THE ALBERT SHANKER INSTITUTE

NATIONAL PRESS CLUB FORUM

EARLY CHILDHOOD ASSESSMENT: PROBLEMS AND POSSIBILITIES

MURROW ROOM
NATIONAL PRESS CLUB
WASHINGTON, D.C.

12:00 – 2:00 P.M.
OCTOBER 6, 2004

GREETINGS:
NAT LACOUR,
SECRETARY-TREASURER, AMERICAN FEDERATION OF TEACHERS

FEATURED SPEAKERS:
FRED MORRISON,
PROFESSOR OF PSYCHOLOGY, UNIVERSITY OF MICHIGAN
JACQUELINE JONES,
DIRECTOR OF INITIATIVES IN EARLY CHILDHOOD
AND LITERACY EDUCATION, EDUCATIONAL TESTING SERVICE

MODERATOR:
BARBARA BOWMAN,
PRESIDENT EMERITUS, ERIKSON INSTITUTE

Transcript by:
Federal News Service
Washington, D.C.

Go to Full Text of Forum
Go to List of Attendees

LIST OF ATTENDEES

Alicia Ardila-Rey
American Association of Colleges for Teacher Education

Amy Berg
Century Foundation

Heather Callister
Abell Foundation

Jennifer Carlson
Head Start Bureau
Administration for Children and Families

Toni Cortese
American Federation of Teachers

Yasmine Daniel
Children's Defense Fund

Carol Brunson Day
Council of Professional Recognition

Shyrelle Eubanks
National Education Association

Danielle Ewen
Center for Law and Social Policy

Toks Fashola
Johns Hopkins University

Michelle Figlar
National Head Start Association

Mark Ginsberg
National Association for the Education of Young Children

Danelle Gonzales
Trust for Early Education

Darion Griffin
American Federation of Teachers

Beth Hale
Institute for Educational Leadership

Jane Hannaway
Education Policy Center, Urban Institute

Lisa Hansel
American Federation of Teachers

Janet Hansen
RAND Corporation

Warren Harris
Teaching Strategies, Inc.

Fiona Helsel
American Institutes of Research

Marilu Hyson
National Association for the Education of Young Children

John Jackson
NAACP

Linda Jagielo
Head Start Bureau
Administration for Children and Families

Laura Logerfo
Urban Institute

Anna Lovejoy
National Governors Association

Jana Martella
Council of Chief State School Officers

Denise McKeon
National Education Association

Dori Mornan
Center for the Child Care Workforce

Mildred Otero
Children's Defense Fund

Jennifer Park-Jadotte
Teaching Strategies, Inc.

Nancy Reder
National Association of State Directors of Special Education

Cheryl Riggins-Newby
National Association for Elementary School Principles

Miriam Rollin
Fight Crime, Invest in Kids

Stephanie Rubin
Trust for Early Education

Karen Schulman
National Women's Law Center

Tom Schultz
Head Start Bureau
Administration for Children and Families

Paul Skoczylas
The White House

Joan Baratz-Snowden
American Federation of Teachers

Gerald E. Sroufe
American Education Research Association

Howie Schaffer
Public Education Network

Jon Vaupel
Children's Defense Fund

Marci Young
Center for the Child Care Workforce

Early Childhood Assessment:
Problems and Possibilities

NAT LACOUR: *Assessment of preschool is important so we can see how individual children are progressing, guide their teachers on what they need to do to help them, and get a sense at whether or not programs are working.*

We haven't done it right in respect to K-12 children so it certainly is a challenge for those working in the preschool years. But we need to get it right because we know from many research studies that learning gaps formed in the early years can stay with children all their lives. Our goal must be to avoid a gap all together, before it even sets in.

As many in this room know, there are tensions about how to do this right. We don't want assessments to narrow the scope for what children learn, and we want them to be fair, to measure what children truly know without distorting the true achievements of preschoolers from a variety of backgrounds. We need to know more about the problems in assessment.

So, as a former teacher and union leader, I will tell you that of all the reforms we have attempted and talked about -- going all the way back to 1983 -- I think that the best thing that can be done for children who are in our schools, particularly those that have yet to start, is to work on putting together a program in early childhood that will prepare them so that they will not fall behind. Most of the kids who fall behind remain behind throughout their schooling.

Early childhood education is a very critical arena. Not enough people have come to appreciate it, but a growing number of people are. And I think that it is the job and the responsibility of each of you in here to help make this a true high priority throughout this country, at all levels – national government, the state government, and local schools. We want this to be a program in which there is great participation by all of you.

Now I want to introduce our moderator for today's session.

Professor Barbara Bowman is one of the three faculty founders of the Erikson Institute, and served as president of the Erikson Institute from 1994 to 2001. She is an

authority on early education, a national advocate for improved and expanded training for practitioners who teach and care for young children, and a pioneer in building knowledge and understanding of the issues of access and equity for minority children.

She is past president of the National Association for the Education of Young Children, has served on numerous boards, including the High School Educational Foundation, the Institute for Psychoanalysis, Business People in the Public Interest, the Great Books Foundation, the Chicago Public Library Foundation, and the National Board for Professional Teaching Standards.

She earned a B.A. in Child Development from Sarah Lawrence College and an M.A. in Education from the University of Chicago. Among the honors she has received are honorary doctorates from Bank Street College, Roosevelt University, Dominican University, and Governors State University. We know she is among the most respected people in this field and we are very pleased she agreed to be with us today. Thank you.

(Applause.)

BARBARA BOWMAN: *Thank you. That's quite a nice introduction. I'm just delighted to be here today. I thought I would begin by building a little context because I'm sure there is nobody in the room that has been in field as long as I have. (Laughter.)*

And perhaps it would be useful to remember how we got to this place. And this place that we're at right now is how we should be assessing young children. So let me remind you that about sixty years ago, we had our first inkling that early childhood was going to take center stage when Ben Bloom, and McVickers Hunt, and Bronson Benner alerted everybody in the educational world to the fact that the kind of experiences children have early in life make a difference.

And certainly by the early sixties we were pretty well assured, at least in the academic community that early education could change the trajectory of children's lives, particularly children at some risk for educational success and for special needs. Forty-five years ago, we began our experimental research on programs for young children. And, of course, many of you remember the High/Scope research. I was appalled to find out that the High/Scope children are now in their forties. (Laughter.) That doesn't seem possible. (Chuckles.)

But the latest iteration of the High/Scope work is coming out in another new book, "The High/Scope Children in Their Forties." They were the first one to alert us to the fact that we could build programs that could have a long-term effect on children's lives. And certainly that has been followed with the Abecedarian Project now, so that we now have very good, robust evidence over the last 45 years that, if you have high quality programs, it can make a difference in kids' lives.

Beginning about 35 years ago, we began to be concerned about the differential effects of those programs, and many of you will remember the first Westinghouse Report

that pointed out that the summer programs weren't having too much of an effect, but, *indeed, some of the full-year programs did.*

And that was followed very soon by the plan variation studies, which showed that there was enormous variation, both within and between programs. So we had an increasing body of evidence, beginning in the 1970s, that all programs didn't result in wonderful things happening to children, but certainly some of them did.

About 30 years ago, we began to look at the programs that were having effects on kids to see if we could determine what effects different types of program variables were having. And, again, those of you who have been in field a long time may remember the study where observation of children in the centers pointed out that group size, teacher-child ratio, and teacher education were critical variables in what effects the program would have on children.

That was followed, certainly, about in the next 10, 15 years by an additional stream about interactions and the important of the quality of the interactions between teachers and children that affected the outcomes of programs, and we have Debbie Philips and Carollee Howes, who summarized that information for us.

Then about 25 years ago, we began to see a switch from concern about program variables and performance standards, if you will, to a question of – well, individual children, how are they doing? And I think we began it with the requirement that we assess all of our children for developmental delays, and certainly in the last 25 years we've made that pretty well institutionalized, and, certainly for low-income children. And the new assessment instruments for special needs are certainly ubiquitous.

We also began to look at classroom performance, and work sampling, a variety of other classroom performance assessment instruments, where we tried to figure out what kids knew and were able to do, if the notion of changing curriculum in order to be responsive to them.

About 10 years ago, we began this – as the rest of the education world did – we began being concerned about accountability, and certainly the 10 letters, and which 10 letters kids ought to know hit our early childhood community, followed very quickly by the national reporting program and other ways of looking at testing as a way of assessing what young children know and are able to do.

And that's kind of where we are today. I think there seems to be a general acceptance of the notion that we ought to know what kids can know and can do, and that we ought to know which systems and which configurations of systems achieve those kinds of benefits. Certainly the National Head Start Association has gone on record as being interested in and approving of our efforts to assess children's learning accurately.

On the other hand, the current proposal by the – or the current program of National Head Start has drawn a lot of controversy and a lot of concern. The concerns

tend to focus around, I think, three major issues. The first one is age. Can we really individually assess children of such early age? Anybody who's been around three and four year olds knows that they may or may not respond appropriately to questioning; and some concern that, in testing them, we're simply getting an incorrect assessment of what they can know and can do.

The second concern is content. What we can ask children that is easily coded in a test tends to be the most superficial knowledge that they have. And so there is great concern that in order to devise a test that is economically feasible, we have to stick with content that is unimportant.

And the third that we're concerned about is its predictive value. Do any of these tests that we give young children have much to do with their later educational achievement?

Those are, I think, the major questions that I see that we ought to have under discussion today and we've got two experts here to help us discuss them. And I'd like to tell you a little bit about the format we're going to use today. First, Fred Morrison is going talk for a few minutes, and then we will have a chance for you to ask him questions or to raise other issues. Then Jackie Jones is going to talk, and we'll talk a little bit about what she has to say, and then we'll open the floor up for general discussion about both papers.

We will finish, I promise you, absolutely on target, at 2:00. So I would encourage you, as you're asking questions or making comments that you keep them as brief as possible. This is being taped, so when you make a comment or ask a question, please give your name, and we'll be able to attribute the very elegant question that you ask to the right person.

Let me start introducing Fred Morrison, a professor of development of psychology at the University of Michigan at Ann Arbor. I knew Fred back in the olden days when he used to be a professor at Loyola University and a colleague of ours at Erikson Institute, so we were very sad to see him go. His research focuses on the nature and sources of literacy acquisition in children during the transition to school. His research has uncovered surprisingly large individual differences among children in important cognitive language and social skills, even before they begin school. His current work examines the impact of child, family, and schooling factors in shaping children's growth and contributing to early problems in school. Fred, take the mike.

FRED MORRISON: Thank you, Barbara. It's a pleasure to be here and I want to thank you, Genie [Eugenia Kemble, Executive Director, Albert Shanker Institute] and Nat, and the Albert Shanker Institute.

Genie and I were talking on the phone last week. Albert Shanker was one of my heroes growing up. Didn't have the pleasure of meeting him, but

I never missed a column in the Sunday Times and he was both wise and commonsensical. I can't promise to be that this afternoon, but at least maybe be informative.

As Barbara said, my work and my expertise really is in early literacy, particularly the complex factors that shape children's trajectories over the school-transition period. I, obviously, for that focus, have become very interested in the whats and hows and whys of assessing children, and so it's on that basis that I offer some of these comments.

[Intro]

I think there are a number of important questions around preschool assessment: why do we care about assessment at this point in the history of our concern about children? What's the role of assessment, what should we assess, and how should we assess it, and, hopefully, we'll be able to get to all of those in the course of the next hour-and-a-half or so. But I thought maybe the most useful thing that I could do would be to try to lay the foundation by talking mostly about what we should be assessing, and more particularly, what has research told us over the last 10 or 15 years about what we should be looking at? What are the key component skills, if you will, that are most uniquely predictive of success in school in which we should be focusing on in assessment? But also, I think there's a related question, which I hope we'll get to; namely, what are those skills that we can work on? I mean, I think assessment is important, but ultimately, assessment for improvement, I think, is what we all ultimately are aiming for. We want all of our children to be as functionally literate as they can be.

And, so, assessment gives us a part of that story, but it raises questions about what are we using these particular skills and these particular assessments for. So I want to be able to get to that, and maybe I'll weave it in and out of my comments about what we should assess, but we may have to get to it later.

[Three primary domains of functioning in early childhood development]

I think it's fair to say that over the past 10 or 15 years, three major domains of functioning have surfaced as key to understanding early childhood development, successful school readiness, and successful school transition. That's language, as one major global area; literacy, as a related and separate area; and what I will call, for the moment, self-regulation. I

don't want to call it social skills because I think that is bit too broad, and we don't know enough about things like peer acceptance or sociability and its predictive value; but we do know a lot about self-regulation or executive functioning as a predictor.

There's a fourth domain that I'm just going to mention -- that I think is going to turn out to be important, but we don't know enough about it during the preschool years -- and I'll call it motivation. As some of you know, if you had college courses – or like Laura, who is a student at Michigan, had courses with Jackie Ekles (sp) -- motivation is a big topic where I come from; but we really know most about that with regard to older children and adolescents.

There is some work by John Guthrie on the concept of engagement. So we can talk about that if you want. I think ultimately there are going to be these four domains that are most diagnostic, most revealing about children's early development; but three of them I think we have a reasonably good research base on.

[Language Domain: Vocabulary, Semantics, Syntax, Phonemic Segmentation & Pragmatics]

The first is language; and originally I think we thought we had a fairly delimited domain within language, essentially things like receptive and expressive vocabulary. But that has expanded in the last 10 years because of research of Jane Ellen and Hunten Locker , and Hart and Risley and other people like that.

But the key components that predict, uniquely, children's early reading skills -- at least to about kindergarten, and then they begin to predict children's comprehension skills, starting about third grade -- is vocabulary. By that I mean both receptive and expressive vocabulary. We also know that those two skills -- or research has demonstrated, at least, that those two skills can be relatively separate from one another. They don't necessarily go hand in hand. So there is evidence, for example, that receptive vocabulary is in fact very responsive to parental input – amount of talking as Hart and Risley, and others have demonstrated.

Much less so is expressive vocabulary. Expressive vocabulary is actually more responsive to parental warmth and responsivity; in other

words, how well you listen to your kid, so that those bedtime stories – when you want to whip through that story, you know, you're tired and you want to get to get to bed, too, and you sort of clump three pages together -- (laughter) – well, that's having an effect on expressive vocabulary, particularly your child's ability to point at the pictures, and talk about the pictures, and go on and on and on, when you want to go to sleep. (Laughter.)

The research demonstrates pretty straightforward -- and I think it's probably the most impressive -- that individual differences in vocabulary, as young as three years of age, will predict entering reading skills at kindergarten and also predict comprehension skills, starting in about third grade. So I think as an important domain to look at, it's absolutely critical.

Now, the other thing I'll say at this point – I'm talking about these as individual component skills, but they really do not act in isolation; so ultimately, we've got to understand the interaction. For example, vocabulary obviously interacts with core knowledge or semantics; it's a very critical area that I'll talk about in a minute. But there is also evidence from Russ Whitehurst and Chris Lonigan that vocabulary may, in fact, shape phonological awareness over three and four years of age; so that the larger your vocabulary, the more likely you are to hear lots of different words; the more likely you are to hear words that begin to rhyme or have the same starting letter; the more likely, earlier on, you are to make language a kind of object -- begin to look at it; and that's a critical step in early childhood for reading acquisition.

The second component that has really recently resurfaced is syntax. I think most of us who took a language course were taught that, well, by the age of four years, all kids have syntax. It sort of develops from ungrammatical to grammatical in a relatively straightforward fashion. And for a lot of kids, perhaps the vast majority of kids, that's true. However, it is not universally true. And the most at-risk children often do not have the sophisticated syntactic skills they need to comprehend in a listening fashion and begin to learn how to read.

There's recent evidence -- and it's very new, so I don't want to necessarily emphasize it too much, but it could be quite provocative – from Hollis Scarborough about the impact of African-American vernacular English on early reading can be so profound that the greater use of African-

American vernacular English has been associated with poor reading skills. Now this is separate from children who code switch. That's a whole different phenomenon. But the emphasis on syntactic irregularities or syntactic different with traditional English has been noticed by Hollis as one potential stumbling block for African-American students learning to read in traditional English.

Semantics is the third domain. Essentially, there, that refers not only to word meaning but to core knowledge. And we know that core knowledge in lots of different areas, particularly in young children in biology, and psychology, and physics, is a strong predictor of early word decoding and vocabulary, and later comprehension.

I guess, probably the biggest discovery in the area of language in the past 10 or 15 years is in the area of phonology. That's a complex area, but as most of you know, it refers to the sounds of language. And there are different levels of phonological processing. There's syllabic processing or syllabic segmentation; sub-syllabic segmentation; and phonemic segmentation. And there's a developmental progression, so that many children, by three years of age, do have some sensitivity to syllabic segmentation. Fewer children, but some, do have sensitivity to sub-syllabic – if you've never heard of sub-syllabic, that's essentially, if I have the word back, and I ask you to split off b and ack, okay, that's a sub-syllabic segmentation, as opposed to syllabic.

Phonemic segmentation, or phonemic awareness is very difficult for young children, and that has posed the most significant stumbling block for people who want to assess it because it turns out a very important predictive skill for early reading. Whereas vocabulary and phonological awareness both uniquely predict reading around kindergarten, starting about first grade, vocabulary starts to drop out as a unique predictor. It actually seems to operate indirectly through its effect on phonemic awareness. And the importance of phonemic awareness really continues throughout the early elementary school years.

So we can probably assess things like syllabic segmentation and, to some extent, sub-syllabic segmentation maybe about four years of age pretty reliably. We can do some assessment of phonemic awareness in middle- and upper-middle class kids around four years of age; but it's still very difficult to get a measure of phonemic awareness.

I must say, in the National Reporting System Technical Working Group that Jacqueline and I are on, that's one of the things that we've really been struggling with. So in terms of assessing, it's the Holy Grail at this point. If we could come up with something that would really look at that, it would be very, very important.

Finally, there's pragmatics, which really hasn't been looked at much at all. If you're not familiar with pragmatics, when I lecture about this to my students, I give them the example – we used to live in a colder place than even Chicago up in Edmonton, Alberta. And my little son who was a preschooler at the time had the habit of sort of bolting in the door, running to watch television or his room – leaving the door open. And you can imagine what that would be like in January in Edmonton, Alberta.

Well, instead of saying, please close the door, I would say something like, wow, it's really getting cold in here. Now that's a pragmatic use of language. I didn't really mean, gee, it's getting cold in here. I meant get your fat butt back here -- (laughter) -- and close the door, okay? So it's a higher-order skill that hasn't been really looked at much across the early childhood years.

So vocabulary, syntax, semantics, core knowledge, and particularly, phonology – that core set of language skills, which uniquely, but then in interaction, shape those individual difference before kids ever get to school.

[Domain of literacy: Alphabet knowledge, phoneme awareness & print concepts]

In the area of literacy, probably the foundational skill which people have already mentioned is the alphabet knowledge, and by alphabet knowledge, we include both knowledge of letter names and knowledge of letter sounds. Now, I should mention, if I were amongst scrapping academics, some of them would tell me that phonology is actually a literacy skill, not only a language skill; that's kind of a footnote to all of that for those of you who are wondering why I'm sticking it in with language.

Alphabet knowledge turns out to be a very solid predictor of early reading. And in connection with phonology, especially an understanding or an awareness of, if you will, phonemes, that allows the child to begin to put

together letters and sounds to, as we say, crack the code. But the research suggests, anyway, that unless you have a fairly solid knowledge of those little sounds, those phonemes, then it's difficult to link them to the letter sounds, the letter names; to be able to, in a sense, now look at words and be able to sort of decode them in a systematic way, at least during your early stages.

So letter-sound correspondence is a separate, but a stage up, from letter knowledge and phonology. It, in and of itself, has a whole series of stages that Lenia Erie (sp) has described that children go through: from partial alphabetic to full alphabetic to consolidated alphabetic decoding; but those are some of the key components that begin to be constructed around five years of age for some children, but most, say, in first grade.

Finally, there's the whole notion of print concepts. And this refers to the sort of knowledge you gain about books and about reading. So knowledge of what – when a book is right side up versus upside down, knowledge of reading from left to right, knowledge of what's a picture and what's a word – making all of those distinctions turns out also to predict early reading skill.

Now, I must say, though, that there is some controversy about how important or unique print concepts are. Some people feel, from data interfaces and other projects, that it is a strong, unique predictor. Others like Hollis Scarborough think that, early on, and particularly for very, very at-risk children, it might be a predictor; but for must American children, particularly when you take into account social class, it doesn't have the predictive value of these other skills. Okay?

[Domain of self-regulation]

Finally, I think probably one of the most critical issues, at least from a practitioner's point of view, that we're only now beginning to get a handle on is self-regulation. Some of you saw the Newsweek article a few weeks ago about parents that can't say no, or if you see these news reports of kindergarten children who are suspended for beating up or stabbing their teacher, well, you can recognize the domain we're talking about here.

There is growing evidence that differences – individual differences in the ability to concentrate, in the ability to sit still, in the ability to inhibit

inappropriate responses, and related social skills, if you will -- what we're calling self-regulation – emerge very early. And we know from several studies in our own lab and other labs that, even controlling for IQ and controlling for the literacy environment, variability in self-regulation uniquely predicts kindergarten literacy skills, and reading and math in second grade. So I don't think we need to belabor that any more, but one of the areas that's been most difficult to try to assess in some objective way is self-regulation.

So I think with that, maybe I will just stop. I think the good news, in one sense, is I think we're getting a fairly solid handle on the core set of processes that do predict successful school transition and that we need to think about in terms of assessment.

[Conclusion]

Now the last thing I'll say, though, is, how you assess these depends on a lot of different factors. So as some people have indicated, the extent to which you want to use assessment for, say, program purposes, or just to get a sense of how different programs are going, well, assessing vocabulary, or assessing math skills or whatever would be fine. But some people have argued that that's kind of an outcome approach to processes and assessment, and what we really want to be looking at are the more underlying mechanisms or processes, especially if we're thinking about intervening. So rather than just saying, okay, this child has x numbers of vocabulary items in his or her repertoire at this age, maybe we might want to ask, well, how are vocabulary words learned?

And people have been working on a concept called fast mapping, which is a way of looking at how children literally learn vocabulary items. That, in a sense, may give us a handle on how to increase a child's vocabulary over and above, or more than just assessing vocabulary with the PPBT; but that's an item that we can discuss. So thank you.

MS. BOWMAN: *All right. We now have a chance to comment or ask questions of Fred about his analysis of what's available out there.*

MS. BOWMAN: *Eugenia. (Pause.)*

Q: Oh, okay. Eugenia Kemble, Shanker Institute. I wonder if -- you know, this might take more time than we really have – but if you could briefly say – I mean you've

referred to where it's difficult to assess along this spectrum of skills. Could you say something about what you think the state of the art is right now, with reference to those places that you said you needed to be able assess and those that were the most difficult to be able to assess?

MR. MORRISON: It's my assessment that with regard to the component skills that I outlined, except for self-regulation, that we have a reasonable set of instruments that are reasonably valid and reasonably reliable. Now, I think each instrument will vary on the reliability scale at different ages. So, for example, I think assessments of vocabulary are very solid and you can get some of those even at two years of age. The predictive value of those is pretty strong.

As I said, measures of phonological skills begin to be more solid, if you will, reliable, around four; but I think even for the most at-risk children, that's still a bit of an issue that I think we need to deal with. And things like alphabet knowledge and letter-sound correspondences, I think they're pretty solid by three or four years of age.

So with regard to the questions that Barbara raised, there's some variability about our confidence with regard to age; but I think most of these -- maybe others who have experience might differ -- but most of these, except for self-regulation, are really quite valid and reliable by four years of age. The problem really is with three-year-olds, I think with some of these things, especially phonemic awareness.

With regard to content, I think Barbara raises a good point. I mean, I don't know whether you feel like most of these skills are simple or superficial – they turn out to be the strongest predictors of early reading, up to about third grade. So we can talk about what we mean by simple or superficial, but overall, I'm not sure that we're really missing a whole lot that's going to make a huge difference in how we assess or what we predict about the trajectories of children over the school-transition period.

And I think, somewhat distressingly, the predictive value of these tests is quite high, and certainly in our own data, we have very strong predictability from four-and-a-half to five years of years, all the way up to sixth grade. I mean, it used to be the case that we thought that, well, children's trajectories were kind of set by about third grade; but I think with Barbara's work and Deb Phillips and others, we've realized that the stability

of these individual differences seems to set in a whole lot earlier than we had thought. So I think the predictive value – it's not perfect, obviously, but I think it's surprisingly high.

Q: I'm not a researcher, I'm a policy wonk, so forgive me if I don't have all the terminology down. Oh, I'm Miriam Rollin, I'm with Fight Crime: Invest in Kids. But I understand that there's some research in the High/Scope area, that was referred to earlier, that talked about different approaches that -- more directed instruction, less directed instruction, some of which may result in short-term learning gains up to third grade, but actually may not have the longer term results that one would want in terms of crime reduction, in terms of graduation rates, et cetera, by age 18.

And I was wondering if any of your work touched on those, on relating the short-come outcomes, our relatively short-term outcomes of school readiness and successful transition, and good success in second and third grade with correlations of later successes – well, graduation rates, and crime outcomes.

MR. MORRISON: Well, I would have to admit not being sure exactly what findings you're referring too. Maybe Barbara can speak to it, but my own sense of that issue is just that all of these things are cumulative; that we're not talking about an inoculation model here; that the appropriate assessment and the appropriate intervention prior to school will solve all of our problems. You still got to have high-quality instruction all the way through and high quality parenting, too.

So I don't know if that's really responsive, but perhaps I'm not as familiar with the findings you're talking about. Do you know?

MS. BOWMAN: *The third grade test scores do seem to be fairly predictive later of achievements, so that whatever predicts third grade test scores does seem to predict eighth grade graduation rates and so on. But, I think, maybe what you were talking about was the fact that there were certainly some studies that showed that the (disk star?) and the more directed instruction programs – the kids learned to decode very quickly, but they didn't have very much content, understanding, comprehension, therefore they didn't read very much from third grade on, and then by fifth grade, they began to show the effects of not reading very much. They weren't practicing -- (chuckles) – if you will. That's the only I can think of.*

MR. MORRISON: Oh, yeah, well, that quickly just suggests another theme about the specificity of these influences. That's the other part, that essentially – for example, we know book reading to children has a big effect on their vocabulary, but not much effect on alphabet naming, alphabet knowledge, or letter-word decoding; so, in a sense, you've got to focus your

instruction on what it is you want the kids to learn. It isn't just one global rosy picture.

MS. BOWMAN: *And perhaps we ought to add that Russ Whitehurst's data suggests that it's not just reading the book, it's the interaction of around reading the book with core knowledge that makes the difference, not just reading as reading.*

Q: Danielle Ewen from the Center for Law and Social Policy. I was glad to hear you say that pre-K isn't inoculation. Thank you for that. But my real question is, can you address the question you just did about validity and reliability for children who are immigrants, for children that are non-native speakers of English, for those children who are at risk, who are most likely to be in our preschool programs right now.

MR. MORRISON: Well, that's something we're really looking very hard at in the technical working group. I would say – to be honest, I'm not as much a part of that group as others. So I think that is a serious issue. I don't think we're as close to solving that issue as we are with traditional native English speakers. Maybe Tom could say something about where he feels we stand on that.

MR. SCHULTZ: I'm Tom Schultz from the Head Start Bureau. As you know, we are assessing children in both English and Spanish in the Head Start assessment. We did find that the Spanish versions of the assessments had somewhat less technical quality. As we looked at them originally, we have found, as the programs have gained experience in administering the Spanish assessments, that the reliability has been improving. I think we're also, basically, using this as a mechanism to look at comparison within the universe of Head Start programs, and so we're not trying to compare these children against national norms or other populations of children. I think we need to do – we really don't have measures, at present, that we can use for children who come into preschool or Head Start that speak other languages than English or Spanish. So that's a big issue that we need to work on.

Q: Okay, you mentioned that the use of, okay, use of a black dialect has been found to hinder African-American children – (off mike) – in the development of syntax –

MR. MORRISON: No, no, no. The finding is the greater proportion of use of African-American vernacular English, that that predicts their reading skill in kindergarten. So it's not necessarily – I mean African-American vernacular English is largely – it has a lot to do with syntax, so that is what the measure is.

Q: (Comments off microphone relating to Morrison's discussion of recent research on African-American vernacular and possible effects on reading levels in early childhood and elementary school.)

MS. GRIFFIN: I'm Darion Griffin and I'm one of the co-authors of that articles. And the study was not about the use of African-American vernacular English, but it was about familiarity with school English. We intentionally looked at it that way. So what we found was that students who were more familiar with school English had less difficultly, had higher achievement in comprehension by third grade, as opposed to the inverse, which is we did not study the extent to which students used AABE and then correlated that to reading comprehension levels measured by standardized tests.

Q: Okay, so when you said school English, you mean standard?

MS. GRIFFIN: Standard English, academic English. We chose to call it school English.

MR. MORRISON: Thanks. Yeah.

Q: I'm Marilou Hyson from the National Association for the Education of Young Children. This is an observation that I would appreciate some comment on. It seems to me that one of the challenges is making decisions about what to do with assessment results. As you were describing, many of these assessments are assessing fairly straightforward aspects of early language and literacy development. Kind of a quick fix approach may be to remedy the narrow deficiency that a particular child seems to have with respect to certain skills. Is that necessarily the appropriate or effective way to use those kinds of assessment results – and this may be previewing some of what Jacqueline might be talking about.

MR. MORRISION: Marilou, I'm not sure what you are getting at.

Q: Okay, so for an example, let's say that a child – and I think you kind of alluded this a little bit – a child is deficient in vocabulary on the basis of a certain assessment. So is the remedy for that gap to teach the child more vocabulary words? And that would be just one example.

MR. MORRISON: Yes, but, I mean, not in isolation, I think. How you would do that and how you would try to accomplish that would be a separate issue. But no, I think that would be the point. You would try to enrich that child's vocabulary, but in so doing, you would also be presumably, you know, enriching other things like core knowledge, you know, semantic diversity and things like that. So I don't – I tried to emphasize at the beginning, these things are interactive and intertwined so that the isolated effort to just sit down with vocabulary items is probably not, you know, really what we are talking about here.

Q: That's something that (has been said?)

MR. MORRISION: I know, yeah, but that's what I think we need to emphasize, that these are interrelated.

MS. BOWMAN: *Hold your question, write it down because we are going to have some time at the end to have more general comments, but I do want to give Jackie time to make her comments.*

Jackie Jones is the director of initiatives in early childhood and literacy education of Educational Testing Service in Princeton, New Jersey. Prior to this position, Dr. Jones was a senior research scientist in research and development –

(Side 2 of Tape)

[Introduction]

MS. JONES: Allow me to weigh in on the topic of assessment. I live in this world in Princeton of folks who like to test things and like to measure things – (laughter) – because it's fun and interesting, and whenever they see that I'm going to do some kind of presentation or a seminar then they go, god, we've got to hear about these little children again, and that's hard. So I want to start this by telling you that the folks in Princeton who sit at their computers crunching numbers really think that the assessment of young children is a very difficult and complicated task, and I would like to be able to pose to you a range of fabulous models around early childhood assessment systems. We don't have a lot of good models of early childhood assessment systems, and what I'd like to do is to throw out a couple of questions that I think are the questions that people are asking and then talk a little bit about these systems and why we're having such difficulty trying to shape them and make them useful and effective.

[The Question of Assessment]

It seems to me that as we look at this question of assessment and keep talking about child assessment and programs and how are programs doing, how are we testing children, there are a couple of questions that might be at the heart of things. We have every right, and probably an obligation when we're talking about funded programs, to ask, are programs helping children to become literate? Are they working? Are they doing what they're supposed to be doing? On the other hand, the real question that's going to have an impact on children is, how will I teach Johnny to read? That's the question that's going to, if we get the right answers, have a tremendous

impact on individual children and therefore have an impact on that first question. It'll make the programs work.

Each of these questions requires evidence, and as we think about evidence, what will be compelling evidence that programs are doing what they're supposed to be doing, what will be compelling evidence that Johnny is learning to read, then we get ourselves into this world of assessment. And I want us to think large about assessment. And I say this every time I start one of these talks, that I'm not talking particularly about testing right now, though testing is a part of it, but I'm talking about how do we collect compelling evidence that we can use in an appropriate way to make inferences about programs and children?

[Program Evaluation]

So if we look at program evaluation -- are programs helping children to become literate -- one of the first real assessment issues is, what are the program goals? We have found – and I think you find this across programs – is that in many program evaluations you've got some evidence that doesn't always match what people think are the goals of the program. So if you have a comprehensive program, are you going to have measures that are comprehensive – as comprehensive as you think the goals of the program ought to be? The evidence of progress for a program: is it going to be limited to child outcomes? Well, child outcomes can be an important aspect of how a program is functioning, but if that's not the total scope of the program, then you need to look at other measures. So you see increasingly, in program evaluations, that people are trying to get a lot of contextual variables, a lot of background variables, trying to add to simply the child outcomes, other measures that tell you, in a large way, is this program effective?

Now, what do you get from a program evaluation? What do you get from doing the evil standardized test? You get a broad picture of how children are doing. If you've given that measure to a sample of children in the program – and you can do that – you get a broad picture of how kids are doing. You get some trends across time. You get some information that should help you with, as Fred has said, the goal of evaluation and assessment; that is, program improvement. If we can't use the information to improve programs, if we can't use the information to dig a little deeper

and find out what's going on in the programs, then it's hard to justify doing it.

So the goal is to find appropriate evidence, to collect it in a way that's useful, to report it in a way that's clear, and to be able to use it in a way that gives us information that we can use for program improvement. But there are challenges here, and as we talk about what we think we know about the development of literacy in young children, how do we get all of that in a measure that we can administer to young children? Can we ask the whole scope of what we know is appropriate for literacy development, or do we have to make some hard choices? Do we have to say, what will be the proxy that I use for literacy development in this particular measure? It's not the thing. The items are proxies for a particular task, a particular skill that we think is important. So once we agree that we've got some things that are important, how do we then find that task that we think will be compelling and administer it, realizing, again, that it is not the thing itself?

The question was raised about culture and language – and these are real questions and they're getting more real every day in this country, and we are grappling with that? Remember, if you have a large program evaluation and you're trying to find some items that you think will tap those core skills that are important, you've got to make some choices, and quite frankly, there are times when we come down to the least common denominator. What is that thing that we can use that will broadly address this question to the most children?

Marylou's question was also interesting because it really comes to what is the purpose of any kind of assessment; that is, making some kind of inference. So how are we using the results that we get from any kind of large program evaluation becomes a critical factor. What kinds of inferences are we making about programs, about children, about teachers? When I look back – and I always think about the Shepard, Kagan & Wurtz, the purposes of assessment. And there's that fourth purpose that we talked about: screening and informing instruction and evaluating programs and accountability. And I think when I first read that I always thought, accountability? Why would kids be held accountable? And now it's really hard to separate the program evaluation from the accountability because it's programs being held accountable, and that whole notion of what will this mean is really at the heart of it. So as much as we look the items and we want good technical aspects of reliability and validity, we're also looking at

what are the inferences, what are appropriate inferences that will be made from any kind of measure?

As I look at large program evaluation, you've got pilots that have to be designed. No one sits down in a year and designs the perfect assessment, not even SATs. You don't. You look at what will be reasonably good measures. You go out there and you do a pilot – I love pilots because you're always surprised. You're always surprised. You always get some group doing something you hadn't anticipated, some item acting strangely. I have to tell you that folks at ETS get rounded up periodically to look at responses from some of the big testing programs -- SATs or GREs – and there is always – "differential item functioning" is the name of this process, and so you have to figure out, why did a particular item get the kinds of responses by all kinds of different people that it got?

And sometimes you've got the data that says, here is group that acts this way when answering that item, here is a group that acts the other way, and you just can't figure it out. There's nothing from that sort of surface look at that item that would indicate that. That's why you have a pilot, to give you that data, to go back and collect it so that you can say, I don't see a problem here, but there clearly is and we have to do something probably to get rid of the item, or change it in some way.

So my point is that program evaluation and the design of assessments that will evaluate programs is an ongoing kind of task. It requires a lot of constant review of how items are functioning, how children are demonstrating their knowledge, and you can't necessarily assume that because we've got something that we've put together and it looks good, it's going to be the final version. So I think that that's my comment on program evaluation.

[Assessment Options]

However, if we think about assessment that really gets to the heart of individual children, then we look at what teachers are doing in classrooms and we look at the question, how will I teach Johnny to read? And there, I think, is an even more difficult array of questions because for that kind of assessment to be done, and done well, we're assuming that folks will be able to understand what they want children to know and be able to do, understand the normal course of that, be able to look at the evidence, or decide what

they think is compelling evidence that children are getting that, provide those experiences in their classrooms that will allow children to be able to grow and be able to achieve the tasks we want them to achieve, and be able to use whatever kind of evidence they have in a way that informs their instructional practice in a way that is productive for children.

Quite frankly, we don't see a lot of that going on, and as we talk about good classroom-based assessments – a lot of interesting work has been done by Rick Stiggins, and I looked in the winter issue of the NCMA's – National Council of Measurement in Education, and there are a lot of discussions on what does reliability and validity mean in the context of classroom-based assessment? And quite frankly, we're still working through it, because reliability in the context of a program evaluation, we think we've got that straight. We want interrater reliability; we want alternate forms to be equivalent. We're pretty clear about that. In the classroom you've got – that's what you have is context. You've got all those things that we think are problematic for program evaluation. They're out there in the classroom, and maybe they should be.

So if you look at our history of portfolio assessment and alternative assessments, there have been times when the country has been hot on this kind of thing. Whole states have decided that they wanted to engage in portfolio assessment and yet we come back and we look at the metrics and what's going on: well, you know, we're not too happy with the reliability; well, we're not really sure. And we haven't really done a tremendously good job as measurement folks or in teacher ed in getting everybody to understand what it means to have an assessment – a really good classroom-based assessment that looks at the context of the classroom and is bound by those goals that we think we want for kids.

How do we insure interrater reliability in the context of performance assessment? We're working with the state of New Jersey to design a performance assessment for three- and four-year-olds. Teachers are engaged in this task, and yet, how do we make sure that every teacher is going to be evaluating performance in the same way? So what we really believe, and I think it's becoming increasingly clear, is that for classroom-based assessment in early childhood in particular, we're looking at an enormous amount of professional development, of assessment-related professional development that hasn't been there in the teacher preparation programs, in

in-service. It hasn't been there. And so we're asking people to do new and different things to understand a different way of talking.

I remember having just a wonderful conversation with a staff developer in a district in New Jersey, and I said, we can't use this measure; it doesn't discriminate. And she was appalled. And I just – I can't use this. What good is it? I can't tell one group of kids from another. And it occurred to me that for her, discrimination is just – of course, nothing – you can't discriminate, and it's the discourse that we use that's even very different. I need a measure to be able to tell me how there are differences among people, not because of ethnicity or gender, but because of the knowledge.

And so, the very basis of language that we use can be really kind of serving cross-purposes. But I think that the times have changed, and I think that for this period in early childhood – because if there's ever been a time when the spotlight is on early childhood, it certainly is right now – there has to be a commitment from the measurement community to work a lot harder on these measurements of reliability and validity in performance assessment, in classroom-based assessment to look more carefully at some of the measures that Fred's been talking about, especially the self-regulation that is really important: how do we get good measures of self-regulation, how do we get good measures of social/emotional development. We're really not comfortable with those, but how do we work hard at those kinds of things?

How do we get good measures of vocabulary for folks for whom English is not their native language? How do we feel comfortable that we have a good estimate of that? And we don't have that right now. But I think if we think about a system – let me just pose this to you – that the question of how will I teach Johnny to read and how are programs helping children to become literate are not mutually exclusive questions. Those questions should be asked, and many more questions. I think parents have questions that they should be asking about their children's growth and development. Teachers have questions, administrators have questions, and funding agencies have questions.

So that means that a range of assessments need to be out there and appropriate but coherently aligned to each other and to an agreed-upon set of standards. And so, if we're going to be able to really get the real picture of children's performance, we need to be able to have screening measures that

make sense, that don't sort of over-represent minority kids in some way but really give us a sense of who's in trouble and how we can start doing further diagnostic evaluations. We need to have classroom-based assessments that really are reliable and valid and have teachers looking in a purposeful way about children's performance on an ongoing basis, because that's where the real evidence – and I've always said this, that the real evidence of children's learning lies in the classroom, if we can find ways to capture it and make it real. We need ways to report that classroom-based data that make sense when we have to abrogate that data maybe for accountability purposes sometimes, but it has to be reported in a reasonable way. We have to have people who have confidence in it, because I don't believe that at this point there is a lot of confidence in teacher observation and classroom-based data, and maybe for some good reason.

We also have to have accountability measures that are valid and reliable and as close as possible to authentic tasks. On the other hand, we have to understand the limitations of all of these measures, and so that the inferences that we make from accountability measures that may be using a sample of children don't necessarily tell us what to do with individual children unless we can combine that knowledge with the classroom-based evidence that we have.

So I think it's, in many ways, a difficult task – a much more difficult task when we look at young children, because the picture is very complicated. These are children who can't necessarily, or won't necessarily, engage in this sort of interactive testing agreement. You know, it's hard to get a three-year-old to say, yes, I'm really ready to be assessed right now, thank you. And that's fine, but that means the obligation is to look in an ongoing way at children's performance, in a careful way, and have an appropriate set of standards that we've designed.

So as far as assessment is concerned, it goes on all the time in every early childhood classroom. We just have to make it better and we have to make sure that as we're assessing programs that serve young children, that we really are looking at the kinds of skills that we think are important, that we're looking over time to make sure that they have some predictive validity, that they make some sense, and that means constantly reviewing that assessment, and then I think we have to make a huge commitment to assessment-related professional development, in teacher preparation programs, in in-service, throughout the spectrum of what we do for young

children. If the professionals in the field aren't what Stiggins calls "assessment literate," we can't argue the case and talk the talk, then we're going to have these constant battles of not understanding the data and really missing opportunities to define the evidence of learning for young children.

MS. BOWMAN: *All right, now you can work on Jackie a little while. (Chuckles.) Particular questions for Jackie?*

***Q*:** Hi. Heather Callister, Core Knowledge Foundation. Thank you very much for your comment about professional development vis-à-vis teachers understanding the process of assessment and the language of assessment. And I don't think this is a question; it's kind of a statement that hangs over my head throughout everything that all three of you – or four of you have been speaking about. You mentioned a lot of areas, skill areas important to early reading. I have worked in Head Start for a long time and in the childcare arena for a long time. A lot of the skill areas that you mentioned are skill areas that staff need development in.

So I appreciate what you're talking about in terms of areas that we need to be focusing our teaching upon. The challenge that I continue to have is this: that whether it's vocabulary or open-ended conversation or print awareness or familiarity with books, time and time again, in the staff development that I have and the curriculum development that I have, and getting to the whole arena and level of accurate, meaningful assessment, it's not so much the skills that we're trying to measure or even identifying the skills as they relate to this, the iceberg of – as you mention, the proxy, which is a good reflector of a much larger developmental domain, I come back to what the staff whom I work with can manage to do vis-à-vis those areas, and that has been a stumbling block for many years.

MS. JONES: Part of the work that we've been really interested in looking at is designing professional development models that can meet people where they are. What's fascinating to me always about early childhood is that you have folks who have high school diplomas, folks who have national board certification. The range is enormous. And I think that as you work in classrooms with teachers, what you have to do is to meet them where they are and to provide those folks who have limited skills, if you will, with just the right kinds of tasks to move them along as they're trying to develop their understanding of young children. For example, if your intent is to enhance someone's powers of observation and to be able to question children – and you see sort of scaffolding notions – then I think you provide those kinds of tasks in the everyday context of a classroom, and you have them do it, and you need somebody to move them along.

So I think it is possible to move folks along in a way – it is very difficult and it is slow, and that I think is the difficulty, that we have to have a lot of time and patience, but unless you change the workforce dramatically, you have what you have, and I think we are all facing the fact that in many cases, while we're asking that children are exposed to talk and talk, and lots of talk – there's talk and there's talk – and sometimes you're not getting the kind of talk that you want.

MS. SCHULMAN*:* I'm Karen Schulman, National Women's Law Center. You talked about how important the training is as part of professional development. I was wondering if you could be more specific from either estimates you've done or from examples of a model where it's worked well – like, how much resources it does take, how much time it does take, how many trainers it takes?

MS. JONES: Sure. You know, some of the work that Stiggins has done with school districts trying to develop what he really calls assessment literacy is work that he will start but he really says, you're going to go on and take this over and continue it. So it isn't just sort of a shot, it's developing a different culture – a culture of looking at evidence and talking about children's work and talking about whatever kind of assessment data you have. So it's giving administrators, as well as classroom teachers, this sort of notion that it is part of the everyday culture of that place to talk about whatever kind of assessment data they have.

In my own experiences – and with other folks who are trying to work with school districts setting up assessment systems, I've always said, don't ask us to come in if you're not going to be ready to deal with this for four years. It'll take that long to design the system, to start working with teachers, to get something that could move along, because new teachers come; they have to be mentored by more experienced teachers. You've going to have to keep institutionalizing this work.

So it's just a different way of functioning, a different way of looking at how you talk about children's progress and how people talk together in schools. But I'd say assessment system – give it four or five years.

Q: Toks Fashola -- hi, Jackie -- Johns Hopkins University and American Institutes for Research. You know, we started off with Fred discussing certain things that are clearly known to affect literacy and school success, and then we moved over to Jackie and in her area she talked about the fact that there were certain things that have not necessarily been developed, i.e. professional development. When is there going to be a concrete

conversation – and I'm a former early childhood educator and a former Heat Start administrator and teacher.

I remember the developmentally appropriate movement that took off really well. Basically people put their feet down and said, this is and this is not, and that was it. As we know these professional development things that are not working, that are not good, when is there going to be a serious, concrete conversation between the findings of what works and what does not work among preschool children? Specifically we've talked about low SES students, we've talked about African American students and the quality of the conversation that they bring to the school, i.e. school English. If you have teachers who are not fluent in school English, who are not doing what they're supposed to do because of qualifications, if there's something that's going to take place in the movement to say, since the children don't have it, it is incumbent upon the school and the classroom to give it to the children – not next week, not next year, but now.

Developmentally appropriate took off really, really well. Can that happen in literacy? Thank you.

MS. JONES: I guess I try to be a pretty optimistic person, and some people call me delusional but I try to be optimistic. (Laughter.) And we look at how do we know things are working? This is the what-works clearinghouse that IES has. And I really am going to try to be optimistic and think that one of the things that we need to do is to be pretty clear about how we know things. And so, if we can use the present atmosphere of evaluation of programs to really look carefully at those programs that we think are effective, I think we should make the most of it.

So I think that we should look at some of the work – the work that's coming out of IES and even though we may not like the notion of random assignment, we can still think about how are we going to get evidence that some of the things that we think are effective really are? So how do we design better studies that really do say, developmentally appropriate practice, defined in some way, really is better than something else? So from my perspective, I think it's a matter of trying really hard to get better evidence about what really does work with children.

Now, I have to tell you, I actually had an interesting conversation with Reid Lyon, who said that he agreed: professional development was an interesting and an important issue. And so I said, does that mean money? And he said, yes. So he said, yes. So there may be hope, but I think it is really becoming increasingly clear that unless we can help people to

understand the results of assessment, all the assessment that we're doing is not going to have any impact on student learning.

Q: Hi. Amy Berg, the Century Foundation. I guess a lot of the research I do is on teacher training and education programs, and when you talk about professional development, I think it always shocks me that in college courses, college preparation programs for teachers, there aren't course in research or design or assessment, and how can we expect teachers – I mean, I think it just starts earlier – it starts at the college level. How do we expect teachers to understand research and understand assessment if they don't have that background, and what can we do about that I guess is my question. I find it shocking that there aren't those kind of courses.

MS. JONES: You know, whenever we have conversations with the teacher ed folks – and there are people here that can speak to this better than I -- the comment is we've got 30 credits or we've got these constraints, and those are real constraints. On the other hand, I think we need to think about certificate programs, sort of advanced kinds of programs, an extra year, a special program – something -- as we're trying to incorporate assessment knowledge into the regular curriculum. I don't think it should be something that's, you know, okay, now you know about early children, now you're going to know about assessment. It's part of it -- it's part of teaching and I think that may happen, I don't know. But I think to make it even a special certificate program might be a way to get it started in a place that has constraints around credits and time.

Q: Jane Hannaway, the Urban Institute. I don't know anything about early childhood learning or teaching and therefore I have in front of me about a thousand different questions. And I've been very taken with the literature about the predictive power of performance at early grades on later achievement, so I'm, you know, really very seriously interested in understanding early learning and how to intervene in early learning.

Some of my questions – and if I can just throw a bunch of them out, you know, either of you can pick up on any ones you want. But I'm curious about what the correlation is between and among the different domains that Fred mentioned in the beginning. Are they in fact somewhat independent or are they indicators of some underlying factor or condition? And the reason I think that's important is because it may have implications for how to intervene. Are some -- let me just throw a couple out. Are some domains easier to affect than others? You know, I'm getting the impression that they sort of feed into each other. If some are easier to affect than others, then the payoff associated with one strategy may be better than the payoff associated with another strategy. Are some of these domains more powerful predictors of later achievement and some weaker predictors of later achievement? Again, that would give us some indications about where to intervene.

Then when we jump -- so I've got all these questions about early learning and then all of a sudden we're jumped into professional development and I'm like, whoa, you know, how can we possibly develop a professional development program that really makes sense and is well grounded without some understanding of the basic causal mechanisms associated with early learning? And then we jump from there into accountability and what works and then I'm thinking, wow, we need value added measures of something. How much grounding do we really have in understanding this so that we can move forward? I know that's a lot, but my head's spinning.

MR. MORRISON: Well, in some of these areas I think we know more than in others. I think it's pretty clear -- well, you're actually talking about research that's very ongoing at this point -- in other words I think we've identified these components, but probably one of the most important unanswered questions right now is how do these fit together dynamically across age?

So there are several studies that have come out in the last year and a half or couple of years using these large modeling techniques where they will take measures on kids of vocabulary, phonological awareness, oral language as somewhat separate from vocabulary, print knowledge, and then they'll look at word decoding and comprehension. And they'll start that process -- in the case of Russ Whitehurst and some of his work with Storch (sp) and Lonigan -- start at three years of age, and they'll test kids. So the issue is, well what predicts when? What's important? And it's not absolutely agreed upon, but the sense we're getting is that early on the components are associated with each other -- point three maybe point four correlations between, say, vocabulary and phonological awareness -- but that if you do the appropriate statistical manipulations like forced entry regressions or structural equation modeling, they each predict independent amounts of variance up to kindergarten.

Now, at kindergarten, the bulk of the evidence, except for one study which is in press – (chuckles) – says that the key skill for predicting first grade and second grade word decoding is phonemic awareness. Vocabulary no longer uniquely or directly predicts, but if you look statistically, it goes through phonological awareness. Does this make sense? This is the dynamic that people are working on. However, if you look at third grade reading comprehension, three- and four-year-old vocabulary uniquely predicts that. So it's like it surfaces. So the question of what do you do -- what should I be doing with my preschoolers? Well, you should be doing vocabulary even if the people who say it's phonemic awareness that's critical at four and five years of age, well, you've still -- the importance of

vocabulary is still there because it's likely to influence listening comprehension, core knowledge, and eventually comprehension.

It's exciting for researchers to try to put this together, and for once we have the statistical tools to do it. And so I'm not sure if that actually answers all the questions. But, so, are some more important than others? Well, yes and no. At some points in time, some appear to be more important -- so phonemic awareness -- but everything seems to play a role in different ways and there's this leapfrog effect that we haven't really appreciated. And syntax is another issue, that syntax, until some recent data, didn't seem to predict early word decoding as much, but we do know early syntax will predict comprehension later on. It's really fun, yeah. (Laughter.)

Well, and it also, I do think that it probably has some real practical value, I think in terms of what we do. We sort his out. But this is a journal of ed-psych and developmental psychology, journals like that, if you read them, the recent research.

MS. JONES: You said there was this leap to professional development, and I think the leap came from really looking at assessment systems, trying to figure out how we help people interpret information, and finding that there needed to be a lot more work in that. So I think there's a huge amount of work to be done in understanding how young children develop literacy and then in understanding those indicators that we would use to see how children are progressing. So it's not so much a leap as I think it is an integral piece of the picture -- that we have to all of that.

Q: My name is Jerry Sroufe and I work for the American Educational Research Association and currently with the National Council for Measurement in Education as well. And I just wanted to make an observation and raise a question.

The observation is, my experience has been that people who are psychometricians and design assessments are much more modest about their abilities to do this in a compelling and useful way than are policymakers and others who rely on them. And so I agree with what I thought Jackie was saying earlier.

My question is, it seemed to me when we had the discussion of the variables that make an impact, that many of them are largely determined or influenced by parental experience, and I thought maybe portions at least of all four that were identified, and yet we shifted from that to what happens in the school and what happens in professional development, and it would seem, looking at the policy aspect, that you would want to look at the place where you might have the most opportunity to intervene earliest, and

those things are susceptible to intervention through some of our programs. And particularly, several people raise the idea of cultural impact, and it seemed to me when Tom suggested, well, we had some exams in Spanish now, but the language is the least important part of the parental preparation in terms of the variables that you discussed. Thank you.

MR. MORRISON: Yeah, I'm really glad that you raised that; it's something that I have talked about a lot. I have a couple of boxes here that I thought I was going to try to emphasize if I had a minute, namely that one of the purposes of assessment under any circumstances I think is to inform parents and involve them and that we should be communicating with parents as much as we can.

In several pieces of research that we and others have done – most recently the NICHD study of early childcare; we published a paper in the Harvard Ed Review – which actually looked at the differential prediction of child factors, family factors, preschool factors in the larger socio-cultural context on children's first grade reading and math behavior. And while it's controversial to compare across domains, the parental context was about three or four times more powerful than the preschool context. So not being able to make causal statements or whatever, but I'd emphasis that myself, that I think the larger context of all this really has to bring parenting into the picture.

Q: John Jackson, NAACP. Up until this point I think when we -- I've been assuming when we talk about the desired outcome or predicting success we are talking about what are those factors that will lead a student to our desired outcome? And my challenge has been that we've approached in such a standardized way, but when we began to talk about the parents that are actually really preparing their children for preschool, it's not standardized. Has there been any research out there that discusses what are those components that predict success in students or in environments where the students are not coming into the environment with the vocabulary necessary, with the language necessary, because we've had instances where even those students have reached the point of that desired outcome but they didn't follow this linear, standardized map that we're outlining today.

MR. MORRISON: Well, I think there isn't nearly enough research on that. I have a student in my lab now, in fact, who is attempting to look at that in a more systematic way. The kinds of parenting dimensions that people talk about are the learning environment, warmth responsivity, and controlled discipline. And the only research that I'm aware of that seems to show what you're calling kind of, you know, a different pathway, if you will, is the

research looking at the differential effect of what's been referred to as authoritative versus authoritarian parenting; that in a sample of at-risk African American families, authoritarian parenting was more associated with successful outcomes. That's also true in a sample of Asian parents, who, by standard definitions, were viewed as more authoritarian.

So I think that's all that I'm aware of in terms of what these dimensions are and how they could differentially impact in a different culture.

Q*:* (Comments off mike) -- research from 25 years ago in case studies of families that are making it and families that are not -- African American families -- and his research seemed to suggest that parents who were warm and yet demanding and authoritative were the ones who had the children who were most successful in school.

Q: I mentioned parents, but not just from a parent's standpoint. What is the teacher to do with a student who comes from a background that doesn't follow the standard form of the teacher that has the student who doesn't have the school's English? What can be done – is there a body of research that shows us what can be done to increase that student's likelihood of success, or what can be done to -- because there are students each and every day that succeed that don't enter school with these components, and what actually occurred to help them reach that desired outcome?

MS. JONES: I can't give you a body of research to answer your question directly. What I can speak to you from is the sort of perspective of teachers looking more carefully at what students know from a classroom perspective. And I think if we can help teachers – what we've seen is that as you look at the evidence of student learning in a very systematic way, with teachers who've been trained to observe and have standards and they know what they're doing, they find that children are doing a lot more than they ever imagined they were doing, and so some of the work of Black & William sort of indicates that with good training of teachers, early childhood teachers, to show them how to observe, how to get the evidence of children's learning, they're finding that while they thought, well, this kid doesn't know how to do this, that there are many more skills and abilities that those children have than they ever imagined.

So it's a matter, in many cases, of helping people to understand what kids know, what they're really able to do, and to see that more clearly and then provide the early childhood programs that we think are appropriate and teach –

MS. BOWMAN: *The other piece of that is that increased education of the teacher does lead toward the teacher doing more of the kinds of things that get beneficial results for children. So that teachers who go further in school do more of the good things that are good for children.*

Q: My question is actually – I'm Carol Brunson Day, and my question is in the same vein as the question that was just asked, and it has to do with how we can organize ourselves to look at how to overcome the obstacles that are being predicted – the failure that is also being predicted. We are talking about the assessment and its efficacy in predicting success, but as it becomes better at predicting success it also becomes better at predicting who's going to fail.

And when I think about what we also know about expectations and so forth, that there are examples of children who were predicted to fail who indeed did not fail. We don't spend a lot of time really looking at they dynamics of those situations, but I hear you suggesting that because of the limitations of assessment we have lots of challenges before us, and I guess I would ask if you would speak a little bit more to this question of our abilities to predict failures and how that impacts children as well – children's futures – and how we can become better advocates in the field for the belief that kids don't have to fail if they get good instructional practices or the benefits of adults who believe they can succeed.

MS. JONES: I want to go back to this notion of what we get from more standardized measures that sort of – as I called it, the sort of least common denominator tasks, if you will, and what kind of belief system we put into that. If we're building a system of assessments that we think are appropriate for young children, then we're going to get some measures that for some children will indicate that they stand a really good chance of not succeeding. But think about the inferences that we're making from that. Does that mean that we say there's nothing we can do? I think what that means is, this child is at risk, now how do I provide the best kinds of opportunities to make sure that this doesn't happen?

There is a kind of what we call the power of the score thing that goes on, and it doesn't mean this is this child's destiny; it means that there are a set of circumstances in which this child finds itself, but the business of teaching seems to me to provide those opportunities for children that will enhance their ability to move beyond that. It's not preordained.

MS. BOWMAN: *If anybody feels the need to get up and leave, please – I promised it would be over at 2:00 and it's not, but I'm going to let Fred finish and anybody who would like to leave, don't feel uncomfortable.*

MR. MORRISON: Thanks. Yeah, this is actually a major focus of our research right now. I mean, I think we can predict who's going to fail. But one of the things that we've been doing is going into classrooms. When I was in Chicago and we worked in Evanston we did observations in classrooms in first through third grade and we started doing observations in preschool classrooms in a district just outside of Detroit, and what we have found is that essentially if you look at sort of the dimensions of instruction that are going on – and we've tried to find sort of a classification of, well, what are teachers doing? How can we try to sort of categorize what's going on?

And so we've done a simple sort of classification where we ask, is the instruction, say, teacher-managed versus child-managed, or is the instruction in word decoding explicitly focused on word decoding or implicitly by having the children do sustained silent reading?

The point of it is that, one, we have found that for children who start out first grade with very low decoding skills, the more of this teacher-managed explicit instruction they get, the nuts and bolts, the better they do. But not all kids get it. So I think the exciting part is that we have identified the amount and type of instruction they need in order to make a difference. And it also changes over the school year. And we've just recently found a very similar pattern of what we're calling "child by instruction interactions" in the preschool sample.

So we've been arguing for more of an individualized instruction focus where you look at – where assessment obviously is absolutely critical for the teacher, and you find out exactly where a child is at the beginning of the year. But based upon the research on what works, if you will, in terms of these dimensions of instruction, we've been able to actually prescribe amounts and types of instruction over a week or month period. And a colleague of mine, Carol Connor, has an IES grant to actually look at the impact of an intervention based on this.

So now only do I think we have a sense of how to improve, but I think I really think that we can actually do it. So I'm very excited about that.

MS. BOWMAN: *I'd like to thank you all for coming – again thank the Shanker Institute for sponsoring it, and I think our speakers will be a little while longer if you want to come and speak to them directly yourself. Again, thank you for coming.*

(Applause.)

(END)