

Changes in academic achievement and the use of online learning applications in Pittsburgh Public Schools during remote instruction in the COVID-19 pandemic Webinar transcript

Friday, November 19, 2021 1:00–2:00 p.m. Eastern

[webinar producer] Hello and welcome to today's webinar, Academic Achievement and Online Learning Application Use in Pittsburgh Public Schools During Remote Instruction in the COVID-19 Pandemic. Before we begin, please take note of the following functionality features. The webcast console you are viewing can be completely customized. You can expand each window by clicking on the maximize icon in the top right corner of each window or by clicking and dragging the bottom right corner of each panel to set your preferred display and layout. There are a series of interactive buttons on the bottom of your screen that you may explore.

The slide deck and additional resources are available for download in the Resource List widget here in your console or via the green file icon. If you have any questions for presenters during the webcast, enter them using the Q&A widget or by clicking the purple Q&A button at the bottom of your screen. You may submit questions at any time.

If you have technical difficulties, please click on the Help button represented by a yellow question-mark icon. You may also submit technical questions through the Q&A window and we will do our best to address your needs. Many technical issues can be resolved by pressing F5 on PCs or Command R on Macs to refresh this console. This webinar is being recorded. Upon conclusion and processing of this event, the recording will be available via the same link you used to join today. Now I'm pleased to turn the webinar over to Brian Gill. Brian, the floor is yours.

[Brian Gill] Thanks so much, Rick. Welcome everybody to our presentation on findings during COVID and remote learning in Pittsburgh. This is a presentation sponsored by the Mid-Atlantic Regional Educational Laboratory. I am the director of that laboratory. For those of you who aren't familiar with the Regional Labs, or the RELs as we call them, they are organizations funded by the U.S. Department of Education to support research and analysis and dissemination of research and analysis in regions. The Mid-Atlantic lab serves states and local stakeholders in Pennsylvania, Maryland, New Jersey, Delaware, and the District of Columbia. We did this work in collaboration with the Pittsburgh Public Schools.

So, I'm here with my colleagues from the REL Mid-Atlantic, Whitney Kozakowski and Patrick Lavallee, along with Ted Dwyer who is Chief of Data, Research, Evaluation, and Assessment in Pittsburgh. Thanks, everyone, again for being here. We're going to go through and tell you what we've learned about student learning and the pandemic in its first year. And we're looking forward to taking some of your questions at the end. You can feel free to submit them throughout, as Rick just mentioned. So, many of us, not just in Pittsburgh but across the country, have been concerned about how the pandemic and associated disruptions to instruction affected student achievement. You're probably aware of national evidence that students didn't do as well in the 2021 school year as in a typical year. There's also evidence from other places that students who learned remotely during the last school year not only experienced less instruction but also had more difficulties in getting work done than those learning in person. And there's evidence that student participation in online learning applications declined relative to some pre-pandemic participation.

Meanwhile, there were reports from districts – a handful of districts across the country suggesting that an increased number of students were receiving failing grades. So, not just declines on standardized assessment results but in course grades as well. And Pittsburgh was interested in doing some deep analysis of how its students were doing during the pandemic. So, the data here are all from the Pittsburgh Public Schools, but, as I think you'll see, they have implications that are much broader and I believe they have implications that are broader than Pittsburgh for a couple of reasons.

One is that the findings, in general, are pretty consistent with findings that have been found in other places across the country on the pandemic year. Another is that, as you all know, the kinds of challenges that schools and students faced last year, and to some extent this year as well, are challenges that are being faced by schools and students all over the country. So, what we have are some findings that are broadly consistent with findings nationally but that allow us to dig a little deeper on some of these findings because the data we are looking at in Pittsburgh are so rich. I'm going to give you a quick preview of those findings before I turn it over to my colleagues Whitney and Patrick to talk about the findings in detail, and then we'll turn it to Ted Dwyer to talk about what Pittsburgh is doing with this information.

So, the first finding is that, on average, students in Pittsburgh, in most grades, experienced test score growth. So, we don't like to use the phrase "learning loss" because, in fact, in absolute terms, they did experience growth from the winter immediately preceding the pandemic to a year later. But the growth was less than what you'd expect in a typical year based on national norms. The growth lag – we'll call it a lag rather than a loss – was largest for students in elementary grades, younger kids, and it's generally consistent with some evidence of growth lags nationally.

Meanwhile, when we look at the course data, the failure rates increased substantially, especially for the older students in middle and high school. They increased a little bit more among economicallydisadvantaged students than others. And chronic absenteeism we find really strongly predicted course failure. We'll talk in more depth about the connections among chronic absenteeism, course failure, and participation in the online learning system.

So, data from that learning management system showed that there was a decline in logins and in the rate of completion of assignments over the course of the school year during remote instructions as Pittsburgh went from fall to spring during a time when instruction was entirely remote. Those daily participation data show that there is a strong relationship between those data and chronic absenteeism and course failure,

suggesting that there really is an identifiable group of students who were most negatively affected by the pandemic and remote instruction and who are going to need the greatest attention in the current school year and beyond. So, that's the preview. I will now turn it to Whitney to tell you in more detail about the data and the specific findings there. Go ahead, Whitney.

[Whitney Kozakowski] All right. Thank you for that introduction, Brian. So, I'm going to first start off by sharing a little bit about the test scores that we'll be using for the test score analysis. So, we'll be using data from the NWEA MAP exams. The MAP exam is different from your standard state assessment in a couple of important ways that inform the analysis. So, first, the MAP exams are being offered in fall, winter, and in spring, unlike the typical state assessment which is just done once a year.

Another important feature is that it's what's called a vertically aligned test. So, in a typical state assessment, they are typically not vertically aligned. So, in those cases, you wouldn't want to take, say, the state score from a third-grader and compare that to a fourth-grader, and you wouldn't have the expectation that third-grade scores relative to fourth-grade scores, that the fourth-grade scores might be larger. The NWEA MAP exams are different in that sense because students are all placed on the same scale. So, what you would expect to see in a typical year is that students' scores would increase from, say, grade three to grade four to grade five. Because it has this nice feature, we'll be able to actually look at how students grew during the pandemic even though they were moving across grades.

We will also do a separate analysis in which we standardize the scores relative to national norms. These national norms are based on pre-pandemic data. So, what this will help us to do is to contextualize the changes that we're seeing in scale scores so that we can talk about how this might compare to a more typical year. So, I'll get to that in a moment when we actually get to the analysis.

We'll also be focused on students that are in grades two to eight for the test score analysis. This is because the test-taking rates are just lower in the other grades. As Brian mentioned, we'll be focused on this time period of winter 2019/20, just before the pandemic hit and disrupted instruction in spring of 2020. We'll be comparing those test scores from winter of 2019/20 to winter of 2020/21. We'll also be using pretty standard demographic data. This will include race and ethnicity, gender, economically-disadvantaged status, absences, as well as students' IEP status.

For the grade analysis, we'll be focused on the first semester of 2019/20, that will be our baseline period, and then we'll be comparing that to the first semester of this past academic year, after almost a year of remote instruction at that point. Lastly, we'll be looking at data from Pittsburgh's Learning Management System, and this is used by students in all grades in Pittsburgh, so K to 12, and we will have daily records of students, whether or not they log into the system and what they do in the system. So, we'll look at the number of materials, course materials that they open and submit. We can also see information about whether they're submitting assignments, taking assessments, or actually posting to discussion boards.

So, as I just alluded to, the first question that we're going to be looking at here is whether or not students actually learned, as measured by the test, at all during the period of remote instruction. For this analysis, what we'll be using are the scale scores from the NWEA MAP exams. What we'll do is we'll take individual students who took the test in the base period, so winter of 2019/20, and we'll look for those individual students and see how their score changed from that period to winter of 2020/21. We'll calculate a growth score for each student and then we'll take an average of that growth score for all the students who took the test in both periods. Now I'll talk through the findings for that.

Okay. So, we're first going to look at the change in math scores. So, over in this figure on the right, the yaxis here is showing you the change in scale scores. You can think of this as the average growth for individual students in their math scores. And then this x-axis is showing you the grade of the student in the base year, so in 2019/20. The gray bar is showing the change in the scale score from winter to fall and then the blue bar is extending that and then looking at winter to winter.

So, what you'll first notice here is that the median change is generally positive across all grades, with the exception of fifth grade for the winter to fall period, and that it continues to grow for the winter-to-winter period. And the main takeaway here is that during this period of remote instruction students were learning things that were measured by the tests and their scale scores were going up on the MAP exams as a result of that. You can see that they were going up not only from the winter to fall period, despite all that disruption that occurred in the spring, but they continued to go up from the fall-to-winter period when they were continuing to learn remotely this past academic year, too.

Next, we shift over to the reading scores. Again, you see a similar pattern here in that, from winter to fall, students' scores increased typically and in all grades for reading, and that they then continued to grow from winter to winter – rather, from fall to winter, so that you see that from winter to winter it's also a positive change. As with math, the key takeaway here is students did learn. Their test scores went up, but, of course, the second question is how does that growth compare to a more typical year when students weren't learning remotely and then they weren't experiencing all the disruptions of the pandemic. So, that's what we're going to turn to next.

For this analysis, we're going to use the same general method of taking the difference for individual students, but instead of taking just their scale scores and looking at that change in scale score units, what we're going to do is take the scale scores, standardize them relative to the pre-pandemic national distribution in both periods before differencing them. So, what that means in practice, if you see that a student's score stays the same in standard deviation units, what that means is that from one year to the next, say moving from third grade to fourth grade, the student's scores increased enough to maintain the same position in the national distribution. So, if you are at the 50th percentile in the baseline year, you would, again, be at the 50th percentile the next year if the change was zero for you.

If, instead, you see a decline in standard deviation units, what that means is that the student did not grow sufficiently to stay at the same position in the national distribution. In other words, their growth was lower than what we would have expected average growth to be for students starting at that place in the distribution. So, next, we'll go ahead and take a look at this.

So, on the y-axis here, you can see the change in standardized scores in standard deviation units. This is the graph for math. You can see students in grades – that were starting in grades two through seven in 2019/20 who then would have moved to grades three through eight the next year. As with the prior graph, you can think of this as – so, for example, the second-grade bar, this would be what was average growth in standard deviation units for students that were starting in second grade at the start of the pandemic.

What you see across the board here is that, on average, there's a decline of about .15 standard deviations in math. That's about the equivalent of six percentile points. So, if Pittsburgh students in these grades were, on average, scoring at, say, the 40th percentile, they would then be scoring at the 34th percentile at the end of this period, by the winter 2021 period. Of course, there's a lot of variation here, though, that's pretty interesting.

So, if you look at students that were starting in grades six and seven, there's no decline for them in standard deviation units, which means that they grew about what you would have expected them to grow over this period of remote learning. The lags are entirely concentrated among students that were in the elementary school grades at the start of the pandemic, so these students in grades two through five. The other thing I would add is that the test-taking rates were slightly lower in 2020/21 versus 2019/20, and that's pretty consistent nationally. You see this pattern in test-taking rates across the U.S.

So, in Pittsburgh, we do a couple things to address this. So, one, we focus on calculating growth for individual students. Two, we also impute scores. We do an imputation analysis just to see what – how our results might have changed if we imputed scores for those that were missing them. The general takeaway from that exercise is that once we imputed scores the findings were quite similar to what we're seeing when we just calculate growth and ignore the students that we don't have the follow-up score for. So, across the board, we think the findings are pretty consistent.

Next, we move onto reading. So, on average, across students that are in grades two through seven at the start of the pandemic, we're seeing declines of about .1 standard deviations in reading in standard deviation units. So, what this means in practice, this is about four percentile points, but, again, you see the same pattern where students that were already in middle school grades, so grades six and seven, you barely see any lag in their scores. The lags are very much concentrated in the elementary school students.

Next, we move over to looking at differences across different groups of students. So, for this analysis, we're looking at the math score change and we're using all of the students in grades two through seven and then we're splitting them into their different demographic groups. So, we can see here we're reporting results for boys and for girls, for Black and white students, for those that were economically disadvantaged or not, and those that have an IEP plan or those that don't. The main takeaway here is that we see larger declines for boys. They're about -.2 standard deviations compared to a decline of only .1 standard deviations for girls. So, that's the biggest difference. The differences for other demographic groups are fairly small in magnitude.

Next, we're going to shift over to reading. So, you can see here, again, there are declines for all of the groups, but the biggest difference between groups is actually between Black students who have larger declines than white students. The decline for Black students is about .15 standard deviations compared to .04 standard deviations for white students. While there are some differences for the other demographic groups, they're not as large as the Black-white difference that we're seeing.

So, next, we're going to take a look at how course failure rates have changed during remote instruction. The main takeaway, as Brian was mentioning, is that we do see large increases in course failure rates very much concentrated in grades six through 12. We also see larger increases for chronically absent students and those that were economically disadvantaged.

Okay. So, this chart is showing the percentage of students that are failing at least one course. This is their grade in 2020/21 or in 2019/20, depending on the bar. So, these blue bars are for 2019/20, the orange bars are for 2020/21. It shows the share of students that are failing at least one course in the fall semester. What you'll see here is that, starting in about grade five and all the way up to grade 12, we're seeing fairly large increases, certainly over five percentage points, in the share of students failing at least one course.

So, for example, in this eighth-grade bar, what you see is you're moving from about nine percent of eighth-graders failing at least one course to almost over 20 percent failing at least one course. So, these are large increases in course failure rates. We also looked at the grade distribution to see is this just failure rates that are increasing or are we seeing shifts in the grade distribution more broadly. What we see is that it's actually just a general downward shift in the grade distribution. So, we're seeing fewer As and Bs and an increase in the number of Ds and Fs across the board in grades six through 12.

Next, we did the same analysis where we look for all the grades, so for grades one through 12, and then look for differences in the change, rather the increase in the percentage of students failing at least one course by demographic group. And what you can see here is that there's been a nine-percentage-point increase in the share of students that are economically disadvantaged that are failing at least one course. But for those that are not economically disadvantaged, there's been barely any increase. It's about a fourpercentage-point increase for those students. You can also see slight differences by race here as well.

The biggest change by far is between students that are chronically absent versus not chronically absent. So, for those that were chronically absent in the first semester of 2019/20 compared to those that were chronically absent in the first semester of 2020/21, there's been a 20-percentage-point increase in the share of these students that are failing at least one course, which is enormous when you compare that to the share that were not chronically absent. For them, there's only been a four-percentage-point increase in the share failing at least one course. So, it really suggests that the students that were chronically absent, that that's where most of the course failure rate increase is occurring. So, this sparked us to look a little bit more closely at what was going on with chronic absence rates. So, we first wanted to know, well, how did chronic absence rates change between 2019/20 and this past academic year. You'll see here that the share of students that count as chronically absent, so that's missing more than ten percent of instructional days in the first semester, it's about the same, it's about a quarter of students in both periods. But what's changed is that the share of – rather, the number of days that they're absent has increased quite a lot.

So, for chronically absent students, they were only missing about 19 days, on average, in 2019/20 for the first semester, and then that bumped up to 27 days, on average, in this past academic year for the first semester. So, that's a big increase. And then you can also see that there's this very strong relationship between the number of courses that students are failing and then the number of days that they're absent. So, for students that aren't failing any courses, they're missing fewer than five days in the first semester of 2020/21, compared to students who are missing – rather, who are failing three or more courses are missing something like 33 days, and that's almost the equivalent of about a third of the first semester. So, I'm next going to turn this over to Patrick who's going to talk about some of the learning management system data and what we can learn from that.

[Patrick Lavallee] Great. Thank you, Whitney. So, as Whitney mentioned, we'll actually look at how much and in what ways student access and use learning management system while learning remotely. Pittsburgh's own Schoology, which students use to access classroom resources and also click through to join class virtually. The data we get from Schoology gives us a measure of student participation and engagement. We can see when students log into Schoology, which course materials they open, the assignments, assessments, and discussion posts that they submit. So, at a high level, we find that students who are chronically absent or failing their courses logged in on fewer instructional days and opened and submitted fewer course materials on average per week.

So, we'll start by looking at this with student-tracked course materials. The chart on the left shows the students who fail more courses open and submit a decreasing number of course materials. The count of course materials opened includes any resources which are posted to the classroom Schoology, so this would include handouts, study guides or links to external resources, but not anything that students can also submit, such as assignments, assessments, and discussion posts to not overcount interactions with the same items.

So, focusing on just the course materials that they submitted, students who did not fail any course submitted a little over eight items per week while those who failed three or more courses submitted less than three per week. We think that this is a big difference that adds up over the course of the semester. The chart on the right describes a similar contrast but by chronic absenteeism. Again, just narrowing on course materials submitted, students who are not chronically absent submitted almost nine items per week while those who were chronically absent submitted just shy of 4.4 or less than half. We also find that logins show a similarly strong relationship with course failures and chronic absenteeism.

Now, this chart looking at logins has a circle from each instructional day during remote learning colored by instructional day type. The y-axis shows a percentage of students who logged into Schoology while the x-axis shows which instructional day during the school year that we collected this data. So, you'll notice that most days are full, synchronous days, which are in red, and that fewer students log into school during other instructional day types. So, for example, full asynchronous days, which are in black, tend to have many fewer students logging into Schoology at all. We also see this steady decline through the year in course materials submitted. Just as a point of reference, students returned for winter break on day 74, so about halfway through this graph.

Now, remember that students are also logging into Schoology not only to open and submit homework but often to click through to join class. So, this pattern, if true across the country, might explain the slower learning the NWEA found in the second half of the last school year. That was a study released this summer that compared spring-to-spring scale score growth.

So, wrapping up here, students who are Black, economically disadvantaged, were on an IEP, were chronically absent, or have lower test scores opened fewer course materials each week. This chart shows those differences. So, you'll see that the difference exists where we would expect them, unfortunately. So, for example, you'll notice students open more course materials as they moved into higher math test score quartiles. We find similar patterns across characteristics, these same characteristics when we look at the average course materials submitted each week as well as the percentage of days logged into Schoology. With that, I think I'll turn it back to Whitney for implications.

[Whitney Kozakowski] Great. Thanks, Patrick. I think we just want to take a moment to talk through some of the takeaways from this work. So, the first thing we'll point out here is that there's evidence now that both older and younger students appeared to struggle during the pandemic and during remote learning. So, we're seeing the largest lags are for the students that are in elementary school grades but we're also seeing for the middle and the high-schoolers there are really large increases in the share that are failing at least one course. So, there's evidence of distress for across the age range, unfortunately.

The fact that the course failure rates were concentrated among the students – so heavily concentrated, I should say, among the students that were chronically absent, and that this group was also the one that was opening and submitting fewer materials and logging in for fewer instructional days, it suggests that that's a group that could certainly use additional support, that school districts can identify, that Pittsburgh can identify and can offer additional support to help them reengage and catch up.

The other thing I would say here is when we look at the magnitude of the declines in test scores, the lags that we're seeing, I think that there's both cause for concern and also cause for hope. So, the concern is that these lags are substantive and we need to be putting resources toward addressing them. Business as usual and just hoping that they go away is probably not going to be enough. But that said, the size of the lags, they're about the size of some positive educational interventions that we know can raise test scores by the amounts of the declines that we're seeing. So, for example, high-dosage tutoring has effect sizes. Those kinds of interventions have effect sizes that are large enough to address the kinds of lags that we're seeing. So, there's a range of educational interventions that could be used to help target the lags that we're seeing, but it's a matter of actually organizing, getting the resources to do it, and enacting them, enacting effective programs to address them.

The fact that we're seeing these declines in the use of the learning management system over the school year, it suggests that perhaps engagement might have declined over the school year under remote instruction. So, one thing I'll add, what Patrick was showing you in the fraction of instructional days that students log in, that was all under remote instruction. It wasn't this hybrid period. We didn't look beyond the point when students actually returned to the classrooms. And so the fact that we're seeing a decline over that period of time suggests that perhaps the students were just logging in less as the year kind of went on, and that's consistent with some patterns that have been seen in the Zearn data, too, which is another online learning application that students use.

The other thing I would add here is that the fact that we're seeing, for the asynchronous days, students are a lot less likely to log in suggests that if school districts that have to have a return to remote learning need to do it again, that they might want to focus on the synchronous days because that seems to be better at helping students actually engage with their coursework as measured by the logins to Schoology. With that, I'm going to hand this over to Ted Dwyer who's going to talk to us about how Pittsburgh's kind of processing these findings and what some of their takeaways form this work is.

[Ted Dwyer] Thank you, Whitney. So, I'm Ted Dwyer. I'm the Chief Accountability Officer over Data, Research, Evaluation, and Assessments at Pittsburgh Public Schools. Before I start, I just wanted to say that the REL assistance on this project has been a force multiplier for our district. We had a lot of questions and we all have a lot of data. While we knew we had the personnel who had the acumen to conduct the analysis, with our operational tempo over COVID and all of the new, different data types that we had, we knew we couldn't dig into the data like REL could. So, thanks, IES, for continuing to fund the REL. And thanks, REL, for being a valuable partner to us and the region.

So, how has the work been received in Pittsburgh Public Schools? We've used it in a lot of different ways. The most visible use of the work is that we've used it to inform conversations in relation to the ESSER funding, both internally and with our community partners. It's helped to frame the discussion and recommendations from the community for using the ESSER funding. Internally, it's provided confirmation and clarity towards some of the learning lag our students have experienced, and it's also reinforced the importance of leveraging our attendance flags in the early warning system to identify and address students who are not in school to receive their education. Additionally, we had the REL team present results at a board meeting which was also a public meeting, and we look forward to having them share additional results as they become available.

The steps that the districts are taking to address learning lag and reengaging students, we started the school year with an explicit focus on providing a venue to reestablish connections with our students who had not been in school for 18 months, and that opportunity was provided by setting the expectations for teachers to focus on social emotional activities, reengage the students, and the core curriculum for the first nine weeks. To support the effort, the curriculum department communicated the use of supplemental materials would not be monitored during the first nine weeks. And based on usage and click-through information, most of our supplemental materials were not being used as much as they had in the past.

The challenges that we're – what challenges are there in addressing learning lags and reengaging students? So, there's been a host of challenges for adults and for students. So, we started out the school year with a delay with minimal early warning because our external contracted transportation services didn't have enough bus drivers. Transportation has historically been and continues to be a challenge.

Reestablishing rituals and routines in the schools for both the adults and the students, we see this reflected in higher discipline incidents, attendance, both because of transportation and continuing chronic absence issues. Establishing workflows, processes, and expectations for how to leverage and use individual student devices. We were a none-to-one district and now we're a one-to-one district.

What I think is the largest challenge is the social, emotional, and mental adjustment to what everyone has been going through and what they've been dealing with. I did not say that first because it pains me to only have extemporaneous, qualitative data and, as a mixed method researcher, I think that that's not enough. Finally, once again, I mentioned it earlier but attendance. It was a pre-pandemic issue, especially in middle and high school, and it continues to be an issue. We've got an average daily attendance that lags prior years and it's been exacerbated by transportation issues and other issues across the district, and we're still in a pandemic. So, those are probably the biggest things that we're trying to deal with. Thank you.

[Brian Gill] Great. Well, thank you so much, Ted. I'll just add that it has been great for us researchers at the REL to have the opportunity to work with Pittsburgh on this project as well. And I do hope that it's been useful for Pittsburgh and that it will be useful for others across the country. So, we've got now plenty of time to do some Q&A. And I see that we've got a handful of questions that have come in already. So, I'll get started on some of those. Meanwhile, anybody else who has a question, feel free to type it in the Q&A box and we'll get to as many as we can in the next 24 minutes or so.

The first one I think is a quick one for Ted, which was, how does Pittsburgh define what economically disadvantaged means since that's a particular term that's different than what some other districts use, I know?

[Ted Dwyer] Yes. So, the district is a CEP district, Community Eligibility Program district. So, we don't collect free and reduced lunch information. We collect direct certification information. So, economically disadvantaged would be students who are identified in that process. The multiplier, if you use the federal multiplier, we're actually 100-percent free and reduced lunch, the entire district. So, that represents the direct cert percentages.

And the economically disadvantaged information comes from a different source than the free and reduced lunch?

It comes from our lunch program but by – in order to participate in the CEP program, we cannot collect the traditional free and reduced lunch information. So, it is flowing in and through our lunch program, but it doesn't include students who are not being direct certified and who are eligible for reduced lunch, either.

[Brian Gill] Okay. Thank you. Okay. Here's one for Whitney. We have somebody asking, "Are these just descriptive data? What sort of statistical analyses" – by which I think he means significance tests, for example, was any of that done, and what about interactions between variables? Did you take a look at, say, results for Black students who were chronically absent, or other sub subgroups, so to speak?

[Whitney Kozakowski] Yeah, I'm happy to talk more about the methods used and why we used them. So, I'll start off by saying, identifying the effects of a pandemic are very hard because everybody has been exposed to the pandemic and there's no true comparison group of students that have not been exposed to a pandemic. So, everything that we've done to-date is descriptive. That said, something we've thought a lot about is what is a relevant comparison to try to put the results that we're seeing in context.

So, for that, we actually – you know, there are a couple different ways we could have done this. One option would have been to say, let's compare this to the year before the pandemic hit and we can look at growth in that period of time and then compare it to growth over the pandemic year. Unfortunately, Pittsburgh actually just started offering the MAP test in the 2019/20 school year, so we didn't have a pre year within Pittsburgh that we could even use to try to say what does typical growth in Pittsburgh look like before the pandemic hit.

So, instead, what we decided to do was to use national norms to set expectations about what growth should be. Because NWEA is a test that's used by lots of districts nationally, they conduct what's called their norming study each year, and they determine what average growth should be, say from second grade to third grade and third grade to fourth grade. So, what we do is we use those national norms about what typical growth is for, say, a second- to third-grader or a third- to fourth-grader to put the growth in Pittsburgh in context and to say how did Pittsburgh's growth compare to that growth. And that's why when we say what we're seeing lags, it's relative to that comparison. So, I hope that helps to kind of set the broader context.

The second question, let me see if I remember this, Brian. The second question was about interactions, was that right?

[Brian Gill] Yes. Yes, sort of sub subgroups.

[Whitney Kozakowski] Yeah. Let me jump back to this slide here because I think this will be relevant. So, one thing that we did – so, we did look – obviously, we broke the results down and disaggregated by the groups that you're seeing here. We did also look for interactions among these different groups. So, for example, we didn't – couldn't do this with the IEP students because there are just not quite enough of them that when you start interacting that with, say, race or with gender, you're getting into really tiny samples very fast. But what we did do is look at this for the male-female split, the Black-white split, and the economically disadvantaged and not. We looked at the interactions of all of them to try to see is there a particular group within this set so, say, a student who was male and Black and economically disadvantaged that perhaps had a larger decline than some of the other groups.

Across the board, we actually didn't find evidence of any interactions. So, that's why we stuck with reporting it at this level. Really, the gender difference here was the main one that really stood out. We did the same thing for reading and, again, it was the same takeaway. Let me see if I – oh, this is reading. Yeah, we did the same thing for math, I should say, and it was exactly the same takeaway, that the real difference was through teen boys and girls, and we saw that difference consistently between, say, Black girls and Black boys and white girls and white boys, et cetera, when we looked at the interactions at a more granular level.

[Brian Gill] Great. Thank you, Whitney. Another one for you. So, one person in our audience has noticed that we don't have test score results for high school grades and, in fact, at least some of those students were taking NWEA MAP assessments. So, could you say a little bit about why we don't have – we've only got the grade level – the course grade outcomes but not the NWEA outcomes for high-schoolers included here?

[Whitney Kozakowski] Yeah. That's a great question. So, as you probably have noticed, in Pittsburgh, they offer the NWEA MAP exam actually starting in kindergarten all the way up to grade 12. So, before starting this analysis, we looked at test-taking rates in every test administration from the fall 2019 period through the last date that we had, which was winter of 2020/21. The goal with doing this is we wanted to first see which grades there seems to be a representative number of students actually taking the test.

It was through that process that it became pretty clear that we needed to stick with students in grades two through eight, and that's because the ninth through 12th-graders and the K through 1st-graders, the test-taking rates were just not very high. So, we were getting something like 60 percent, maybe 70 percent for some of them. So, because you're not getting kind of a representative sample, it's difficult to then look at the growth for those students and then try to make representative statements about the grade. So, we didn't do that. And then K and one didn't take it remotely. They weren't offered the test in the 2020/21 school year. So, for them, we couldn't include them obviously because we didn't have – nearly didn't have enough students to have included them, either.

[Brian Gill] Okay. Thank you. Let's see. Here's one for Ted. We have someone asking, "How did Pittsburgh Public Schools track absenteeism during this period when all the instruction was remote? Was that related to the Schoology logins and, if so, is it possible that kids were counted as absent just because they were having technology problems?" What did absenteeism mean during remote learning?

[Ted Dwyer] It's a good question. So, absenteeism and attendance was actually tracked by our teachers. During a synchronous educational day, students would log in to Schoology and, through Schoology, into a Teams chat. If the students weren't there, teachers were able to call them in through Teams. So, they could actually access them if they were – so that the child could actually access the Teams meeting through the Teams app. And if the student was participating that way, the teacher actually maintained the attendance record in our student information system. Additionally, the students could also call into the Teams chat if they couldn't get into the online systems.

And one of the things that we did at the beginning of the pandemic is actually secure hotspots and Internet access for a lot of our families who didn't have it. So, we were very purposeful to make sure that students actually had Internet access. I'm sure it wasn't 100 percent, but that's one of the things that we worked on really hard with both the community and internally to make sure students both had devices and access.

So, there is a possibility that students may not have been correctly identified, but we did have several processes in place, and that included providing teachers the ability to see both logins in Schoology, in our Clever system, which is actually how a lot of students would get to Schoology but it's also an online system that allows them to get to resources. And we also had a process set up so that they could use the Microsoft Teams program to communicate with students. Additionally, there's a texting app that, if a student couldn't log in or if there was an issue, the teachers actually had access to be able to text the parents of the students and provide them the link or the telephone number for the student to call in.

[Brian Gill] Okay. Thank you, Ted. Another question for Whitney, I think. We have someone who noticed that the course failure rate was highest in ninth grade and wonders, were the course failures concentrated in any particular subject area, for example, Algebra I, which, of course, lots of research indicates is a stumbling block for many kids? Were we able to take a look at any differences by subject matter?

[Whitney Kozakowski] Yeah, that's a great question. It's something that, you know, if we had had more resources and time, we definitely wanted to look into, but I don't – we did not have a chance to do that. It certainly was at the top of the list, though, if we had had additional time. We just know – so, we haven't broken it down by subject, so it would be nice to see, is it concentrated in electives, is that where the increase in failure has been, or is it really in these core courses that we know are very important for students to be able to move on to the next subject and are related to graduation requirements, for example? But, yeah, we appreciate the question.

[Brian Gill] Ted, I'm wondering if that is something that you've looked at at all in the data in Pittsburgh, about subject matter differences and course failures?

[Ted Dwyer] We've done it historically. We haven't actually completed that for this data.

[Brian Gill] Okay. Okay. We have an interesting question here asking about the differences between the test score outcomes and the course grade passing outcomes. I guess this is perhaps most notable in the middle school grades. So, in middle school grades, we were showing the kids are learning, according to MAP, but those failure rates are still increasing. Whitney, do we have any insight on that? Or Ted, do you, from your knowledge of what was happening in Pittsburgh, have any insight on what might explain that?

[Whitney Kozakowski] I'll let Ted start. If you don't, I have ideas I'm happy to offer. I didn't know if you have a smoking gun, Ted.

[Ted Dwyer] No, we're in a school. But could you repeat the question so I'm sure that I'm answering it?

[Brian Gill] So, you know, we see outcomes on the map, at least in terms of absolute learning, suggesting that there's growth. In the middle school, there's some indication that, even relative to prior norms, there wasn't substantial decline, and yet, at the same time we see that, the same middle school kids are failing. So, what – do we have any hypotheses about why kids might be failing more if their standardized test results are not going down substantially?

[Ted Dwyer] So, no and one of the things that several board members have asked are related to grades and how students are performing. One of the things that our superintendent has asked for and that we're working on is establishing a process to provide information on a link between grades and standardized test scores. We didn't have a national norm assessment for and it's difficult to compare across the state assessment because it's not linked across years or across grade levels. So, that's one of the spaces that we're trying to move towards is providing an analysis that can be used to have conversations with teachers and with principles about the distribution of grades versus how students are doing if we hold students consistent to how they're doing on an assessment.

So, if the students are proficient on the assessment, what's the distribution of students who are proficient on the assessment in that same area of MAP and what's the differences – are there differences internally, too? So, more of an analysis of our grading processes. Because grades are within the control of teachers, we need to be very careful about what we say about what those grades mean because they are specific to what the teacher is doing in the course. So, we need to be in a space where we're talking about what's the difference across so that can be a conversation, not a "You're doing this right or you're doing this wrong," because it really does depend on what the teachers are doing and whether or not we have consistency in grading practices across the district and across the content area. So, it's a –

[Brian Gill] Yeah, you know, it occurs to me that one possible explanation for this is that, you know, a failing grade doesn't necessarily mean a student hasn't learned anything; right? That a failing course grade is going to be related, to a great extent, to whether they've completed their assignments, whether they've shown up for things; right? And you could certainly imagine that students – this could well be related to our data on chronic absenteeism and participation in the online learning system – that students who are failing to show up frequently but still manage to show up on the day the assessment happens, and show up for half or two-thirds of the days of instruction are learning something, but they're not perhaps turning

in enough assignments to have those teachers give them passing grades on their assessments. I don't know, Whitney, if you may have thoughts about this as well.

[Whitney Kozakowski] Yeah, I agree with that, Brian. I think the other option here is that, you know, it could be that students are putting less effort in some courses but not in others so that they're more likely to fail, say, not their math class but their French class or another class, but they're still showing up for math and we don't want to actually see absence rates by source. So, it's possible that it's a subject-specific thing.

I've heard anecdotal reports from other people from districts that have said, "Oh, well, the students are allowed to fail up to one course in a semester," and I don't know if PPS had this kind of policy, Ted, that's why I was kind of deferring to you on this. But I've heard from other districts where they had said, like, we're kind of loosening some of our policies because we're going to allow students to – specifically it was that "We're going to allow them to drop a grade in their GPA. So, the students just stopped trying in one course because they knew it wasn't going to carry forward. So, Ted, I don't know if Pittsburgh had any policies like that in place, but, yeah.

[Ted Dwyer] No.

[Whitney Kozakowski] Maybe it was just differences across subjects then and how students put effort in.

[Ted Dwyer] We also talked to the students at the end of the year as well. There were several students that said that there were – they felt like there were higher expectations from their teachers on them as well. So, we haven't been able to really explore that, but they were – during those conversations, that's one of the consistent refrains from students is that teachers were asking them to do more in the online environment than they had in the in-person environment. I don't know if that is accurately representative because the students moving into a grade level where they didn't – that they hadn't been in before. So, it may have been that the teachers had the exact same expectations but it was a new environment for the students.

[Brian Gill] Yeah. So, lots of interesting possibilities there. No definitive answers to that question, unfortunately, but it's a very good one and, with more time, maybe we'd take a look at that. We got a nice question here about someone asking about recommendations for analyzing data like this generally. You know, based on doing this work, and I think this could be for Whitney or Ted, but what would be some recommendations for people who need to analyze their data in other districts or other states? This person says, "We're always trying to think about how we can analyze data in a more equitable way by helping identify how to best provide support to underserved students." So, what have we learned here that we'd recommend other places take a look at in their data?

[Ted Dwyer] I would recommend asking. There's Ask a REL. There are local colleges and other districts. And if you're involved in any kind of consortium or anything like that, put your questions out to the people in the consortium and the organizations in the consortium. But, like I said earlier, the REL has been an extremely useful partner in answering those kinds of questions and really thinking through those processes.

[Brian Gill] Thank you, Ted.

[Whitney Kozakowski] Thank you, Ted. I have a lot of thoughts on this. I think test score data, analyzing test score data is a minefield this year. Because of all the different changes that have happened, there's a lot that you have to keep in mind. The biggest problem is that the rates at which students have taken the tests are typically lower and it really presents a challenge if you're going to, say, compare third-graders from a prior year to third-grade performance this year. You're getting a very different set of third-graders taking the test each year. So, these cross-sectional comparisons like that are really kind of potentially misleading.

So, I mean, a couple recommendations. One, check the missingness rates. Two, we ended up doing a longitudinal analysis where we looked at calculating growth for individual students. That's a lot safer because you're comparing within student what happened to them during the pandemic versus taking a group of students where the group who takes the test has changed. So, there's a couple things like that. I mean, there's – yeah, I think the missingness rates and doing a full missingness analysis is really important.

[Brian Gill] Great. Thank you, Whitney. All right. We're in our last two minutes here, so I'm going to go to one last question, which, unfortunately, I can't claim to have a great answer to. This is a question coming from an old friend in Pittsburgh who wants us to get inside the black box, so to speak, and wonders, you know, what was actually going on in schools and classrooms that might explain this. Moreover, what can we learn about how to do things better, you know, which might involve trying to identify particular schools or particular districts to expand the scope that may have been especially well prepared, and did they end up doing better in terms of having results that didn't show these kind of learning lags or increases in course failures.

Again, from everything we've seen, this does not look at all like a Pittsburgh-specific story. These results really are consistent with what's been seen across the country. But, nonetheless, I'm sure it is the case that there are places – some places were better prepared than others. And I think there's still a lot more – probably a lot more research that could be done to try to understand the variation in results across different schools and districts in terms of the kinds of supports that were provided and how those related to student outcomes. I wish we would have been able to do that here, but that would be a much more ambitious research project.

So, we are coming up on the hour. I want to thank everybody for participating today. Thank you to our panelists. We are all reachable pretty easily if you've got follow-up questions for us. I'm sorry we weren't quite able to get to all of them but we got to most of the questions that came in.

My final request for you is if you could take just a minute to complete a very short survey. I think it's got two questions. It shouldn't take you more than about 60 seconds to give us some feedback on the webinar so that it can be shared with our funders at the U.S. Department of Education. There's instructions there.

The teal button in the webinar window tells you how to do that. Thanks again to everyone. Really appreciate your participation and hope to see you again.