

BILL EAST: Hello. I'm Bill East, Executive Director of the National Association of State Directors of Special Education, or NASDSE. I want to you welcome you to NASDSE's 2010-2011 Professional Development Series. Speaking for Bambi Lockman, NASDSE's president, and the board of directors, we want to thank the many states that are participating in our series this year.

I also want to thank Pennsylvania's State Director of Special Education, John Tommasini, and James Palmiero, Director of the Pennsylvania Training and Technical Assistance Network, for making it possible for NASDSE to bring this conference to you from WQED in Pittsburgh.

The coordinator for NASDSE's Professional Development Series this year is Christine Cashman. If you have questions about NASDSE or NASDSE conferences or any of the services and programs that NASDSE has, I suggest that you check our website at www.nasdse.org. That's www.nasdse.org.

The title of today's conference is Special Education Teacher Evaluation: Issues and Answers. The evaluation of teachers, including special education teachers, has become increasingly important, as teachers are held accountable for student performance.

Many special education teachers are assuming new roles, such as co-teaching, coaching, and supporting general education teachers in many ways. This conference will address issues of teacher evaluation with specific references to special education.

We know there are many approaches to dealing with teacher evaluation in states and local districts across the country. We hope this conference will be helpful to you as you deal with the issues in your location. I am very pleased at this time to introduce our presenters for today. To my immediate left is Lynn Holdheide. Lynn is the special education research associate at the National Comprehensive Center for Teacher Quality. Now most people just call that the TQ Center, Teacher Quality Center. And it's located at Vanderbilt University. Her work there focuses on providing a TQ Connection newsletter.

She also is known for her work around response to intervention, inclusive practices, and effective teaching practices and strategies. Before going to Vanderbilt, she was a consultant at the Indiana Department of Education. She has experience as a teacher, a transition specialist, a vocational specialist, and a residential provider. Lynn holds a master's degree from Eastern Illinois University. Lynn, welcome to our conference.

LYNN HOLDHEIDE: Thank you.

BILL EAST: To Lynn's left is Tony Bagshaw. Tony is the Managing Director of Human Capital for Battelle for Kids. He leads Battelle for Kids work in strategic compensation and other human capital innovations.

Prior to that role at Battelle for Kids, he was Director of Knowledge Management, which is a key thing that we're dealing with at this time, with all the knowledge we have to manage across our program.

Prior to that time, he has over 20 years of experience in education as a teacher, coach, administrator at the elementary, junior high, and senior high levels, and assistant superintendent of a school system in Ohio. Tony has a master's degree in secondary mathematics from Indiana University Southeast. Tony, welcome.

TONY BAGSHAW: Thank you.

BILL EAST: I'm going to turn the program over now to our presenters. We will be back with you for a ten-minute break at some point during our conference. And I look forward to listening to our presenters, to hear what they say about this very important topic of teacher evaluation. We're going to go to Tony first. Tony?

TONY BAGSHAW: Thanks, Bill. We really have two goals today.

We're hoping to paint for you a national perspective on where teacher evaluation is, some best practices, some things that are going on around the country that you might find of interest. And then I think we're going to take a pretty strong position on the fact that special education teachers and administrators need to be involved in those conversations, and need to insert themselves into the conversation around evaluation, because of the impact that it's having on the profession.

What you see before you is really a slide that comes from a full-scale literature review that we did at Battelle for Kids when we were helping a 23-district consortium in Tennessee redesign teacher evaluation across that consortium. And I'm going to point out some things I think you might find of interest, and sort of, you know, how have we arrived at this point.

You'll see on the slide some things that really give you just a bit of history. And I want to point out the Coleman Study, that is not news to anybody in education, that really was pretty unflattering in terms of the influence of schools on student performance.

And you'll notice in 1980 that we really got our first study that really starts to ask some serious questions about teacher evaluation in general. Is it effective? Is it doing what it was designed to do?

You'll see a 100-district study a bit later. And then what I find a little disturbing, frankly, is that we replicated that 100-district study about 10 years later, and found really

very little progress in terms of teacher evaluation. Then we have the advent of some research coming through in the mid-'90s until now. And what that research is showing, and we have a slide we'll spend much more time on later in the presentation, it's showing that the number one influence on student achievement is teacher quality.

So we have a situation where we have teacher quality being brought to the forefront, and at the same time we have the realization that our evaluation systems are probably not very good and probably not doing what they were designed to do. So you really have a clash. And that is why you're seeing this issue really get so much attention across the country, and be the focus of many states and LEAs.

If you look at the first three papers on this slide, you may have read some of these. They certainly will all tell you the same thing. They're very, very well done, and they will tell you that teacher evaluation in America is not very good in general. I'll call your attention to the Gates Foundation sites. The Bill and Melinda Gates Foundation is putting approximately \$100 million into each one of those four sites, and the whole purpose of that work is to measure educator effectiveness. And so it would behoove all of us to keep tabs on that as that work progresses.

If you think about the policy that is before us now, in essence, you look at, you see an alignment between all of the policy at the federal level. You see the ARRA funds, which specifically spoke to reliable evaluation systems. And certainly those of you who have had any dealings with Race to the Top know that of the four assurances, the weighting of quality teachers and quality administrators was the highest point value.

Certainly Lynn and I don't make the decisions about where this goes at the federal level, but we are not hearing anything that would lead us to believe that those four assurances are not going to be a major part of the reauthorization of ESEA, and thus that focus on quality principals and quality teachers will continue to be a focus. If you look at value-added research, and we're going to talk about value-added in much more depth later in the program, and there are several studies that are referenced here.

These folks might argue a bit about the degree to which teacher quality impacts student growth rates. We might argue about whether it's 65% of the variance or whether it's 40% of the variance. But what they won't argue is that it is the most important factor. And the advent of value-added analysis has indeed brought a lot of this to the forefront, and we now have some really solid research that says teacher quality really matters.

When I speak to and meet with teacher groups, I usually tell them, I have good news and I have bad news. The good news is, you make most of the difference. The bad news is, you make most of the difference. And value-added analysis and good research in that area has shown that to us.

If we looked across the literature review, and we really looked at what were the findings, what do we see, what are the results of the fact that the systems that we currently have are really not very efficient?

One thing is that excellence goes unnoticed. And I think we miss a great opportunity as a profession to learn from those folks who are producing extraordinary gains. The other thing that . . . stays typical. And we don't help those folks move from the middle of the distribution to the top of the distribution. And obviously, chronically low performance goes unaddressed.

And at a time when resources are being increasingly taxed, and knowing that teacher efficacy is probably the most powerful tool in our toolbox, I'm not sure we can continue to have evaluation systems that don't address any of those. It's pretty clear from the research that teacher evaluation to date has had little or no impact on human resources decisions.

It has not played a major role in retention, it has not played a major role in promotion, placement. There are those with increased interest in the degree to which robust teacher evaluation systems can inform strategic compensation systems. It really just hasn't played a role. And so the result of that is, once again, the fact that our most powerful impacter is going unaddressed. You don't have to start with a blank piece of paper as you think about diving into this work.

There are professional evaluation standards out there. A quick search will yield some pretty fruitful results. And every professional organization has standards that they publish, that they believe are critical for success in that profession.

So I would encourage you to look there. There are many states that are taking on the issue of teacher evaluation. I have pointed out Tennessee, because it's the one I'm most familiar with. And they have laws on the books that were passed in January that are statutory, that mandate that, for instance, they mandate that teacher evaluation must play a role in human capital decisions.

I really invite you to give a tremendous amount of thought to this third bullet. Teacher evaluation, most would agree, really serves two purposes. One is it gives us a summative decision on how someone is performing. Those are decisions that impact retention, etc., etc. Others advocate that the primary role of teacher evaluation is formative in nature. It gives us information that can, so that we can help folks get better.

I think it's really important, if you delve into this process, and redesigning a teacher evaluation system, that you're really clear about what your purpose is. There are a lot of folks that believe that those are definitely at odds. For instance, one of the great debates is, can you put a building principal in a position where they are in charge of both the summative evaluation and the formative aspects?

It's sort of that notion that, can you really be a coach, and really say to someone, bare your soul, tell me what you're struggling with, let me help you, and then at the same time be the person that makes a summative evaluation of whether or not they work there the next year? So I just really encourage everyone to understand that that is really something you have to grapple with. And then there are a ton of SEA efforts and LEA efforts, and some of those we're going to highlight a bit later, that are ongoing.

And so you don't really have to start with a blank piece of paper. And I'm going to hand it over to Lynn.

LYNN HOLDHEIDE: Thanks, Tony. The need for accountability in evaluation systems, which Tony just described, is really no different for special education teachers, especially when we consider the fact that we have performance-based compensation systems really bubbling up across the nation. And it is really our job to figure out, our job being special educators, to figure out what are those special challenges to fit into those evaluation systems that are out there? And also to determine where we do fit, and how we need to differentiate, if we do need to differentiate, for those systems.

There are some particular challenges, or persistent challenges, that are really happening in special education, that really emphasize the need for accountability within special education teacher accountability. And one of those is, of course, the persistent achievement gap for students with disabilities. It's been there for a while, and it is also a main reason why many districts and schools are not meeting AYP.

We also have issues with teacher retention and recruitment in special education. We have many special education positions across the nation that are going left vacant or unfilled. And then ultimately, my passion is that we have a limited use of evidence-based practices that are happening in the classroom. We do have sufficient evidence about what works for students with disabilities, and we need to make sure that our teachers are using those practices in order to elicit improved results for students.

Some other challenge that Bill alluded to earlier is the challenge of how the special education teacher roles are changing. We have a more diverse classroom. We have increasing caseloads, and we're serving in many, many different capacities. We have teachers who are in resource rooms. We have teachers who are working in co-teaching situations, and those co-teaching situations vary across the nation, some of which, they're truly co-teaching, where they're both party to designing the lessons through assessment.

And then we have co-teaching where the teacher is maybe working with a group of students within the classroom.

Knowing all of this, one of our charges at the TQ Center is really to work with states and regions in helping them design teacher evaluation systems. And many of the questions that came to us, quite frequently, is how do we do this, and what do we need students, teachers, special education teachers?

So we set out to do this particular inquiry or study, and taking a look at what are those challenges in evaluating special education teachers. Our goal was not, was to identify those, but also to determine what was currently going on, identify some promising practices that are out there, and then ultimately to provide guidance and policy recommendations to districts and states that are really entering down this path of designing a teacher evaluation system that works for all teachers. To do this, we conducted a review of policy and literature.

We had a survey inquiry, and we interviewed a series of state and district level practitioners and researchers across the nation. This study was conducted between December and April of 2010.

We did do a survey. We had two surveys designed, one designed for state directors of special education, and one designed for local directors of special education. We worked with the Council for Exceptional Children to design the survey, and several other national experts. And our respondent pool was state directors of special education, as well as local directors. We did have a total of 1143 respondents in the survey.

We know that local control and teacher evaluation, really, historically was a local responsibility. But we know also that with Race to the Top requirements, states are becoming more and more involved. One of the questions we did ask our respondents, our group of special education directors was, how is your evaluation system designed currently?

And as you can see here on this graphic, 22% said they were using the state's design system, but a little under 50%, almost 55% of the respondents, indicated that they use their own locally-designed system for teacher evaluation.

One of the questions we wanted to ask then is, do you modify that system at all for special education teachers? And what we found is that 26% of our respondents indicated that they did indeed, to some degree, modify the evaluation system to accommodate special education teachers.

We also found out, for our local respondents, 81% of them indicated that they were not able to modify the teacher evaluation process at all, because of contractual agreements within their schools. So we know that there's a little bit of modification going on, but for the large part, most of the teachers are being evaluated the same, the special education teachers. What we did find in our review is that many of the evaluation systems indicated that they were measuring the teacher's ability to meet the needs of diverse learners.

But what concerned us a bit, was that we were wondering if we were missing some of those special skills, those skills that we know special education teachers play in the classroom, like writing an IEP, facilitating those meetings, collaborating with teachers, parents, the students themselves in helping them achieve their goals.

Secondary transition, and the social and behavioral interventions that we do and work with students on, as well as compliance with legal mandates, those are things that we know, just a few of the things that we know as a special educator, are really important, and a big part of our job.

And then we were also concerned about those, as I alluded to earlier, those specific evidence-based practices that we know are effective with our students. Are those included and looked at and evaluated, whether a teacher is using those, when they are looking at evaluating a teacher? When we designed this survey, we wanted to take the opportunity to get the collective opinions of these special education directors.

These are folks that have a lot of experience in supervising, mentoring, and leading this population of teachers. So we really wanted to get their opinions as to what they felt about teacher evaluation, and how those should go. It was on a Likert scale. The directors, the respondents were given a statement, and then we asked them to indicate if they, went from strongly disagree to strongly agree with the statements.

And what you see here, starting from the bottom, is that 84% of our respondents, or experts, if you will, indicated that special educators are required to have knowledge, skills, and expertise that general education teachers are not. We also learned that 32% of them felt like, using the same evaluation process, it should be the same, using the same evaluation process for general education teachers as well as special education teachers. And the last bullet, I'm going to wait until a little bit later to talk about. And I'm going to go ahead and turn it over to Tony.

TONY BAGSHAW: Thanks, Lynn. As soon as you start talking about teacher evaluation, and certainly the trend that measurement is becoming a bigger part of teacher evaluation, you start talking, you get into talking about what are we going to measure? And really that is, to some degree, a values conversation. In districts that highly value collaboration, that's going to receive more weight. If you value statistics about raising test scores, that's going to receive more weight.

So the point is really that most of these systems occur in districts that have their own DNA, and the district really has to have a significant conversation around what they value, before they embark on a redesign of a teacher evaluation system. We have some technological advances that are really, really helping, in terms of what we can measure.

One really simple example is smartphones, and the software you can put on smartphones, and the fact that you can put your walkthrough software on your smartphone. You can breeze through a classroom in ten minutes with dropdown screens, and really start to gather a lot of really good information. We, when we look at the data and the evidence, we have to be able to measure it in an effective way.

Oftentimes, when you start having these conversations with folks, they will quickly approach things that, believe me, I think are important, but they're virtually impossible to measure. And if the whole premise is that evaluation is going to become more evidence-based, we are, to some degree, certainly at this point, limited by those things where we can actually acquire the information. And we have staff, we have money, we have resources, we have policy, we have things that have always been at our disposal to do teacher evaluation well.

Yet the evidence is pretty clear we haven't done it so well. So one of the big questions that I have is, as a, for school districts, for states, are you truly, truly willing to put the resources into this that it's going to require to do this well? Because I do not believe that we can shoehorn a much more robust evaluation system into the existing resources. So I think that's a fair and legitimate conversation that has to be had.

What are some of the challenges when we think about effectiveness data, particularly things like value-added? What are some of the concerns that come across? One of those is that if you look at, in essence, value-added data, and overweight that, then you are devaluing some things that are really important in producing quality kids.

And I couldn't agree more. I'm a very, very big supporter of the use of value-added. But having said that, there are other things that really matter. Everything that counts can't be counted. And we're going to talk about some innovative practices, and some really promising things that are going on a bit later.

One of the things about the use of things like value-added that you need to understand is that value-added is not treatment. It is world-class diagnosis. It can tell you which programs are being effective. It can tell you which particular groups of kids you're getting gains with and which ones you're not getting gains with. But it is not going to tell you that you are not using formative instructional practices well, therefore, this is the professional development that you need.

So I think we have to be fair and reasonable about what the limitations are. And then the areas, the other 69%, the places where we can't get that value-added data, those are real challenge areas for everyone across the country.

Folks are really just struggling with, what do I do with K through 2 in art, and we're going to talk about some of those things as we move forward.

This really is a slide about, based on our analysis of what's going on in measurement across the country, where we believe it is going. I do believe

value-added, with all of its warts, and all of its limitations, and all of its challenges, and certainly there's no shortage of that debate in the blogosphere or out on the Net today, I do still believe it's going to be a core component of things moving forward. I believe it's going to be a core component of accountability right down to the teacher level. One of the things that I'm really excited about is this notion of softening measurement error through multiple measures. We have a lot of folks talking about multiple measures.

One of the challenges we have right now is, we don't have great other measures. There are some promising things going on, and I think, once again, the Gates research around the MET project, it shows real promise. We'll reference some later. But the whole idea here is that you don't want to rely on a single measure. And by bringing in multiple measures, you do soften that measurement error.

The third bullet we've discussed, and I think that it's very, very clear. Measurement, things like value-added are increasingly becoming part of human capital decisions. I think one of the things that's really exciting is, if you think about, if you fast forward a bit, and let's say, for the sake of argument, that you have developed a much better teacher evaluation system, a much better principal evaluation system. We have an opportunity now to take that information, for instance, how a particular principal might score teachers on a more robust rubric, and we have the opportunity to correlate that with other data.

And hopefully those agree to a great degree. But if they don't, then it really gives us a professional development opportunity to discuss the fact that this principal may be consistently, their evaluation scoring is consistently not correlating with the more objective data that we've got. And it gives us a chance to intervene with that person, and hopefully help them get better at that process. And once again, we're going to talk about some promising practices as we move along through the presentation.

LYNN HOLDHEIDE: Thanks, Tony. One of the questions I typically ask when I go out and do training is, if you were to observe an effective teacher, what things would you see or what things would you hear? And after doing that, would those things be different for special education teachers? Would those items that you list be different?

And one of the questions that we continually get asked is, should the evaluation system be the same for all teachers, or should it be differentiated for teachers of students with disabilities?

And I just listed a few pros and cons here. A pro of keeping it the same system is that the measurements and the dimensions are identical, making it more simplistic to implement. A con for that is what Tony alluded to earlier. It could devalue some of the work that we're required to do as a special education teacher, like the IEP facilitation,

the paperwork, the legal requirements, the collaboration with other teachers, all the things that we know are very important.

Also, we lose a focus on the social and behavioral outcomes for many of our students, that we know we work hard to work with those students to succeed. So those are things that, really, if we are designing, or we're being evaluated with the same system, may be a con for that, because it could devalue a lot of the work that we do. Having the same system, obviously is more simplistic, as Tony alluded to earlier.

Designing a comprehensive evaluation system that not only measures teacher effectiveness, but also helps the teacher grow in terms of professional development, is difficult, and a complex thing to do. And so the more that we try to diversify it for different groups of teachers, the more complex it's going to get. So it would be more simplistic.

But, again, we're back down to, does it devalue those roles and responsibilities? And as I alluded to earlier, what about those evidence-based practices that we know can work with our students? Do we want to make sure that we are enforcing those, or trying to encourage teachers to use those? We talked, Tony mentioned earlier, what do we value? What is it that we value, that we want our teachers to do? And is it going to be those evidence-based practices?

Fairness. Fairness is very important. Teachers need to buy into whatever system is developed. They need to know that it's fair, and that they, it is measuring their effectiveness just the same as anyone else. So the same system leads to more fairness.

But in that, it also becomes difficult, because what happens if you are a co-teacher and you are trying to work in that classroom side by side, working with the teacher in designing the lessons, assessing the students, working with all students in the classroom? That's great. You would feel pretty fair about being evaluated against that.

But what if you're one of those teachers that's really trying to get into that classroom as a special educator and help, but you're really relegated to helping just a couple of students? Or maybe you're relegated to copying papers? I wouldn't feel very comfortable about being evaluated that way, because I know I'm a better teacher than that. So that's some of the issues with the fairness.

And then, of course, the reliability. If we're using observation instruments of whatever nature, if the evaluator is not familiar with, or doesn't have the content expertise to evaluate a special educator, you start to question the legitimacy of the evaluation in that regard. So there may be some need to be explicit, and have explicit criteria for teachers of students with disabilities.

Many states are dealing with this. Do we do the same system? Do we try to differentiate it in any way? And really, it's been quite a struggle for many states and districts, trying to kind walk through this. We have a need to not want to separate ourselves as special educators, but at the same hand, we have this need to really start to hang on with the things that we know that we do.

And it starts to make you question a little bit, why do we certify and train special education teachers differently, if we're going to evaluate them in the same way? What some states and districts are doing is assembling a focus group of special educators, to really dive into these issues. One of them is to take a look at the current system that they have in place and what they're recommending, and assessing whether special education teacher effectiveness can be determined with that system, or is there a need to differentiate in one way, shape, or form.

They also pull a group together to identify what those current challenges are. Many of them are really sitting down and focusing on, what are the skills, those tasks that they're required to do as a special education teacher, that should be included or explicitly stated within the evaluation system.

What I see a lot of states and districts really struggling with right now is coming up with other measures. If they're not included in value-added teacher scores, what other measures are they going to use to determine whether or not the students have grown academically? And so what these states are doing is assembling this group of special educators, and saying, okay, what measures can we use, that are reliable and valid, that can show student growth, that can be a factor in teacher evaluation?

And then obviously, once the system's up and running, continually being involved in assessing the effectiveness of the measure. We might say, this is what we're going to start with, this is what we want to start to evaluate. And we might see there are some skills and some strategies that we hadn't put in there that we're not picking up on in our evaluation we want to reinforce. So continually being involved in that assessment of the process is a really critical and important key.

One of the things that we did ask, I mentioned to you before I would come back to this bullet, if you look at the very top, when we asked our experts, if you will, over a thousand individuals who really work a lot in special education, 92% of them indicated that they would like to see the special educators use of evidence-based practices as a component in the evaluation process. Remember I alluded earlier, there's research out there showing that our teachers aren't using those evidence-based practices.

And what we're hearing, I think, pretty loud and strong from the field is that we need those to be part of the instruction that our teachers are providing. And if that's the case, then should those be, those evidence-based practices be a component of the teacher evaluation process? And I will go ahead and turn it back over to Tony.

TONY BAGSHAW: Thank you. As we've thought about the measurement issue, and really looked across the country, and tried to look at everything that was going on, I

think we've arrived at a pretty balanced approach to where this is going to go, and what seems to be fair.

The first thing is multiple data sources, and I talked about that earlier, and multiple measures. The second one is across time. And as I get into value-added a bit later, we really seem to have two pretty extreme camps in value-added. One of those is a camp that says that it is the do-all end-all, it tells us everything we need to know, we should be using it in high-stakes personnel decisions, with one of your snapshots, and frankly, I don't agree with that camp.

When we talk about looking across time, we'll talk later about the fact that, if you look across three years, that's a much more fair and reasonable way to look at performance.

We talk about linking teachers and instruction to students. We're going to spend a lot of time on that lately, or later, excuse me. This idea, that the data that we have in our current education system really supports some of the uses, it just, our research says that's simply not true. I really do believe we are heaping really robust data uses on top of a system that really wasn't designed for that.

Frankly, the current data system that we have was designed to count kids, and take a snapshot, so that it could inform funding. It was not these more robust uses. It's really not what some of these systems were designed to do. I do believe that data should inform instruction.

And I'm not really willing to go back to the days where we didn't look at the data, we didn't look at the information, and we really just proceeded based on opinion. And the final piece is, I believe this has to be about improvement, not about judgment. And there certainly is a degree, there is going to be a degree of judgment in these processes. Once you start measuring things, and once you start embedding them in someone's formal evaluation, there is clearly a degree of judgment in that.

But I'm back to the summative versus formative aspects of things. And I think if you make, whatever you do in these systems, if you make it more about judgment than improvement, you really miss the point. We believe every child can learn and grow and improve. And I think we owe the adults in the system that same consideration as we design these systems.

Here are some things that Lynn and I think are worth looking at in terms of some sort of cutting-edge things that are going on across the country. Certainly if I were an LEA or an SEA and, you know, tomorrow we sat down in a meeting and said, we're going to redesign our evaluation system, either at the state level or at the local level,

these are some places where I would start, so that you don't have to start with a blank piece of paper.

The TEAC committee in Tennessee, the Teacher Evaluation Advisory Committee, was mandated by legislation. It's broad stakeholder involvement. The legislation in Tennessee mandates that where available, value-added data must be used for a minimum of 35% of evaluation, with 15% coming from other measures. These folks have done a lot of really good work.

Their minutes are public record, so you can go to the Tennessee Department of Ed. website and find the minutes of the TEAC committee. And I would encourage anybody looking at this on a statewide level or a local level to go read those. And you can read some of the things they're being successful with, and some of the challenges that they're having.

Really on this slide we put, I wanted them all in one place as a reference for you. I'm not going to speak right now to Austin and D.C. impact, because Lynn's going to cover both of those in depth at a later point. Harrison School District 2 in Colorado, I find this one really intriguing. They, this is a very innovative practice. Here's their solution to getting at the idea of, how does one measure growth in, let's say, art?

What they're doing is each art teacher in the district is required at the beginning of the year to take a sample of artwork from a particular student. Then at the end of the year, they're required to take a sample of artwork, representative sample samples, from those same students.

Those pre and post, I guess, is probably a good way to refer to them, those are brought into the gymnasium. No one knows which artwork is which. My understanding is, teacher does not score their own students, but then they, it's arena scoring. And then they score the growth, and rank those as a team.

Now I think that holds some promise. Do we know it's going to work? Do we know it's going to be valid? Do we know it's going to be reliable? No, we don't. And I think anyone that embarks on this endeavor needs to understand that there is not a clear-cut solution to many of these problems. We are clearly, I think, in the infancy of doing this well.

Another interesting piece that's going on is Hillsborough County Public Schools in Florida.

Hillsborough is one of the Gates' MET sites that I referenced earlier. I think what Hillsborough is doing, which is particularly interesting, is that they are literally developing/have developed pretests and posttests for every subject in the school, every subject. Career Tech, Latin III, you name it. So once again, do we know that that's going to be a valid and reliable way to measure this? They're certainly making every effort and have a lot of assistance, and I think there's some promise there.

I'll be very curious to see what comes out of both Harrison and Hillsborough in terms of the research. I'm sure as the years pass they will refine those more, and be able to center some of those efforts to get a better sense of what true growth is in a year. I referenced the Gallup student polling, and Gallup's is not the only organization that is attempting to create measures in this space.

But this one I find particularly intriguing. I've always believed that one of the most powerful things that teachers do, and wonderful teachers, great teachers over the years, because if you talk to any successful person and asked them about why they are where they are, most of them will reference a teacher.

So, and there was a special connection there. And I believe fervently that great teachers have that ability. They have the ability to connect with kids. And those kids will go to great lengths for them, because they know they're cared for. I certainly, I have three children of my own, and I won't bore you with all the stories, but I have personal evidence that that is the case as well. The Gallup student poll is designed to measure three things.

It is designed to measure hope, engagement, and wellbeing in students. You can give it to students in grades 5 through 12. It takes them about eight and a half minutes to take it online, and here's the really great news. It's free at the building level. I think there's a lot of promise there. I'm just really curious to see as we, once again, one of those other measures, what is that going to tell us? How is that going to correlate with value-added gains? Which one's going to tell us more?

So I would highly encourage you, it's probably worth your time to jump online and at least go look at that and see if you think it makes any sense for your school district. Those are some things that Lynn and I talked about that we thought were some promising practices, that, and once again, she's going to cover Austin and D.C. impact in more depth later. All right. Value-added. I've referenced it multiple times.

It's not a huge secret that Battelle for Kids has been engaged in value-added work for ten years, so full disclosure. Probably most people wouldn't see us as completely unbiased. We like to think we're at least balanced in discussing value-added. What is it?

Really, value-added is a statistical measure that is designed to measure growth of students versus some standard. This is, we're not going to spend time today to talk about different ways of modeling value-added. But that's what it is designed to do. It is, in most cases, it is typically based on standardized test scores. That's certainly one of the criticisms. One of the things that bring value-added its reliability and validity is the fact that it is based on those standardized test scores. One of the criticisms is, it's based on standardized test scores.

It's, you, in essence, what you are doing, if, probably the best parallel that you can think of is life insurance.

You are, if I go and ask, if I attempt to purchase \$500,000 in term life insurance tomorrow, I'm going to get a particular rate. That rate is going to really be based on my health habits, my age, my gender, all of those things, okay. So you're, in essence, using what you know about me personally to make a prediction, frankly, about when I'm going to die, one way or the other. And that's what's going to, you know, that's what's going to drive your premiums.

With students, you're doing the same thing in a value-added analysis. You are taking their prior testing history, and you know how students with that same testing history have scored over the years, so you're able to make a prediction about how that student should score. And then based on how that student does score, you can say there was value added, or we came up a little short of what would be predicted.

That's really, that's a gross oversimplification of value-added modeling. But at its core, that's really what you're attempting to do.

I want to be very clear. There are multiple providers of value-added data. There is tremendous debate in the space about which is the best way to model. You, one, probably the most raging debate is, do you have to include socioeconomic factors in the modeling of the value-added? There are certainly different camps on that issue.

Some models are extremely robust. They are very, very complex. Others are relatively straightforward. I will tell you that as, when we get into the value-added section in more detail, there are tradeoffs, depending on which way you want to go. But here's the important part to me, and the reason that I can tell you exactly where I was standing the first time I had it explained to me.

Value-added is considered a productivity measure. And value-added is correlated with what we do in the system instead of what comes to us from outside the system. And that's why many see it as a very, very, it's a true sea change in terms of measurement in education. And I'll actually show you some slides later that will, I think, make you feel a little bit better about the fact that that may indeed be true.

Do I believe personally that value-added is the way to go? And I would say yes, I do believe that. But you'll also see that it says, within reason. I referenced this earlier. I do not believe in snap judgments using value-added analysis. You drive your repeatability up to about .8 when you start looking at 3-year averages.

I think you really start down a very, very dangerous path if you're going to make large-scale human capital decisions based on one year's worth of data.

We've talked about balancing this with other measures. This is one of my favorite statistical quotes. All statistical models are wrong, but some are very useful. And I really believe value-added is one of those. Nothing is perfect. There are

challenges. And those of you who are statistically inclined probably either know some of those challenges, or once you start studying it, you will quickly get into some of those challenges.

But having said that, I still believe it is a tremendous tool for school improvement and have plenty of practical evidence that supports that claim. You have to be careful about your data. Education data is notoriously noisy. Once again, the data systems that we have are improving rapidly.

We have to credit the federal government for supporting that. I think they're getting better all the time, but we are using them. Value-added analysis is a very, very robust purpose. And if you're building that robust purpose on top of sort of a house of cards with data that isn't very good to start with, I think you run the risk of making judgments that simply aren't true. And you're just, you're not being fair to folks.

Anyone who monitors statistics always knows, the bigger the sample size, the better. So you need to be a smart consumer of this information, and understand that sample size truly matters. I always tell folks, I have more confidence in value-added data at the district level than at the building level than at the teacher level. I have more confidence in a teacher-level value-added report that was created for a middle school teacher that has 150 students in a day versus a self-contained teacher that has 25.

It is simply based on sample size. And when we get into special ed, we'll talk about that in great detail. There are simple models, and there are really complex models. And you're going to make a tradeoff if you decide to use this data for school-improvement purposes/human capital purposes. And you cannot get around that tradeoff.

If you want a simple model that is really easy to explain to folks and really doesn't take a really robust communications effort, those are certainly out there, and they have utility. But understand what you trade off is probably some of the robustness of the model. If you're using a more robust model, I mean, these models take entire weekends of supercomputers to crunch on a statewide basis, you get more robust data.

But the tradeoff that you make is it's exceptionally difficult to explain to folks. And this is a quote from Rob Meyer, from the University of Wisconsin, it's one of my other favorites. It's, you know, simpler is better, until it's wrong. And so there are certainly tradeoffs in choosing a value-added provider as you look at those things, to make you aware of those.

This is probably my all-time favorite value-added slide. It really is the problem that value-added analysis is attempting to solve. If you look at the green bar, that's proficiency. That's proficiency in your state. And if you use pure achievement data to make judgments about the performance of districts and buildings, in essence what you're doing, is you are rewarding the teacher, the building, the district, for Student A, because in a, from a pure achievement standpoint, that student was above the proficiency bar every time.

So in third grade, fourth grade, fifth grade, sixth grade, seventh grade, eighth grade, passed, passed, passed, passed, passed, passed, passed. You get rewarded for that. I don't think any of us want that to be our child. If you look at Student B, that student never passed. So in a pure achievement world, that would have been fail, fail, fail, fail, fail, fail, fail. That's not the case in a value-added analysis. You would get clear credit for Student B, and you would not get credit for Student A.

I referenced earlier, was going to show you a slide that I thought you would, would make you feel a bit better about this whole notion of whether or not value-added analysis can indeed be that productivity measure, i.e., is it fair? Is it fair across different typologies of districts and buildings?

What you have before you is a slide from a state, which shall remain nameless, from a year and subject, which shall remain nameless, which really outlines the value-added gains of school districts in that state.

This is a large state, by the way, that has about, almost two million kids in it. And I've always found this slide to be pretty fascinating. This, these typologies are the way the state classifies these districts. And if you look across there, what you will see, I think you'll see a couple of slides that probably would shock you to some degree.

If you'll look at the third slide, or, excuse me, not slide, but column. If you'll look at the third column and the fifth column, in essence, what that is saying, is that in this subject in this year, that the suburban districts in this state produced virtual identical gains in students as the 21 largest urban districts in that state. Now clearly, there's not a state, if you looked at pure achievement data, that that would be the case.

And so that is why there are folks who advocate that value-added is, to some degree, a great equalizer in terms of how we're examining performance. Just to give you a quick look at what people get, I think here's one of the powerful things about value-added, is while the, it's a really complex analysis, the results you get are pretty straightforward. Green is good, yellow is what you expected, red is not so good.

And you can see by this slide, you can see as a superintendent, that would be a very, very useful slide for you to have, in terms of, once again, really robust diagnosis, and how you're going to allocate your resources. I think this is probably the analysis that's most often used for school-improvement purposes. What you're looking at is, you're looking at bar graphs that represent the students being grouped into quintiles based on what their predicted scores were.

So you'll see that the students with the lowest predicted scores are that first quintile, Quintile one, middle, the third quintile, top, okay, once again, based on their prior testing histories. And you can see that this is a pretty useful tool for school improvement purposes.

We can see that we're pretty clearly getting more gains with kids in the lower quintiles than we are in the higher quintiles. Now having done a bunch of this work, typically what we see is that maybe a curricular stretch issue, if you saw this at the district level. But you can see the utility of the information for school-improvement purposes.

Here's an individual teacher report, and this has always been one of my favorites. And if we were live, we would break into groups and sort of discuss this. This is a report, this is a value-added report for an individual teacher. What you are seeing now instead of the students being broken into five groups, they're broken into three groups.

That's once again based on sample size. So that first bar on the left is, how did the students grow that in, that had the lowest third of predictions. And you have the middle, and then you have the high. And so one of my favorite questions to ask is, if this were one of your teachers, and you were the building principal, would you consider this information actionable? And it's always a great debate, because you can actually argue it either way.

The no argument is that all of those bars are really close to what you would have predicted for those students, which is that zero bar that's right in the middle, the horizontal bar. So I could argue that the answer is no, it's not actionable. The other argument is that you are getting slightly better gains with the high kids than with the low-gain kids, which is exacerbating your achievement gap. And so therefore that information would be actionable.

This is a really powerful slide. I've used it many, many times.

And it is, why should we take this analysis to the teacher level, with all of the statistical challenges that exist when you do that, which are going to be exacerbated when we talk about special ed? This is why you do it. If you look at this slide, and once again, back to the research where folks debate about whether it's 40% or 65%, this particular slide is from Bill Sanders' 2004 National Governors Symposium, so it's a bit dated. But it still tells the same story.

And what this tells us is that 65% of the variance in gains in students occurs at the classroom level. And I don't really believe that we can get serious about school improvement unless we drill that analysis to that teacher level, and talk about it at that level with all of the perils that exist therein, so.

And I think after a break, I think we're then going to talk about some of those issues in much greater depth, so, Bill?

BILL EAST: Well, I want to thank you for your comments so far. I, you know, as I listen to you talk, it's obvious to me that we need to know a lot more to make wise decisions. And I wondered if you would talk just a moment to, and both of you respond to this

question, if you had to, if you had this question, you know, what are the two or three most important things we need to do in the research arena to give us that knowledge that we need? What would it be? And I'll go to either one of you . . . Lynn?

LYNN HOLDHEIDE: I think earlier today we were talking about that, and I said, every time we talk about this subject, I can think of multiple studies that need to be done, most particularly for special ed and value-added, and we'll talk about that in a moment. But typically special education teachers are not being included in value-added scores.

And we need to determine if they can be included. If students are using accommodations on their statewide tests, does that play a role in determining teacher effectiveness, what role? Does that make a difference? As well as, what about those students that are on alternate standards, or are getting an alternate assessment?

How will those students be, will we get teachers, will the teachers get value-added scores for those students? And what do we need to do in terms of that? I would also like to see the places that are using one system for all teachers. Is that working? Is that helping those special education teachers improve? Is it helping the general education teachers improve? And ultimately, is it helping students achieve, and do better in classrooms? I'd like to see some comparison of that.

Do we want it to be a more general type evaluation system, or do we need to get more specific to make sure that our teachers are growing in the areas that we need them to grow?

BILL EAST: Tony, do you have anything to add?

TONY BAGSHAW: I would, I want to see more research done around those other measures. And to what degree do they correlate with and, are they predictive of the things we really care about, which are college/career readiness for students. And I think there's, I think that's really going to be an exciting space as we try to measure those things.

You now, value-added can't tell us about the role that forming unbelievable relationships with children plays in the growth of kids. It can't tell us the role that creating true engagement in students plays. And I think that's the space that I would like to see more research done. And certainly you could correlate those with value-added, or you wouldn't have to. You could correlate them with other things.

BILL EAST: You know, I just had, my youngest daughter just finished high school last year. And I was constantly reminded of the relationships factor, how important that is between teacher and student. And I haven't seen any real research on that.

We know it makes a big difference, but how do you measure something like that? It's very difficult.

TONY BAGSHAW: Yeah, yeah, yeah.

LYNN HOLDHEIDE: Some of those important features that we know make a teacher a good teacher are difficult to measure. And that's where this becomes very complicated. And if we, I would hope that our goal in teacher evaluation is to drive better teaching, and drive improved student achievement. And to do that, we need to start to identify those things that make a teacher click with a student, and make it work. And those are hard to measure sometimes.

TONY BAGSHAW: Right now we're measuring the things we can measure. But I think there's really frontiers out there that we haven't hit yet.

BILL EAST: I have another question before we get into our slides for the second portion of the conference. We know in teacher evaluation there's a couple ways we can go, the single system or a differentiated system delivery. What are, what is the impact of doing one or the other? What learnings do we have at this point around that issue? Lynn?

LYNN HOLDHEIDE: That's a good question. As folks are deciding whether to do single or differentiated, in some ways, there's things to consider. As I indicated in the slide, a differentiated system makes you value, you can add in the things that are valuable to your field. Whether it be a special education teacher or a music teacher or an art teacher, it lets you add value to what skills and, that you bring to the profession. The difficulty with it is, if you're using an observation instrument or you're using any other type of measures, you worry about the reliability. There's a huge component of training your evaluators, and making sure that if you're doing a separate system, you have two separate systems to train folks on how to evaluate, and that is looking at a resource and a human capacity issue right there.

TONY BAGSHAW: Yeah, I would agree. I think the professional development aspect is a real challenge. And folks often forget about the fact that once we dive in and we do a good job of designing these systems, we have to train everybody to use them. And that is really challenging.

I mean, that's a tremendous allocation of resources. And you, if you create two systems, you're going to have to create two professional development systems, because they're going to be, by definition, different. And I think folks need to think about that long term. That's just going to be, not to say it's not the way to go, but it is going to be an additional resource allocation.

BILL EAST: I know if you use a differentiated system, we used the example earlier, Lynn, about maybe using some factors like the IEP or doing, knowing about the legal aspect of special education. But if you look at the IEP, the special education teacher

would need to know a lot about that. Now a general ed teacher might not need to know as much, but they need to know some.

LYNN HOLDHEIDE: Right.

BILL EAST: So there's multiple levels of what's needed, you know, in the area of professional development. And that's difficult to figure out.

TONY BAGSHAW: Yeah.

LYNN HOLDHEIDE: I've often said to folks who ask me, I said, it's almost like it should be a tiered system, almost similar to RTI, where we have one general system, which evaluates all teachers, but maybe add an additional tier for those specialty areas that we'd need. So then you would maybe lessen the PD requirement for that. You would still have more training that would need to occur for those specialty area teachers and the evaluators, but at least then you would have that sameness, or that similarity across the system generally.

BILL EAST: Okay. I'm going to have more questions later, but let's get back into our presentation now.

LYNN HOLDHEIDE: We're going to jump back into value-added. We talked a lot about what is value-added, and how does it work, and why is it important. And I think we all agree here that it's here to stay for a while. And it is a good way to measure how much growth is happening for students.

Many states and districts really have excluded special education teachers from their value-added models. One of the reasons that it's a challenge for special ed is there's really no research-derived value-added model that's out there for special ed teachers.

We simply don't have the research out there right now with special educators. And that's going to be important to talk about research that we need for the future. Also, we don't know a whole lot about the learning trajectory. And what does that mean? If a student gains ten points on the lower end or ten points on the higher end, is that growth the same?

I don't think we can answer that question. Would you say that, Tony, that we can't answer that question yet, as to ten points down here and ten points up here is not exactly the same as far as the learning curve goes? I mentioned earlier, we have many of our students who take the standardized assessments using accommodations. And we don't have research to show how those accommodations impact teacher-effect scores. And there would need to be some research on that.

We also know we have students that are on alternate standards. If they are on alternate standard or alternate assessments, how will we get value-added scores for those teachers? And right now they're being excluded and not part of it.

TONY BAGSHAW: Right.

LYNN HOLDHEIDE: And the other issue that Tony alluded to earlier, small numbers.

Typically, we have smaller number of kids in our classrooms, and we have smaller N's, so it makes our results a little less reliable if our numbers are too small. Student mobility. We know our population of students. There's a lot of mobility. And if you're moving from school district to school district, from state to state, will you have three years of test results to be able to do those type of growth models? So that's a huge issue.

And then our favorite, which is what Tony's going to talk about here in a few moments, is what do you do in a co-teaching situation? Tony and I are teaching a class together. I'm the general ed teacher, he's the special ed. Are we impacting all students? Are we sharing teacher attribution? Do we think that we both are equally attributing to the students' performance? So those are things that we really need to consider when we're looking at growth models for teachers of students with disabilities.

We did ask, we wanted to go to our experts, our special education directors, regarding this, to see what they thought about using student achievement in teacher evaluation models. As you can see here on the top, 60% percent of our respondents indicated that they agree, achievement gain should be a component of teacher evaluation.

Where we didn't find quite as much agreement is if standardized test scores should be a component. And I think that comes back down to the fact that Tony alluded to earlier. We're not so comfortable that standardized test scores are measuring what we want them to measure all the time, or measuring it in the way that we need it to measure. So there was some contention among that, not real confident that we should use standardized test scores for students with disabilities to determine whether or not a teacher is being effective.

And the last component, we asked the teachers, should progress on the IEP be a component of teacher evaluation? And quite honestly, this result surprised me the most. Seventy-three percent of our respondents, and remember, these are state and local directors of special education, indicated that they would like to see progress on the IEP as a component in teacher evaluation.

Quite honestly, it's a little concerning to me, because I don't think that the IEP was made for that. And I'm not so sure that's a role that should, the IEP should play that role. I think, I'm guessing here, but I'm thinking that 73% kind of came away from the fact that we're not sure what we're going to use. And we know we have IEP goals and objectives that are measurable, so that seems like logical sense. But I think that it's a little bit of a scary step to take in some degree.

Also, as I indicated before, we have a lot of our students in a general education classroom, receiving the general education curriculum, which is phenomenal. But we also in that regard have teachers that are serving in a co-teaching capacity. And what do we do in regards to that?

We asked our experts, who should be accountable, or what students should the special education teacher be accountable for in the classroom? And I'm going to start on the bottom. And the statement was, special ed teachers in a co-teaching role should only be held accountable for students with disabilities.

As you can see there, only 13% of our respondents said, we agree or strongly agree with that statement. What I think we're hearing loud and clear is, we think we have an impact on more than just our students in the classroom.

The second statement we asked folks to respond to is that special education teachers in a co-teaching role should be held accountable for all students. And as you can see, 75% would like to see us held accountable for all of the students, because we do believe that us bringing in differentiated instructional techniques or whatever we're bringing into the classroom has value for all kids.

And then the last one, the statement was, both teachers held accountable for all students in the classroom. And as you can see, overwhelmingly 85% of our respondents indicated that we would really like to see both teachers held accountable. And if you look at co-teaching, that's what co-teaching is about. All students are owned by all teachers, and that we're equally responsible for the planning, the instruction, the assessment, all of that for all of the students that are in the classroom.

So that's what they would like to see, or that's what our experts indicated that they would like to see. When I started this research, taking a look at what folks were doing to evaluate special education teachers, that's how Tony and I became to know one another, because I found that the system that they have in place at Battelle for Kids really was one way to try to share that attribution for learning amongst teachers. And I'm going to turn it over to Tony to talk about that a little bit.

TONY BAGSHAW: Thanks, Lynn. I guess the first and fundamental question is, can we get value-added data for special education teachers? And there are those who, typically, they are just excluded from the analysis. And many times, folks just say, well, we can't do that. Well, that's just simply not true. There are many districts that we've worked with that are getting teacher-level value-added reports for special education teachers.

Now they are more challenging. They come with less certainty, once again, based on sample size. I did a bit of research in some of the districts that we've worked with, and said, okay, what is the lowest number of students you're comfortable with, and still creating a teacher-level value-added report? The smallest we found was five. I will tell you five probably gives me a little bit of heartburn. The largest we found was ten.

Now those are not ten discrete students. And as we'll talk about partial attribution here in a bit, you'll understand why the 10 might be, the 10 total may be comprised from 25 students. And we'll explain that in a bit. Here are the tradeoffs you're going to have to make as you make these decisions. And at the close of the day, it's a local decision.

If you lower N, i.e., let's say you use five, you're going to bring more people into the mix, and you're going to be able to get more reports from more special educators. The problem is, you're going to drive up the standard error when you do that. If you choose a larger N, there will be more teachers where you just can't get a report. But you will increase your certainty.

And certainly we've seen this done a lot of ways. I've seen six, I've seen seven, I've seen ten. It, there seems to be some consensus that somewhere in that neighborhood is about the, I've never seen anything lower than five, in terms of creating a teacher-level value-added report. But you can get the information for special educators. Devil's in the details.

I really believe strongly you cannot ask professionals to make significant changes in their practice unless the information you are putting in front of them they believe is true and it measures what it purports to measure. I just think that's the only way you can do it. And Lynn talked about mobility, co-teaching.

In essence, when we first took the analysis to the teacher level in Ohio, in collaboration with the Ohio Federation of Teachers and the Ohio Education Association, we knew, we're really blessed that our CIO had worked at the state level. And he knew that things like mobility, like co-teaching, that how, in essence, unclear the data was. He said, this won't work. He said, we will not get good information out of this.

And it was a very, very bright fellow who took the linkage software that you're about to see from concept to code in about three months. And I will tell you, Version 1.0 was not very good. But we're now on Version 4.0, and feel like it's a pretty good way to get at these things.

When we started, the whole purpose of the linkage software was for data accuracy. What we ended up finding out through the training and the use of it, that data transparency was really even more important.

I don't want to date myself and use a Madeline Hunter reference, but I will. In essence, when teachers complete that linkage process, it creates an anticipatory set for them, that they know they have a teacher-level value-added report coming back. So they'll start asking for it, and they'll start looking for it.

Okay. What does this do? What are we trying to do through this process? What we are trying to do is we are trying to get the data in the system to exactly mirror what transpired in the classroom. Teacher of record doesn't work. It doesn't work for all kinds of reasons revolving around co-teaching, etc., etc. So what we are trying to do is

we are trying to use a software package to facilitate a conversation between the co-teachers to say, okay, what is fair?

Here's a prime example. Let's say for the sake of argument that Lynn is a seventh-grade math teacher, and I am the special educator. Actually, we probably should flip that, because those were our roles in real life. But anyway, so I have, let's say I have 13 students on my caseload. And we're using a push-in model, and Lynn and I are using, really we're good co-teaching, okay, in the model.

So let's say that all of my students just happen to land in Lynn's fifth-period math class. And there are a total of 28 students in there. So I have my, did I say 13? I have my 13 and Lynn has, you know, we have 28 total. What Lynn and I can do is we can have a conversation around what is fair and justifiable. And you can break it down as low, by 10% increments.

We could, for instance, agree that on my 13 students, since I'm pulling them out for some small-group instructions, but we're probably in, da, da, da, whatever we decide, we could say that I am 80% responsible for the instruction of those students.

So if Johnny is one of those students, I get .8 for Johnny of his gains, whether they're good or not so good. I get .8 of that. Lynn gets .2. If we flip all the other students in the classroom, okay, the other 15, all right, maybe we would decide that we're going to flip it for the others. So for all of those other students, I get .2, or a 20% attribution, and Lynn gets .8, or an 80% attribution.

And you can do it for any, we could elect to do 50/50. Really, my point is, that it is, at its core, it's a professional development opportunity, because Lynn and I have to have good discussions about what is fair, and who's responsible for these, the instruction of these children. And the other thing it does, in the value-added analysis, you've now taken care of that co-teaching piece.

You now are accurately attributing to those students, good or bad, by the way. Okay. So if you'll look at the screen, you'll see that it basically starts with a principal approval process. And I'm not going to spend a ton of time on this. It's a very transparent process.

It starts with the principal. The principal, in essence, makes sure all the teaching assignments are right. Then it goes to the staff. It's a, and then the staff, really the principal, can assign people to the linkage support team. They can elect, I've seen everybody from counselors help with this, to lead teachers, to lead secretaries, to teacher leaders. It's really up to the principal who they assign to help with the process. And then those people have administrative rights, so it's not the full burden of the principal to get this done.

And then what you would see on the screen is really, you would see, the teacher would see their assignments. And they click on their rosters.

And this is the screen I think you'll have the most interest in. You'll see that the first column is their roster. They go through and verify that information. Once again, I talked about data accuracy, but I also talked about data transparency.

In the end, a person is going to get a report back, which is extremely personal. I mean, this is their profession. This is what they have put their heart and soul into. And if we are asking them to make significant professional changes, my experience has been they're more willing to do that if they've verified the data.

So the first thing they do is get their roster right. They can exclude kids, they can add kids. You'll notice that the second column is mobility. It's preset to, I had that student all year long. We know special ed kids are pretty transient a lot of times. So we want to account for that. If I have a student from November to March, they should not count the same in my teacher value-added analysis as a student that I had the entire year. And then the last column is percent of instruction.

Regular education teachers, if they're not in a co-teaching situation, will, they will always set that to 100, because they provided all of the instruction for the students.

We did add, we found, actually last year, that we added an advanced mode, because we were struggling with the in/out kids, the kids that were with us from August to October, and then they're gone, and then they come back in February. And there are a lot of those kiddos out there. So you can literally now go month by month, and click in terms of claiming instruction.

And we talked about this, so I won't belabor the point, but you can actually go student by student, and then map out that attribution, and then that attribution is weighted in the value-added analysis itself, okay? Question, Bill.

BILL EAST: Let me ask a question here before we get too far.

One of the data that you're talking about is what I call, for lack of not knowing what to say, hard data.

TONY BAGSHAW: Yes.

BILL EAST: You know, it's, you can verify the roster. You can verify the mobility. What has me hung up just a little bit is two teachers sitting down and coming up with, how much time attribution do you give, you know? Do I get 80% for those number of students, or do I get 20%, or what? At this point in the development of this teacher evaluation system, are there suggestions, are there, is there a chart, is there something to help these teachers make that decision?

TONY BAGSHAW: That is a . . .

LYNN HOLDHEIDE: That's a . . .

BILL EAST: It sounds real subjective to me, you know?

LYNN HOLDHEIDE: It's a very good question. It's the exact same question I asked Tony the first time we met. I can't imagine that that's an easy conversation for teachers to have, to sit down and say, okay, I get 50%, you get 50%.

That has to be difficult. And remember I asked you, how many calls do you get about that? I mean, how do you, how have you helped them in making that? What kind of professional development needed to occur for them to make that decision?

TONY BAGSHAW: Yeah, yeah. Well, I oversaw the program statewide for two years. And I really recall getting two phone calls around that issue. Now it's entirely possible they didn't get all the way to me, because I understand your point. We purposefully refused to make those decisions. We would usually get those two people on a conference call and say, now look, we're going to walk you through the decisions that you are making.

But I really feel like those have to be local, because I've seen, I've always said that, having observed a ton of teachers, the worst teaching I ever saw in my life was in a co-teaching situation. The best teaching I ever saw in my life was in a co-teaching situation. And I've seen just about everything in between. And we have not experienced folks really, really struggling with that conversation.

And they don't really tend to push back against the notion that it's their decision very much. So we haven't. We have not published guidelines. We give examples. We say, for instance, you know, the one I just shared, just really broad examples. But we really kind of refuse to make the decisions for them.

LYNN HOLDHEIDE: Well, I see it as a great professional development opportunity for the teachers. One of the struggles we have with co-teaching is really not identifying what roles and responsibilities each teacher is going to have in the classroom. It's a real struggle. Sometimes the general education teacher plays more of a role and the special ed teacher's in the background. Sometimes they're more equal. It's very varied across the nation.

But I see it as an opportunity to really sit down at the table and start ferreting out who's going to be doing what. When you start to talk about how much percentage, you start to think about how you're working together within the classroom. And maybe, hopefully, that's facilitating some more discussion as to how you can work together, and how we're all responsible for all of the students that are in the classroom.

It would be my goal that you could walk in and not know which one the special education teacher is or which one the general ed teacher is. And I think this discussion, albeit it might be kind of crazy in the beginning, but is a good discussion in professional development about how co-teaching is going to work in that building.

BILL EAST: Good point. Thank you.

TONY BAGSHAW: Sure.

LYNN HOLDHEIDE: Something like that. All right. We're going to go off of value-added a little bit, and talk about other ways that student growth can be determined. And I'm going to highlight Austin Independent School District, which is in Texas, and talk a little bit about the process that they have. They have what are called student learning objectives. These are developed, all the teachers develop two SLOs, student learning objectives, for the students, for the semester.

One SLO has to address all of the students within the classroom. The other SLO can either address a specific skill, if it's a reading skill or a math skill, depending on your expertise or your content area, or it can target a particular set of students. Of course, I would say students with disabilities could be targeted, and that's how you determine your SLO.

They use a broad range of assessments to determine whether or not there is growth. And they're to be aligned to the state standards and the district expectations. They're developed collaboratively. It's a process where the teacher sits down with the principal and determines the student learning objectives for that particular year, for that particular group or groups of students, or particular skill.

So they sit down and they determine that. They have a rubric, and it's kind of hard to see here, but I'll highlight some things, a rubric that's used to help in determining and writing those SLOs, and determining whether or not those SLOs are met. The first question they ask are, what are the needs? What are the needs of those students within your classroom, or those particular students, or in that particular skill set? What are the needs for those students that we really want to achieve? Which I think is good. It's a needs-based type goal.

The other thing that they look at is what and who is going to be targeted? They have a saying, it says, that targets the needs of an identified population. Here again, that could be students with disabilities, or it could be another population, ELL students, whatever it would be. But they identify a particular population in which their SLO is derived from.

The other question that they have is, what will the students learn? What is it that we want the students to walk away with to show whether or not they've learned a particular skill? The question there is, when they develop what the students what the students are going to learn, they're based on student needs. It needs to be rigorous, it needs to be measurable, and it needs to reflect on students' strengths. Sounds a whole heck of a lot like, what? An IEP.

We're really taking a look at, what are the needs of the students. We're making sure that whatever goals we have written, it's very, it's measurable. Sounds a little bit like an IEP. Then they take a look to see, how do we know if they've learned it?

What is it that we're going to look at? What type of evidence do we need to collect to determine whether or not a student has made adequate progress towards their SLO?

Here they say it needs to be demonstrated as reliable and valid. And I think that's where it gets a little bit sticky. Because these are self-written SLOs, there would be some flexibility into how it's going to be evaluated. And this is where they're really trying to make sure that there's some rigor to the evaluation process, that there is some sort of standards, measures, standard measures of assessment, something that is determining whether that student met that goal or not. And there's some reliability to that measurement.

We also take a look at, what is your goal for student achievement? What exactly do we want them to achieve? And again, is it rigorous, based on past performance? They've set up this rubric, I think, to kind of prevent folks from setting small goals. One of the things that we've been accused of in special education is that we have low expectations, or we may have had low expectations for some of our students, and we need to higher those expectations.

I think this is kind of getting at that. Is it rigorous? Are we taking a look at what type of growth we really can anticipate for this particular student, and not holding them back.

And the last one is, again, how rigorous is your SLO basing that? So the principal and the teachers sit down, and they take a look at what SLOs they have developed, and then they rate them according to this rubric. And that's giving some assurance of the validity of the process, to make sure that it's not just a writing exercise of writing a goal, but actually an exercise to measure whether or not students are, indeed, making the progress that we anticipated them to make.

Some of the strengths of this, one of, I think, the biggest strengths, and I failed to put it up on this slide, is that it's applicable to all teachers. Every teacher can do this, a music teacher, an art teacher, or a special education teacher. This is applicable to all teachers. Whereas with value-added, what we heard before is, in some situations, special education teachers are not included.

We don't have a way to measure non-tested subjects. This is that other 69% we're all concerned about, how are we going to measure? This process, of course, would be able to measure those with all teachers, so that would be a good thing.

It's good professional growth. Nothing's better than sitting down and coming up with a goal with your students, with the teacher, and saying, this is what we want you to accomplish this year. That's great professional growth. The teachers are constantly going to be watching whether or not they're achieving that. They're going to figure out ways to try to make it work better within their classroom. Good professional growth for them.

The weakness is it's heavily dependent on administrators. If a principal is going to sit down with every teacher, that's some time. And it's a lot of training. We need to make sure that the teachers are trained in how to write good objectives, good goals for

those students, and that it isn't just a writing exercise. So there is some resource dependence as far as administrators go.

One of the things that Race to the Top requires is that it's comparable across classrooms. And this becomes very difficult when you are writing independent SLOs, if you are independent from other teachers.

So that comparability is not something that is a strength in this particular situation. And then it's not clear how the student's beginning point is determined, so, you know, what is your baseline, and how do you determine how much growth is there?

Do you have anything? You've looked at the Austin system quite a bit.

TONY BAGSHAW: I have. Yes, I have.

LYNN HOLDHEIDE: Would you like to add anything about the strengths and weaknesses?

TONY BAGSHAW: No, I think you did a great job. I think the one thing that certainly through my evaluation that I sort of asked, I think it's a great system. I think it shows real, there's real promise here. I mean, because of the applicability to everyone, which is a huge strength.

One of the things that I would want to see done in the end is that they would look at the attainment level of the SLOs, and that they would bounce those off of more objective measures, to sort of correlate those. And what I wouldn't want to see is a building with an extremely degree of attainment of SLOs, and other more objective measures say that we're not making progress.

LYNN HOLDHEIDE: Right.

TONY BAGSHAW: But I think as long as those things correlate, I think it holds a lot of promise.

LYNN HOLDHEIDE: Yeah, a bouncing it against these student achievement scores on the standardized tests, something, because with something that can be considered more subjective, you need to be watch, have a watch guard for that.

TONY BAGSHAW: And once again, that's multiple measures, and trying to soften that measurement error by correlating and bringing other things into the measurement mix.

LYNN HOLDHEIDE: Exactly. Exactly. All right. Another example I just wanted to point out, I know I mentioned to you before that I'm a little cautious about using the IEP as a factor in teacher evaluation, but I am bringing this up, because of the fact that 72% of our respondents indicated they'd like to see an IEP as a factor. There is a school district, Norwell Public Schools in Massachusetts, who do use progress on an IEP as a factor in the evaluation process.

Each student has a data binder. It's very data-driven and very, they base their assessments on standard measures that are valid and reliable. And they factor it in.

Now having said all that, the administrator there said, I did a year of intensive, intensive training. It's going to be very, very dependent on the quality of the IEPs, how the goals are written, if they're aligned to the standards, if they're measured appropriately. All of those things play a huge factor in whether or not this would be a viable option for teacher evaluation.

And this last bullet here basically says the same thing. If you're using the SLO or the IEP, which can be construed as somewhat subjective, are you going to counter that with objective evidence? And then you also need to, you need to factor in the amount of time and effort that you need to have in training for teachers to make sure that this is a good process.

I'm going to bring up the Delaware model. Delaware is working very hard on developing their teacher evaluation system. And one of the things that they're struggling with, or not struggling with, but working through, is how to assess special education teachers. And as I mentioned before, they've assembled a group of special educators who've worked for quite a period of time now in identifying how they will evaluate those teachers.

One of the things that they have done is they have had facilitators who've come in, and they've actually trained the facilitators. They come in, they set with this group of special educators, and they have a vast array of different types of special educators sitting at the table. And they're trying to determine which assessments can be used, or rubrics, to determine whether or not there's been growth for students. So they are actually sitting down with the practitioners in the classroom, they're building consensus as to what measurements could be used.

Once those measurements are determined, they will then raise them up to the Department of Education in Delaware, and they'll decide whether or not those assessment meet the muster, I guess you would say, in determining student growth for those kids.

But why I bring them up is that, is to show the need to involve teachers in the process. If they're going to feel like there is fairness, and there's, they have faith in the data that's being given to them, for them to be involved in the process is really, increases your credibility in general with your teacher evaluation system.

We talked a lot about value-added. We talked a lot about other measures where we can get data to show student growth. But we've also talked about there are other things that teachers do and do well, and that also have an impact on student achievement. Some of the things that are being used out there, to take a look at those, are classroom observation instruments.

Large, a lot of districts and states are considering or using Charlotte Danielson's teaching and learning framework, or a modified version of that, some type of

observation instrument, which is almost always augmented with looking at lesson plans, assignments, student works, things that when you look at it together, not only do you do the observation, but you look at evidence. You look at hard evidence of what the student's work looks like, as well as what the lesson plans look like, those type of things. And that is what's being used quite a bit across the nation, in combination, often, with value-added scores or any other type of non-value-added scores.

Other things that states and districts have looked at are administrator or supervisor reports, surveys from students and parents. And I like to think about that quite a bit, hearing from the parents. We all have kids. We all know when our student has been with, our kid has been with a good, really effective teacher. And getting that kind of feedback would be a good thing, as well as from the students.

I've shared this story before, but my daughter's belief is that teacher observations really don't work, because it's a totally different situation when they're being observed than it is on any typical day, and that her idea was that we should have the students, the principal, at the end of the classroom that they've observed, asking the students, if this is a typical teaching that happens within that classroom. But I think we can learn from students. Students really know which teachers are effective and which teachers aren't quite as effective.

And then, of course, there's evidence binders, I think that's what they're called now, instead of portfolios, but some sort of collection of evidence that can be created and presented by the teacher to show what they have done in the classroom that's made a difference.

In terms of observation protocols, we did ask in our survey to our experts of local and state special education directors, if they're using an observation instrument, how many of those align to the state's professional teaching standards? What we learned in that is that 51% said that they aligned to some type of teaching standards. I think there's something to learn here.

We don't have to start from scratch. We've had a lot of committees, a lot of folks really work hard in identifying what it is that we expect teachers to, what skills, competencies we want them to have when they walk out into the classrooms. And aligning our evaluation system to some type of standards is a good thing.

What you can see here is 26% of our respondents really had no idea if it was aligned to any standards at all, or if it were aligned to standards. And what we also saw, and you can see the comments there, is that many systems are just really outdated, and hadn't really been necessarily been aligned to anything, and haven't been used to the capacity that they could be.

When we looked at the observation instruments, we also asked them, how many of you use the same observation instrument as that for general education teachers?

And what we found is 85% of the respondents indicated the same instrument was used for teachers of special education students, or special ed teachers, as well as

general education teachers. Twelve percent indicated they used a modified or different observation system.

And, you know, as I started to explore that 12%, what was going on out there that was different, what I found is, most often, folks were using the same observation instrument, but where they were modifying it for special educators was in the narrative piece of the observation instrument.

So someone with a skill, like a special education director, would go in and observe the teacher. They would, the difference in the observation instrument would be in the narrative part, which I think works if you've got a skilled and trained evaluator. It gets a little bit more, I think, uneasy, when you might not have an educator, or somebody with a special ed background. They wouldn't know necessarily what to look for if it wasn't explicit on the observation instrument.

An example of a state that does modify their observation protocol a little bit differently is Alabama. And they have it slightly modified for specialty area teachers, like speech paths, library specialists, school psychologists. They modify it just a little bit for that.

And then for teachers of students with significant cognitive disabilities, there's some modification in there, in which they use their competencies, and they've extended the classroom to include the community settings, because we know with these teachers, they're often working in the community on life skills with these students.

It was important to extend it beyond the classroom to the community. And then also extending the academic content that listed within the observation protocol to include life, functional life skills, which we know is a very important piece of their education as well.

Some of the contention that we hear from teachers is that, I'm being evaluated by somebody who doesn't know my content, who doesn't know my expertise area.

For example, if I'm going in and evaluating a math teacher, I might not be so good, because I didn't have content experience or knowledge in math. Same thing with special education teachers. When they feel like they're being evaluated by a principal who necessarily understand, perhaps, their roles and responsibilities, or their different instructional techniques used, the legitimacy of that evaluation starts to lose credibility a little bit.

And we asked our experts, how do you feel about that? What kind of experience and knowledge should an evaluator of a special education teacher have? As you can see here, 77% of our respondents indicated they'd like to see the evaluator have specialized training to some degree, so that they know what to look for in a special education teacher, to determine whether or not they're being effective.

Sixty-one percent of our respondents indicated that they should have experience in special education. They should have some background in special ed if they're going to be the evaluator of special education teachers. Another question that we asked our respondents was, do you require any kind of training for your evaluators?

We know training is a huge capacity. If we want any rate of reliability between our evaluators, we need to make sure that they're trained. Sixty percent of them indicated they had some sort of training program for their evaluators, while only 12% indicated they required some specialized training, like for a special education teacher, for evaluating a special ed teacher.

Couple of examples of places that really tried to tackle that issue of the expert or the knowledge base of the evaluator is Toledo's Peer Assistance and Review program. Instead of having one evaluator, they have a team of evaluators, where they can pull expertise from different areas. So if you were evaluating a math teacher, you might pull a couple math teachers in there. Or if you were evaluating a special education teacher, you would pull a special ed teacher in there, and it would be more of a peer review. So that way you would have the expertise sitting around the table when you're doing the review. Another example I wanted to point out was Norwell Public Schools in Massachusetts.

In there, they have formative reviews or observation. There's two that occur each year. One is conducted by the principal, and the other one, then, is conducted by the special education administrator. Each of them are really focusing on their area of expertise.

The principal might be looking at certain instructional techniques, where the special education administrator might be looking for those evidence-based instructional strategies, or taking a look at how to facilitate IEP meetings, and so forth. In the end, they both come together and collaboratively develop the summative evaluation.

TONY BAGSHAW: The other advantage you have with Toledo, certainly, if folks chose to look at that example, would be that that is a long-standing program, and they have refined that many times over the years.

LYNN HOLDHEIDE: Yeah, and it's got a lot of good attention, because, you know, you've got your peers taking a look at you, and which also provides a lot of credibility to the feedback that's coming to you as a teacher, and a great professional development.

If you're watching other teachers and evaluating them, it's really a good professional development opportunity for you as an evaluator, too.

I am going to discuss the District of Columbia Impact program, just to give an example of a district that uses multiple measures. You've heard us talk about the need to use multiple measures throughout this presentation, and it's very important for the validity of the system.

But the District of Columbia has 20 different groupings or categories for teachers. As you can see here, the first group would be teachers where there are value-added scores. Second group would be teachers in which you wouldn't have value-added scores, that other 69% we've been alluding to quite a bit. The third group would be special education teachers, fourth group, the non-itinerant ELL teachers. The groupings

go on and on. But they have different groupings set up for each specialty type, or specialty or content-area teacher.

Within their evaluation system, they have individual teacher value-added scores, which is based on their standardized test in D.C. And that's where they get teacher value-added scores for those that are within that, within those subjects that we can get those scores for.

They also have non-value-added achievement scores. And those are for those kids, that other 69%, where they come up with some standard measures to determine whether there's growth for those students. Those are things like the DIBELS assessment, AIMS. They have a listing, actually, if you go on their website, and I'll show you where a link to that will be. But they have all different types of assessments that can be used for that non-value-added piece for those teachers.

They also have the teaching and learning framework, which is a rubric to observe a teacher and determine whether or not they're being effective in the classroom. That's your observation instrument that they're using. They also have something called Commitment to School, which is taking a look at those teachers and their involvement with other type of activities within school, if they're playing a leadership role.

What I love about the Commitment to School is they have a component on how well they collaborate with their special education teachers, as well as their ELL teachers who are coming into their classroom. They've designed a rubric about this.

So the principal can sit down, take a look at the teacher and, according to the rubric, rate where they feel that they fall in that particular area. They also have school value-added scores. So it's a school-based type score that is applicable to all teachers. And then there's core professionalism, which is actually not factored into the teacher evaluation weighting system, but is also a component that they look at.

What I like about the Impact system is the way that they differentiate for special education teachers. What you see here is that 10% of that rating that a teacher walks away with is based on the non-value-added achievement scores. Fifty percent is based on that teaching and learning framework, which is that observation instrument that I spoke of. Five percent for the Commitment to School, school value-added score is 5%.

And here you see two extra components that their team of folks had decided were very important. One of them is the IEP quality plan, wanting to make sure that the plan is written, it's standard-aligned, there's measurable goals, those type of things.

They felt that that was very important, 15% were of importance. And they also have a rubric that evaluates that. And then IEP timeliness, this is obviously coming back from the regulations and worry about meeting deadlines according to IDEA, and so forth. So they have a rubric to make sure, are the teachers completing their paperwork on time, are reevaluations occurring, those type of things.

The reason I bring this system up is because it does use multiple measures, and it does have a way to differentiate. You see that the core of it is still the same for all

teachers. But the way they differentiate it is the way it's weighted. And then, of course, they've added a couple new components there. So it's something to take a look at, or something to consider.

Within the brief that much of this work is coming out of, we came away with some recommendations for folks that are involved in this process. And to me the first one is the most important one, making sure that special education teachers and administrators are at the table when these frameworks or these evaluation systems are being designed.

We need to be there making sure that our unique roles and responsibilities are being considered in the design process. We want to be part of that, not coming in late and having to change the system that doesn't quite fit us. Whether it's the same system or not, just being at the table, making sure that it's accounted for, special education teachers are accounted for, are very important.

So whether you're a special ed administrator or a teacher, get involved. And if you are the person designing the evaluation system, recruit those folks in, because you're going to get a lot of good knowledge coming from them that'll prevent a lot of redoing things after the fact.

Identifying a common framework that defines effective teaching for all teachers, we talked about that. There has to be a definition. We have to agree upon what is an effective teacher, and then differentiating those if we need to for special educators. And that, I wish I had an answer for that. Should we differentiate, should we not? I think we really need to sit down and take a look at what system's being designed, if we feel comfortable, if it's a good reflection of a special education teacher and their effectiveness, or if we need to differentiate it in some degree.

Integrate evidence-based practices. We do have evidence. We know practices that can elicit improved student achievement for our kids. Let's make sure that those are integrated into whatever evaluation system is there, because what's integrated and what's being measured is going to be valued and, hence, will be happening in the classrooms. Improve data quality. I don't know if you want to speak to this at all, Tony, but . . .

TONY BAGSHAW: I'm for it.

LYNN HOLDHEIDE: Yeah, you're for it. We're all for data quality. But I can't speak to that enough. If the data doesn't feel valid to the teachers, then everything else is just gone out the window. All the professional development, all the things that we're trying to do to help teachers become more effective really goes out the window when they don't believe that the data is quality in the beginning.

In addition to, or in other situations where there's an absence of standardized assessment data, which is going to occur with many of our teachers, we need to come up with other evidence. And I've talked about how Delaware's going down that path, and other districts are doing the same.

Ensure that evaluators are trained, that they have some sort of explicit training for evaluators of special educators, that they know what to look for to determine whether or not they're effective. And last but not least, and I think that this is so very important, is that we need to collaborate with teacher preparation programs. If we want folks to come out that they're effective teachers in the classroom, we need to make sure they're being prepared to be effective teachers in the classroom.

For example, if evidence-based practices does become a large component of teacher evaluation, we need to make sure that preparation programs are arming and tooling those teachers with those skills, so that they can come out and utilize those in the classroom.

And I'm going to just preface this with a link to much of the work that we just talked about. We talked about many different districts that are, and examples of places and what they're doing. If you go into this document, you will find links to the Impact program, to Austin, so you can get more information on those sites as well.

BILL EAST: Well, thank you so much, Lynn Holdheide and Tony Bagshaw, for your excellent presentations today, talking about teacher evaluation. I know I've learned a lot, and I'm sure our audience has done so, too. You know, an observation that I've made just sitting here listening to you talk today, it's the same observation I have made is I have done almost 50 of these professional development conferences now, and it seems like always that presenters provide great content, there's great professional development going on all over this country.

But something that almost every presenter has said somewhere in their presentation is once you do the professional development and start implementing the program, teachers need time to talk to teachers. Teachers need time to talk to administrators. We need time to talk. And what we really don't do, in my estimation, is we don't value time to talk, time to meet, as we try to implement our programs.

And that's a theme that I would hope that we would think about as we try to implement this very important program around teacher evaluation, or any of the others that we've promoted here at NASDSE, like Response to Intervention, the positive behavioral supports, or co-teaching, or whatever. So it's the follow-up, it's time to really talk about how things are going and how can we get better.

LYNN HOLDHEIDE: And how to learn from one another. There's wonderful, great teaching going on out there, and we can all learn from one another. And if we can build in that time and that collaboration, we're going to be farther down the road.

TONY BAGSHAW: Yeah, time is our most precious commodity. And that's when I ask about, do we have the will to allocate those resources to do this well? It's going to take more time to do this well.

BILL EAST: Well, again, I want to thank Lynn and Tony for your time with us today. If you have comments about this conference or any of NASDSE's services or programs, I remind you to go to our website at www.nasdse.org and make your comment.

I want to thank you for your support of NASDSE's Professional Development series. And that's our conference for today. Thank you so much.