Principal Callandra Arrington; Helen Mayo and Tessera Slade, the MSA test coordinators at Fort Worthington; fourteen current teachers and one former teacher; a school social worker; and fourteen students. In addition, Dr. Gresham reviewed student attendance records, benchmark assessments, Stanford 10⁴ and MSA results, and IEP and 504 plans for special education students. The report “incorporate[d] . . . anecdotal and quantitative data” to determine whether the allegations of testing irregularities were substantiated.

The anecdotal data was derived from the interviews. Dr. Gresham noted that there was a “clear divide amongst the staff,” but found that the consensus was that Todd was a “data driven administrator who focused . . . on [AYP]” and that the staff felt “tremendous pressure . . . to make AYP.” Teachers uniformly reported that they were required to provide a “daily list of students who did not finish sections of the [MSA]” to Mayo and

⁴ The Stanford 10 is a standardized test administered to first and second grade students.

Slade, and that Mayo and Slade later removed students from their classrooms for the purpose of completing the test. Mayo and Slade claimed that students only were permitted to complete the test at a later time if they had been late or absent during the MSA administration, but many teachers disputed that this always was the case.⁵

The “quantitative data” aspect of Dr. Gresham’s report analyzed overall performance at Fort Worthington and individual performance by a sample of students between 2004 and 2009, so as to compare student achievement before and during Todd’s tenure as principal. Specifically, Dr. Gresham used three approaches to analyze the data, examining 1) “the overall trajectory of Stanford and MSA scores at Fort Worthington and City Schools by grade,” 2) the “average change in scores for students attending Fort Worthington versus other schools,” and, 3) test scores for students who transferred into and out of Fort Worthington during the relevant period to determine if their scores markedly increased or decreased before or after the transfer. Dr. Gresham explained that quantitative data analysis cannot “definitively identify the presence of testing irregularities,” but if all three “methods converge on similar conclusions,” that is “strong evidence of testing issues at a school.”

The AAO investigation found that the overall trajectory of test scores at Fort Worthington showed large gains in the years after Todd became principal, especially

⁵ Mayo was recommended for termination from BCPS as a result of these allegations. She retired before being terminated, however. Thereafter, upon a request by the BCPS, the MSDE revoked her teaching certifications. She did not appeal.

between 2007 and 2008. In particular, in 2008, the special education students, who historically had a “proficiency pass rate[]” that was “less than half” the rate for the total student body, actually exceeded the pass rate in reading for the total student body.

The average change in test scores per student at Fort Worthington, as compared to students in the BCPS as a whole, was notable after Todd became principal. While students at Fort Worthington showed average gains “significantly worse” than the district average in 2006, in 2007 their gains were slightly higher than the district average. By 2008, the gains were more than 20 percentage points higher than the district average. According to Dr. Gresham, “[t]hese differences [were] so huge that it [was] virtually impossible they could occur by chance.” He noted, however that the huge average gain also could stem from “improved instruction.”

Finally, the AAO comparison of the changes in test scores of students who transferred into and out of Fort Worthington during the relevant time period showed that students who attended Fort Worthington in the 2005/2006 school year and transferred to other schools for the 2006/2007 school year improved their test scores relative to other students in the district. In contrast, students who attended Fort Worthington in 2007/2008 and transferred to other schools for the 2008/2009 school year experienced a “large and statistically significant drop in their [test scores] relative to the district,” which suggested either that their 2008 MSA test scores while attending Fort Worthington were “inflated” or that the students’ “earlier progress diminished” as a result of poor instruction at their new schools.

Dr. Gresham found that the “consistency and corroboration between the teacher and student interviews, the triangulation of the quantitative data, and statistical analyses used in this investigation” substantiated the allegations of cheating on the MSA at Fort Worthington and “provided conclusive evidence that Fort Worthington ha[d] violated the [MSDE] test administration guidelines.” Dr. Gresham concluded that Mayo and Slade, who as mentioned were the test coordinators at Fort Worthington, were directly involved in testing violations, and Arrington also “may have committed testing violations.” He further concluded that Todd “established the protocol of having teachers identify students who ‘may be disruptive during testing,’” and directed that those students be tested separately from the rest of the student body. The report did not otherwise reach any conclusions regarding Todd’s involvement in cheating on the MSA.

After Dr. Gresham completed his investigation, the MSDE referred the investigation to Jessie Pollack, Ph.D, a psychometrician, for an erasure analysis. Dr. Pollack received the 2010 MSA test booklets from Fort Worthington and requested the 2009 MSA test booklets from the BCPS contractor who maintained them. (The 2008 MSA test booklets already had been destroyed.) Upon reviewing Fort Worthington’s test scores for those years, Dr. Pollack noticed that the performance of the special education students was “stellar, to the extent that [she] had never seen performance that high before.” She separated out the 2010 test booklets for special education students at Fort Worthington and noticed “extreme numbers of erasures and extremes in wrong to right answer changes.” She brought that to the attention of the State Test Administration and

Security Committee and it gave her the go ahead to focus on the testing of special education students.

On June 23, 2011, Dr. Pollack submitted to the MSDE her report entitled “Test Security Investigation: Fort Worthington Elementary School.” In it, she explained that she had conducted an erasure⁶ analysis on the multiple-choice question sections of all the 2009 and 2010 MSA test booklets for math and reading for the special education students at Fort Worthington, a total of 142 booklets.⁷ For each test booklet, she counted the total number of erasures; and, for each erasure, she determined whether it was from wrong to right, from right to wrong, or from wrong to wrong. She identified a benchmark of four or more wrong to right answer changes in a test booklet as a red flag.

Dr. Pollack found that in 2009, 48 of 77 test booklets (62%) had four or more wrong to right answer changes and 26 test booklets (33%) had ten or more wrong to right answer changes. In 2010, 39 of 65 test booklets (60%) had four or more wrong to right answer changes and 11 test booklets (16.9%) had ten or more wrong to right answer changes. Dr. Pollack noted that in some test booklets, the erasures reflected that students had chosen several wrong answers before selecting the right answer; in others, there were noticeable differences in the “style and completeness of the erasures”; and finally, in

⁶ Dr. Pollack defined an “erasure” as “smeared graphite either within or around the bubble . . . where the student marks his or her answer.”

⁷ This amounted to 16 percent of the MSA test booklets for Fort Worthington.

some math test booklets, student solutions showing an incorrect answer were next to a wrong to right answer correction.

Dr. Pollack concluded that “[t]he body of evidence,” including her observations of the markings, the test scores overall, and the “magnitude of the [wrong to right] answer changes” “permit[ted] the interpretation that test violations occurred in both the MSA Reading and MSA Mathematics test administrations in Grades 3, 4, and 5 in both [2009 and 2010.] Therefore, staff at Fort Worthington . . . violated test security procedures for State assessments mandated by [COMAR 13A.03.04.05].” Dr. Pollack expressed no opinion as to how and by whom the test violations were accomplished.

D. The Charges Against Todd

Meanwhile, on June 8, 2011, Andrés Alonso, Ed.D., the then CEO of the BCPS, wrote to the Superintendent to request that Todd’s MSDE certifications be revoked based upon her violation of the test security and data reporting policies and procedures in COMAR 13A.03.04.05 & .06. As relevant here, he charged that, in violation of COMAR 13A.03.04.05A, Todd directed a small cohort of allegedly disruptive students to be tested outside of their classroom setting even though they did not have IEPs or 504 plans in place that would permit such a testing accommodation. He also charged that Todd established protocols under which staff coached, directed, or altered student test responses, in violation of COMAR 13A.03.04.05B(5); failed to properly monitor test administration, in violation of COMAR 13A.03.04.05B(7); and directed, encouraged, or failed to report testing violations, in violation of COMAR 13A.03.04.05B(9).

On June 14, 2011, the MSDE notified Todd of the charges and advised her that she had a right to request a hearing.

On June 24, 2011, Todd requested a hearing. Her case was referred to the Office of Administrative Hearings (“OAH”).

E. The OAH Hearing

Over ten days in February, April, and June of 2012, a contested case hearing proceeded before an Administrative Law Judge (“ALJ”) at the OAH. In its case, the Local Board called eight witnesses: Dr. Lewis; Dr. Gresham; Dr. Pollack; Lawrence Rudner, Ph.D, an expert in psychometrics and statistics; Crystal Chestnut, the former school secretary at Fort Worthington; Muriel Olowokakoko, a third grade teacher at Fort Worthington; Mayo; and Jerome Jones, the manager of labor relations for BCPS.

The Local Board argued that revocation was warranted for three reasons: 1) Todd inappropriately established a protocol for testing students with alleged behavioral issues in small-group settings, outside their classrooms, even though these students did not have an IEP or 504 Plan permitting a testing accommodation; 2) Todd resigned after she learned of the cheating allegations; and 3) Todd encouraged teachers and other staff to coach students during administration of the MSA and/or had knowledge of or participated in the completion of or changing of answers after the MSA test administration was over. As we shall explain, the final agency decision to revoke Todd’s MSDE certifications was based only on an adverse finding on the third issue. Accordingly, we will not discuss the evidence adduced on the other two issues.

The Local Board presented a largely circumstantial case in the form of quantitative and statistical test score and erasure analyses from Drs. Gresham, Pollack, and Rudner that showed that widespread cheating had occurred during administration of the 2008, 2009, and 2010 MSA tests. The reports by Dr. Gresham and Dr. Pollack were introduced into evidence.

In addition to confirming the information in his report, Dr. Gresham testified that he had reviewed benchmark assessment data for Fort Worthington students in the 2008/2009 school year and compared it to the students' 2009 MSA scores. He explained that benchmark assessment C, which, as mentioned, is given each winter shortly before the MSA, is like a "mock MSA." The students' average proficiency on benchmark assessment C was significantly lower than their MSA proficiency, however. To give one example, in 2009, just 48 percent of the third grade class scored proficient or advanced on reading on the multiple choice section of benchmark assessment C, but 81.2 percent of that class scored proficient or advanced on the MSA.

Dr. Pollack testified that her review of the literature concerning erasure analyses showed that the mean number of total erasures by students on multiple-choice tests is "typically less than 3." She relied on that data in determining that test booklets with four or more "wrong to right erasures" should be flagged. She opined, consistent with her report, that the high rate of wrong to right erasures was evidence of tampering.

Dr. Rudner testified that he reviewed Dr. Pollack's report, reanalyzed her data, and "came to the same conclusion she did." He began by calculating an "erasure rate." He

multiplied the number of special education students who took the MSA at Fort Worthington in a given year by the number of questions, to reach a total number of “opportunities to erase”; he then divided the total number of actual erasures by that number. He explained that the normal erasure rate is between 3-5 percent, sometimes as high as 8 percent. The erasure rates on the 2009 and 2010 MSAs varied by grade and subject with the lowest rate being 6.7 and the highest being 20.4. Dr. Rudner explained that the literature made clear that test takers are much more likely to change a right answer to a wrong answer and that this was even more likely for a low-performing population. He identified 87 test booklets that were suspect because they had four or more wrong to right erasures. Within that “suspect” grouping, Dr. Rudner found that eighty percent of the erasures were wrong to right. Based on this evidence and his review of the data as a whole, Dr. Rudner opined that there “was no doubt whatsoever there was extensive tampering. Lots of books were tampered with, and when those books were tampered with, lots of questions were tampered.” He characterized the “data” as “so far from usual” that it was “beyond statistics.”

The Local Board also presented direct and circumstantial evidence linking Todd to the cheating at Fort Worthington. Olowokakoko, a teacher at the school, testified that Todd had directed teachers to instruct students to “try again” if the teachers noticed wrong answers during administration of the MSA. Dr. Gresham’s report, which as mentioned was in evidence, included numerous teacher statements about the immense pressure to make AYP after Todd took over as principal; that Todd made student

achievement a competition at the school; that Todd told teachers to “do whatever it takes to make sure the scholars are successful on the test”; and that Todd was highly “data driven.” The Local Board also introduced into evidence the May 10, 2010 letter and summaries of the anonymous complaints made to the fraud hotline.

In her case, Todd testified and called eleven witnesses, including Bernard Lentz, Ph.D, an expert in complex multivariate statistical analysis; Omari Todd, her husband, who at the time was Senior Vice President of Regional Operations for TFA; Shannon Coleman-Siciliano, a former IST and reading intervention teacher at Fort Worthington; Arrington; Maisha Dorsey, a former teacher at Fort Worthington; Darlene Merry, the former chief academic officer for New Leaders; and Peter Kannam, former executive director of TFA and New Leaders.

Todd’s defense was two-fold. She denied that any cheating had taken place and, in the alternative, argued that if there was cheating, it happened without her knowledge or involvement. She maintained that the test score gains could be explained by instructional improvement and supports put in place under her leadership. She asserted that a small group of teachers at Fort Worthington, including Olowokakoko, were angry with her because of a mold problem at the school that the teachers believed had caused another teacher to suffer a miscarriage.

Arrington, Coleman-Siciliano, and Dorsey testified that Todd never directed them to help students cheat; that they never heard her tell anyone else to cheat; and that no staff members ever reported allegations of cheating to them.

Todd’s expert witness, Dr. Lentz, criticized the methodology employed by Dr. Pollack and attempted to undermine the conclusions reached by her and by Dr. Rudner.

In rebuttal, the Local Board called Dr. Alonso, recalled Dr. Rudner, and called Irma Johnson, who was involved in discussions with Todd with regard to her resignation.

F. The ALJ’s Proposed Decision and Proposed Order

On September 24, 2012, the ALJ issued his proposed decision and order. As a threshold matter, the ALJ ruled that the Local Board’s argument that Todd’s resignation, after being charged with test security and data reporting violations, was a basis to support the revocation of her MSDE certifications was not properly before it.

The ALJ then turned to the allegations that Todd improperly directed that general education students be tested outside of their classrooms. The ALJ found that the testing arrangements for these students amounted to an improper setting accommodation in violation of the Maryland Accommodations Manual. He further found that it was “virtually impossible to believe” Todd’s testimony that she did not realize she was violating the Maryland Accommodations Manual by allowing these students to be tested in a small group setting. Nevertheless, the ALJ gave Todd the “benefit of the doubt” and found that the Local Board had failed to prove a willful and knowing violation.

The ALJ concluded, however, that the Local Board “overwhelmingly proved that systematic cheating occurred at Fort Worthington on the 2008, 2009, and 2010 MSAs, and that [Todd] was aware of and complicit in the cheating.” The ALJ found the most compelling evidence to be the “improbable improvement in Fort Worthington’s MSA

scores between 2007 and 2008” and, more specifically, the fact that “special education students scored better in reading than all students.” The ALJ did not credit Olowokakoko’s testimony that Todd openly directed teachers to cheat during staff meetings. Although the ALJ expressed frustration over the lack of eyewitness testimony directly linking Todd to the cheating, he emphasized that because the cheating at Fort Worthington was orchestrated and carried out by “a few people secretly erasing wrong answers and changing them to right answers,” absent testimony from one of the perpetrators admitting his or her involvement, the case necessarily depended upon circumstantial evidence.

The ALJ rejected outright Todd’s defense that no cheating had taken place occurred at Fort Worthington. He reasoned that even if the anonymous tips that instigated the investigation had been made by teachers angered by the mold issue, that would not explain why the BCPS and MSDE investigations substantiated those allegations through statistical analyses. He found Dr. Gresham’s report, Dr. Pollack’s report, and the testimony of Drs. Gresham, Pollack, and Rudner very persuasive. He adopted their view that the number of erasures in the 2009 and 2010 test booklets for special education students was highly unusual and that the unusually high number of wrong to right answer changes was clear evidence of cheating. He rejected Dr. Lentz’s criticisms of Dr. Pollack’s report, characterizing them as “fairly trivial,” and noted that “[t]his was not a borderline case, where the erasure evidence was ambiguous.”

The ALJ also rejected Todd’s defense that if cheating did occur, she was not aware of it. He credited Todd’s own testimony that she was a “model principal with a command of everything going on in her school building, including up-to-the-minute knowledge of each child’s current performance level.”⁸ In addition, he relied on evidence showing that Todd was “data driven” and that the cheating began soon after she took over the helm at Fort Worthington. For all of these reasons, the ALJ concluded that Todd “made a conscious decision to violate many of the test security and data reporting policy and procedures set forth in COMAR 13A.03.04” and that the “magnitude of the cheating on the MSA during [Todd]’s tenure as principal” compelled the conclusion that she “was aware of the cheating, and in fact, directed the cheating.” Her testimony to the contrary was “not believable or credible.” In light of the “egregiousness” of Todd’s conduct, the ALJ recommended the sanction of revocation of her MSDE certifications.

E. Todd’s Exceptions

On October 5, 2012, Todd noted exceptions to the Proposed Opinion and Order. In her first exception, she argued that the ALJ erroneously identified one of her exhibits and part of another of her exhibits as “not admitted.” The second exhibit, Todd’s Exhibit 5, was “a series of charts of MSA results from 13 schools in Baltimore City” in “Area 9,” a group of struggling schools of which Fort Worthington was a part. She maintained that

⁸ The ALJ noted that Todd could have presented evidence in the form of her own records tracking individual students’ progress to rebut evidence that their MSA scores were anomalous, but did not do so.

Exhibit 5 showed that schools in Area 9 had similar MSA test score gains during the relevant period. In her second exception, Todd argued that the ALJ was biased against her and “her beliefs as an educator.” In her third exception, she argued that the ALJ displayed an “unfounded allegiance to [Dr.] Pollack.” In her fourth exception, she argued that the ALJ improperly relied on “anonymous and unsworn statements not subjected to cross-examination.”

F. The Superintendent’s Final Decision

On March 25, 2013, after hearing argument of counsel, the Superintendent issued the final agency decision. She adopted the ALJ’s findings of fact and conclusions of law and revoked Todd’s advanced professional teaching certificate.

The Superintendent summarized the ALJ’s “Key Findings” in a chart divided by year. These findings included that, in 2006/2007, Fort Worthington was a failing school that had not made AYP in years; but a year later, special education students at Fort Worthington were outperforming general education students there on the MSA and the special education students’ scores on the MSA reading exam had improved 70 percentage points. The Superintendent noted that Dr. Gresham had opined that a score increase of this magnitude was “statistically unlikely to have occurred by chance.” The Superintendent further noted that students who left Fort Worthington and attended other schools between 2007 and 2009 “showed significant decline in performance on assessments,” while students who came to Fort Worthington from other schools during that period “showed significant improvement in performance on assessments.”

Moreover, Fort Worthington students almost uniformly performed worse on benchmark assessment C, which was given shortly before the MSA each year, than they did on the MSA. The Superintendent observed that Dr. Pollack’s erasure analysis confirmed that an unusual number of erasures by special education students were made that changed wrong answers to right answers on the 2009 and 2010 MSAs. Test tampering also was corroborated by interviews with teachers and students detailed in Dr. Gresham’s report.

The Superintendent then turned to Todd’s exceptions. She agreed that the ALJ’s proposed decision incorrectly stated that two of Todd’s exhibits had not been admitted. She disagreed, however, that the exhibits, particularly Exhibit 5 showing MSA score increases at other BCPS Area 9 schools, rebutted the allegations of cheating at Fort Worthington or Todd’s involvement in the cheating. The Superintendent emphasized that “[t]he facts surrounding the substantial increases at other Area 9 schools [were] neither before [her], nor ha[d] [she] received allegations of test tampering or had the type of contextual, quantitative evidence presented challenging the test scores at any other Area 9 school than Fort Worthington.” Thus, even if the ALJ had erred by not considering those two exhibits in rendering his decision, Todd was not prejudiced by his mistake.

The Superintendent rejected Todd’s argument that the ALJ was biased against her and her educational methods, and against those who hold the expectation that low income students can excel on the MSA. She concluded that, viewing the ALJ’s statements in context, it was clear that he was skeptical that special education students would be able to improve at such a dramatic rate even with the implementation of new instructional

methods and professional development. The Superintendent explained that the ALJ's comment that Todd was being pressured to "accomplish the impossible" was not evidence of bias or prejudice, but a reasoned judgment based upon the evidence before him. The Superintendent stated that, although she "believe[d] it [was] possible for economically disadvantaged students to achieve remarkable success through excellent teaching, hard work, family and community support," the Fort Worthington 2008 MSA score increases could not be explained based on those factors.

The Superintendent also rejected Todd's assertion that, because Dr. Pollack's erasure analysis had been discredited in another case involving allegations of cheating at a BCPS school, the ALJ's reliance upon Dr. Pollack's testimony was improper. The Superintendent reasoned that the ALJ was not bound by findings made in another proceeding and, in any event, she herself was persuaded by Dr. Pollack's and Dr. Rudner's testimony that the number of wrong to right erasures in the 2009 and 2010 test booklets was a "statistical impossibility," absent malfeasance, and that Todd's evidence did not rebut that testimony.

Finally, the Superintendent rejected Todd's argument that the evidence before the ALJ was insufficient to prove that Todd personally directed or was involved in any cheating. The Superintendent emphasized that hearsay evidence, such as the anonymous tips and the interviews of unnamed students and teachers that Dr. Gresham performed, is admissible in an administrative proceeding if it has "sufficient probative value"; in fact, it "may even serve as the sole basis for the decision of an administrative body." The

Superintendent found that the anonymous calls made to the BCPS fraud hotline, as well as the May 10, 2010 anonymous letter to BCPS, were “sufficiently reliable and probative to have been admitted into evidence.” She noted that three of the five phone calls were placed within a few days of the administration of the 2010 MSA and they merely spurred the investigation by Drs. Gresham and Pollack. With respect to the May 10, 2010 letter, the Superintendent noted that the identity of the author of that letter was known to Dr. Gresham. For these reasons, the Superintendent concluded that the ALJ did not err by admitting into evidence the anonymous reports linking Todd to the cheating, even though the credibility of the declarants could not be tested.

Finally, the Superintendent reasoned that the “nature of the misconduct – particularly erasures in test booklets post test-administration – [was] such that it would be virtually impossible to have direct evidence of Ms. Todd’s actions.” Thus, the absence of “direct evidence” that Todd knew of and/or perpetrated the cheating was not a “fatal flaw.” Moreover, although Todd claimed in her response to Dr. Gresham’s investigation that she “played no role in test administration at Fort Worthington,” as she was neither a test coordinator for the school nor a test examiner, she clearly had a role in administering the MSA because she directed Arrington to test supposedly disruptive students separately. The Superintendent concluded that this evidence showed that Todd had a “deeper level of involvement in decisions regarding the test administrations” than she was acknowledging and supported the “ALJ’s conclusion that Ms. Todd’s staff could not have cheated on the 2010 MSAs without her knowledge.”

G. Judicial Review in the Circuit Court

On April 19, 2013, Todd filed a petition for judicial review in the Circuit Court for Baltimore City. The court heard argument and, on September 3, 2014, affirmed the final agency decision to revoke Todd’s teaching certification. This timely appeal followed.

STANDARD OF REVIEW

In an appeal from the judgment of the circuit court on judicial review of an administrative agency decision, we look through the decision of the circuit court and review the final agency decision. *See, e.g., Motor Vehicle Admin. v. Shea*, 415 Md. 1, 15 (2010) (citing *People’s Counsel for Balt. Cnty. v. Loyola College in Md.*, 406 Md. 54, 66 (2008)). In this case, the final agency decision was that of the Superintendent. As the Court of Appeals explained in *Maryland Aviation Administration v. Noland*, 386 Md. 556, 571-72 (2005), our review of the agency decision is circumscribed:

A court’s role in reviewing an administrative agency adjudicatory decision is narrow; it “is limited to determining if there is substantial evidence in the record as a whole to support the agency’s findings and conclusions, and to determine if the administrative decision is premised upon an erroneous conclusion of law.”

In applying the substantial evidence test, a reviewing court decides “whether a reasoning mind reasonably could have reached the factual conclusion the agency reached.” A reviewing court should defer to the agency’s fact-finding and drawing of inferences if they are supported by the record. A reviewing court “must review the agency’s decision in the light most favorable to it; ... the agency’s decision is prima facie correct and presumed valid, and ... it is the agency’s province to resolve conflicting evidence” and to draw inferences from that evidence.

Despite some unfortunate language that has crept into a few of our opinions, a court’s task on review is not to “substitute its judgment for the expertise of those persons who constitute the administrative agency.” Even with regard to some legal issues, a degree of deference should often be accorded the position of the administrative agency. Thus, an administrative

agency’s interpretation and application of the statute which the agency administers should ordinarily be given considerable weight by reviewing courts. Furthermore, the expertise of the agency in its own field should be respected.

(Citations omitted.); *See also Marzullo v. Kahl*, 366 Md. 158, 171-72 (2001); *Bd. of Physician Qual. Assurance v. Banks*, 354 Md. 59, 67-69 (1999).

DISCUSSION

Todd contends the Superintendent’s decision to suspend her teaching certificates must be reversed for three reasons. First, the Superintendent erred by adopting the ALJ’s proposed findings of fact and conclusions of law when the ALJ failed to consider two of Todd’s exhibits and displayed “a real and pronounced bias . . . against members of the education community who believe . . . that significant positive change and growth is possible amongst students in inner-city schools.” Second, the Superintendent improperly relied upon anonymous and unsworn statements implicating Todd in the cheating at Fort Worthington. Third, and relatedly, given that the anonymous and unsworn statements were the only evidence the Superintendent relied on that implicated Todd in the cheating and could support a finding that she willfully and knowingly violated the test security and data reporting regulations, the Superintendent’s decision was not supported by substantial evidence.

The Local Board responds as follows. The Superintendent appropriately adopted the ALJ’s proposed findings of fact and conclusions of law and rejected Todd’s argument of bias on the ALJ’s part. Moreover, the anonymous reports of fraud properly were admitted not for their truth, but to show why the BCPS began its investigation into testing

improprieties; and Todd has waived any argument that she was prejudiced by her inability to cross-examine the author of the May 10, 2010 letter. In any event, the Superintendent did not rely on the anonymous statements in reaching her ultimate decision. Finally, the statistical and quantitative evidence substantiating that widespread cheating had taken place and the reasonable inference, drawn from the evidence that Todd was a highly data-driven and involved principal, supported the Superintendent’s finding that the cheating could not have taken place without Todd’s knowledge and approval.

Todd’s first argument – that the ALJ failed to consider two of her exhibits and was biased against her – was raised in identical fashion in her exceptions to the ALJ’s proposed decision, and was considered and rejected by the Superintendent. We perceive no error. The Superintendent agreed with Todd (and the Local Board) that the ALJ mistakenly identified two exhibits as not having been admitted in evidence, in whole or in part.⁹ The Superintendent herself considered that evidence in reaching her independent decision to adopt the ALJ’s proposed findings of fact and conclusions of

⁹ As discussed, one of those exhibits, Todd’s Exhibit 5, showed test scores at other BCPS schools in Area 9. The other exhibit, Todd’s Exhibit 1, was an August 2010 email exchange between Dr. Pollack and an MSDE employee. In one email, Dr. Pollack told the MSDE employee that she would need to review 2010 MSA test booklets from “all of the grade 4 and at least half . . . of those from grades 3 and 5” in order to complete her erasure analysis. Dr. Pollack ultimately only reviewed the 2010 test booklets for the special education students at Fort Worthington. According to Todd, this email showed that Dr. Pollack herself believed that her sample size was too small.

law. On this basis alone, the mislabeling of these two exhibits as “not admitted” did not prejudice Todd. Moreover, the Superintendent reasonably rejected Todd’s assertion that evidence that other BCPS schools had shown similar improvement in their MSA scores during the same period of time established that cheating did not occur at Fort Worthington and rebutted Dr. Pollack’s and Dr. Rudner’s testimony that the number of erasures in the special education students’ test booklets during the 2009 and 2010 MSAs, and, in particular, the unusual number of wrong to right answer changes, could not have happened absent adult intervention and malfeasance.

Todd’s argument that the ALJ demonstrated bias against her educational philosophy also is without merit and properly was rejected by the Superintendent. The handful of quotations Todd cites from the ALJ’s proposed decision plainly were not evidence of any improper extra-judicial bias against her. *See Jefferson-El v. State*, 330 Md. 99, 107 (1993) (opining in the context of a recusal analysis that when a judge acquires knowledge “in a judicial setting, or an opinion arguably expressing bias is formed on the basis of information acquired from evidence presented in the course of judicial proceedings before him, neither that knowledge nor that opinion qualifies as personal”) (quotation marks and citation omitted). The Superintendent, as the final decision-maker, stated that she believes that low-income students are capable of “remarkable success,” but she nevertheless was not persuaded that the 2008 test score improvements at Fort Worthington were solely the result of improvements in the

instruction they received. Thus, if there was any bias on the part of the ALJ—and the record does not support that there was—it clearly did not infect the final agency decision.

We now turn to Todd’s arguments that the Superintendent improperly relied upon “anonymous and unsworn statements not subject to cross-examination” in reaching her final decision and there was no other credible evidence linking Todd to the testing improprieties. As discussed, Todd advanced the alternative defenses that no cheating had occurred and, if cheating occurred, she did not willfully and knowingly participate in it. She asserts that, although the statistical and quantitative evidence was sufficient to allow a finding that cheating took place, it was not sufficient to support a reasonable inference that she directed or participated in the cheating. According to Todd, the only evidence to support such an inference was Olowokakoko’s testimony, which the ALJ rejected as not credible, and the anonymous phones calls and letter, which should not have been admitted.

During the ALJ hearing, Todd objected to the admission of a summary of the five anonymous phone calls and to the May 10, 2010 letter. She argued that, although hearsay is admissible in an administrative proceeding, the anonymous nature of the complaints deprived her of her right to confront those witnesses. The Local Board responded that the exhibits were admissible as non-hearsay to show the “effect on Dr. Gresham and why he, in fact, did what he did.” The Local Board noted, moreover, that the anonymous complaints had been produced to Todd’s attorney in discovery and she had not subpoenaed Dr. Gresham or the BCPS to determine the identity of any of the callers or

the author of the May 10, 2010 letter. The ALJ admitted the exhibits, ruling that they were sufficiently reliable and probative to be considered for their “substance.”

It is well-settled that evidence that is hearsay may be admitted in an administrative proceeding. *See, e.g., Md. Dep’t of Human Res. v. Bo Peep Day Nursery*, 317 Md. 573, 595 (1989). In a contested case hearing, an ALJ “may admit probative evidence that reasonable and prudent individuals commonly accept in the conduct of their affairs and give probative effect to that evidence.” Md. Code (1984, 2014 Repl. Vol.), § 10–213(b) of the State Government Article. “Nonetheless, while administrative agencies are not constrained by technical rules of evidence, they must observe basic rules of fairness as to the parties appearing before them so as to comport with the requirements of procedural due process afforded by the Fourteenth Amendment.” *Travers v. Baltimore Police Dep’t*, 115 Md. App. 395, 411 (1997).

We agree with the Local Board that the summary of the anonymous phone calls was admissible as non-hearsay to show why the BCPS initiated an investigation into allegations of cheating at Fort Worthington. Significantly, the Superintendent did not rely upon the substance of the phone calls in reaching her ultimate decision.¹⁰ In her

¹⁰ Todd emphasizes that in the final agency decision, the Superintendent “recites the various anonymous allegations, and does so in a detailed fashion.” The Superintendent only did so because Todd excepted to the ALJ’s admission of that evidence at the hearing. Thus, the Superintendent summarized the phone calls before determining that the ALJ did not err by admitting them.

Similarly, although the ALJ summarized the anonymous phone calls in his findings of fact, he did not mention the substance of the phone calls in his discussion of

(Continued...)

decision, the Superintendent states that the credibility of the anonymous sources could not be tested at the hearing, and then turns to examine other evidence that supported the ALJ's conclusion that Todd willfully and knowingly participated in the cheating. Accordingly, any error in admitting the anonymous calls for their truth did not prejudice Todd.

With respect to the May 10, 2010 letter, Todd emphasizes that the ALJ explicitly credited and relied upon the author's "version of how [Todd] and her subordinates carried out a surreptitious and cynical cheating operation." This was improper, in Todd's view, because it was impossible to test the credibility of the anonymous source and she was deprived of her right to cross-examine the witnesses against her. She argues that the Superintendent's decision was infected with this error and must be reversed.

We agree with the Local Board that Todd has waived any argument that she was denied an opportunity to cross-examine the author of the May 10, 2010 letter. In the letter, the author identifies herself as having been a third grade teacher at Fort Worthington for two years, from August 2006 until June 2008. At the conclusion of the letter, the author states that he or she would "like for this to remain as anonymous as possible," but that he or she recognizes that there are "many indications in this letter as to

(...continued)

the evidence and clearly did not rely on those anonymous complaints in finding that Todd willfully and knowingly participated in the cheating. Thus, the Superintendent, by adopting the proposed findings of fact and conclusions of law, also did not rely on those phone calls.

who I am.” Given these identifying details in the substance of the letter, it seems unlikely that Todd did not know the identity of the author before the ALJ hearing commenced.

In any event, on the first day of the ALJ hearing, February 21, 2012, Dr. Gresham testified that he had interviewed the author of the May 10, 2010 letter as part of his investigation, but could not recall the author’s name at that time. He stated that he “ha[d] the information,” however. On cross-examination that same day, Todd’s counsel asked Dr. Gresham if he “happened to remember the name of the author.” Dr. Gresham replied, “I don’t. I said I have the information, but I don’t recall a name.”

As mentioned, the hearing spanned several months and did not conclude until June 25, 2012. Despite learning on February 21, 2012 that Dr. Gresham knew the name of the author of the May 10, 2010 letter and had “information” identifying that person, Todd took no steps to subpoena Dr. Gresham’s records to ascertain the identity of that teacher over the course of the four months during which the ALJ hearing was continued.¹¹ In light of this inaction, Todd may not be heard to complain that she was deprived of the opportunity to cross-examine the author of the May 10, 2010 letter. *See Travers*, 115 Md. App. 417-19 (holding that a police officer challenging his termination for misconduct waived any argument that the admission of hearsay complaints against him

¹¹ In her reply brief in this Court, Todd asserts that she already had “all of Dr. Gresham’s records” and that those records “did not identify the author of the May 10, 2010 letter.” It was clear from Dr. Gresham’s testimony, however, that the information about the identity of the author of the May 10, 2010 letter was available.

deprived him of the opportunity to cross-examine the hearsay declarant when he failed to subpoena the declarant).

Even if not waived, however, any error in admitting the May 10, 2010 letter was harmless. As Todd herself points out, the May 10, 2010 letter does not actually implicate her in the cheating. Rather, it describes the author's concerns about improprieties based on her knowledge of her students' proficiency compared to their performance on the 2008 MSA, suspicious erasures on test booklets, and suspicious behavior by ISTs during the administration of the 2008 MSA. These concerns spurred the investigation by Dr. Gresham, which substantiated that cheating had occurred. Dr. Pollack in turn conducted her erasure analysis, which substantiated that Fort Worthington staff had changed answers on the 2009 and 2010 MSA tests for special education students. While the ALJ suggested that the allegations in the May 10, 2010 letter set forth a more plausible version of events (in contrast to Olowokakoko's testimony), he made clear that he was relying on the statistical and quantitative evidence – not the May 10, 2010 letter – in finding that there had been widespread cheating perpetrated at Fort Worthington. The Superintendent, in turn, adopted the ALJ's findings in this regard.

We now turn to whether there was substantial evidence in the agency record to support the final agency finding that Todd knowingly and willfully directed, participated in, encouraged, or failed to report test security and data reporting violations. Circumstantial evidence alone may create an inference of knowledge. *See, e.g., Taylor v. Giant of Md.*, 423 Md. 628, 659 (2011) (as an element of a claim for retaliation in

employment, circumstantial evidence may be used to prove that an employer knew an employee had filed a discrimination claim). We “defer to the agency’s factual findings *and inferences* that are supported by substantial evidence.” *Motor Vehicle Admin. v. Karwacki*, 340 Md. 271, 280 (1995) (emphasis added). Moreover, it is the province of the agency, not a reviewing court, to decide between competing inferences that may be drawn from the same evidence. *Bullock v. Pelham Wood Apartments*, 283 Md. 505, 513 (1978).

Here, from the undisputed evidence that Todd was a data-driven and detail-oriented principal who knew everything that happened at her school, the Superintendent drew rational inferences that Todd knew of and directed or encouraged the cheating at Fort Worthington. The Superintendent adopted the ALJ’s reasoning that it was not possible for teachers to have “acted on their own to cheat on the MSA” under the leadership of a principal like Todd, who maintained tight control over every aspect of the instruction at her school. Todd’s detailed knowledge of student performance supported this inference because, had teachers surreptitiously tampered with test booklets, it would have been immediately apparent to Todd that the students’ MSA scores were inconsistent with their performance up to that point in time. Todd’s testimony that the 2008 test score improvements did not raise any red flags in her mind was deemed “not believable or credible” by the ALJ, and the Superintendent adopted that finding. The Superintendent found, moreover, that, while Todd attempted to distance herself from the administration of the MSA at Fort Worthington, there was evidence that she micromanaged aspects of

the testing, such as the removal of disruptive students from their classrooms for testing in smaller group settings. All of this evidence, coupled with the “magnitude of the cheating” at Fort Worthington over several years during Todd’s tenure, constituted substantial evidence that Todd willfully and knowingly committed violations of test security and data reporting regulations by “directing violations” of those regulations, including the alteration of answers; by failing to properly monitor test administration; or by failing to report testing violations that were known to her.

**JUDGMENT OF THE CIRCUIT
COURT FOR BALTIMORE CITY
AFFIRMED. COSTS TO BE PAID
BY THE APPELLANT.**