<p style="text-align:center">**Performance Evaluation Advisory Council (PEAC)**</p>

<p style="text-align:center">**October 19, 2012**</p>

<p style="text-align:center">**Meeting Minutes**
Illinois State University Alumni Center, Normal, Illinois</p>

Members Present
Dr. Bette Bergeron, Ben Boer, Dawn Conway, Dr. Gail Fahey, Steven Isoye, Joseph Matula, Steve Ponisciak, Audry Soglin, Vicki Phillips, Dr. Linda Tomlinson, Dr. Diane Rutledge, Jodi Scott, Donald Daily, Kristen Adams, Kurt Hilgendorf, Hector Garcia, Randy Davis, Matthew Lyons, Stephanie Bernoteit, Michelle Standridge, and Angela Chamness.

Observers Present
Brad White (Illinois Education Research Council); Amy Aslop (Illinois Federation of Teachers); Jack Janezic (Illinois Education Association); Christi Chadwick (Illinois State Board of Education), Larry Stanton, Cassie Meyer, Nick Pinchok, and Jenni Fetters (American Institutes for Research); and Margery Yeager (Education Counsel, by phone).

Meeting Called to Order
The meeting was called to order at 10:04 a.m. by Dr. Diane Rutledge.

I. Welcome, Announcements, and Updates
- Membership updates: Randy Davis, Matthew Lyons, Kurt Hilgendorf, and Stephanie Bernoteit have been appointed as a full members.

- Training updates from Growth Through Learning: Mary Jane Morris provided PEAC members with a written update on the evaluator training program. Vicki Phillips noted that as of Monday, only 32 people have exhausted all attempts to pass their certification tests. Overall, the pass rates continue to improve and suggest the training process is contributing to a consistent level of competency across all evaluators.

II. Approval of the September Minutes
Joseph Matula made a motion to approve the minutes from the September meeting as written. This was seconded by Dawn Conway and approved by the council.

III. Topics in Stage 1: Introduce

> **A. Student Growth Model for teachers—**Guidance related to Student Growth Model (Model, Guidance), Assessments for teacher types, Type I and Type II assessment options (Model, Guidance)
>
> Nick Pinchok (AIR) facilitated PEAC's discussion of student growth for Type I and II assessments for teachers. The objectives for the discussion were to (1) identify interim or benchmark assessments that would be appropriate Type I and II assessments and (2) identify additional information and next steps for making recommendations. After briefly reviewing the relevant legislative, rule, and policy connections, Mr. Pinchok provided the council with a formal definition for interim and benchmark assessments and laid out a set of questions for the council to consider in selecting an assessment of student growth:
> - Is the assessment aligned with Illinois Standards and the Common Core?

- Is the assessment widely used in Illinois districts?
- Can the assessment be administered with a pre- and posttest multiple times during the year?
- Does the assessment measure and report on student growth in a clear and transparent way?
- What grades and subject areas does it cover?
- Is the cost of administering the assessment reasonable?

Next, Mr. Pinchok presented the results of the Illinois State Board of Education (ISBE) Assessment Survey and reported that 375 districts responded to the survey. Mr. Pinchok's presentation provided PEAC members with the five most frequently reported assessments and their frequency of administration in K–8 reading and mathematics, and 9–12 reading and mathematics (for detailed results, see the accompanying slide presentation). In addition, AIR reached out to the publishers of tests commonly used in Illinois and requested information on the number of Illinois districts that had purchased various types of assessments. Mr. Pinchok presented the responses received from AIMSweb (700 districts or other customers), DiscoveryED/Thinklink (150 districts), Scantron (17 districts), and Northwest Evaluation Association (NWEA) Measures of Academic Progress (MAP) (over 300 districts). To assist the council in narrowing their assessment choices, Mr. Pinchok provided a table comparing the four most widely used assessment types (AIMSweb, Discovery, Scantron, and NWEA MAP) on the questions he had recommended for consideration.

Council members offered a number of questions and discussion points. One member requested that the summary table include information on whether the tests were computer adaptive (CAT). The member also wished to know whether using different assessment questions in the pre- and posttests presented a problem for measuring student growth. Steve Ponisciak indicated that in most cases, although different questions are used, they are considered comparable because they measure the same underlying constructs. Mr. Ponisciak stated that he strongly advised PEAC to recommend CAT assessments in order to avoid floor and ceiling effects that are associated with pencil-and-paper tests, which measure a narrower breadth of knowledge and skills. Another member noted that, although CAT assessments were preferable, the council needed to be cognizant of the infrastructure requirements of CAT assessments and the barriers they may present for some schools in Illinois.

One member asked how student growth was measured using a CAT assessment. Mr. Ponisciak explained that CAT assessments rely on historical growth data from fall to spring and require accurate scaling to ensure an accurate measure of student growth. The member requested clarification on how the CAT assessment scales are normed. Another member reported that NWEA and Scantron both use a national sample of students for each band of scores to produce a median growth score for each band, but NWEA has provided a different student sample for Chicago Public Schools that more closely matches their district. The council concluded that they would prefer to have more information on how student growth would be calculated or the measurement model that was possible for each assessment as well as more information on the reliability and validity of the tests before narrowing the list of possible assessments.

One member also noted that one district that tried to use MAP for measuring student growth ran into difficulties when they attempted to adjust the score calculations. The process became very difficult to understand, such that only the middle school mathematics teachers really had

a clear understanding of how their scores were calculated. The member suggested that the committee should also consider how easily the assessments and measurement models can be modified or adjusted and how easily they can be understood by educators and the wider community.

One member noted that AIMSweb was currently used in her district's cooperative and was well liked, particularly because of the charts used to report results. She noted the assessment was being used in their high school grades and was widely known and understood by both educators and parents. Another member raised concerns that AIMSweb was not explicitly aligned with the Common Core and the testing company did not have plans to do so in the future. Other council members requested clarification on the administration requirements for AIMSweb because several council members noted that different assessments may require different approaches ranging from 1-on-1 to small-group administration.

Vicki Phillips noted that she recently spoke with an educator who filled out the ISBE Assessment Survey and was concerned that PEAC might choose to recommend an assessment that is widely used but is not the best option. The educator expressed a strong preference for Discovery/ThinkLink. Several council members concurred that PEAC needed to consider all the best assessment options, even if they are not currently widely used in Illinois. Members requested that additional assessment options be added to the list, drawing on requests for proposals in Chicago, New York, and other states.

PEAC members also raised questions about whether PEAC should be recommending a single assessment for each teacher category or whether PEAC should provide a menu of options. Vicki Phillips noted that for the purposes of the state model, a single assessment would have to be identified for each category because the state default model will kick in for districts unable to reach an agreement between district leadership and the union. Providing a menu in these cases would not provide a clear assessment and measurement model for those districts. Other council members raised concerns that offering a menu of assessments might slow down the guidance and rule-making process when it was critical that PEAC offer districts guidance as soon as possible.

PEAC members concluded the discussion by asking that several criteria be added to the comparison table:
- Administration information (length of time required, technology or infrastructure required, efficiency)
- Technical information on growth calculation for the assessments (e.g., how they are normed or how the percentile scales are determined)
- Level of Common Core State Standards alignment
- District adaptation options
- Ease of calculation and use

Larry Stanton noted that the council ran out of time to discuss assessment options for the high school grades, but in light of the ISBE survey results, he asked whether the council would be okay with pursuing information on the ACT Educational Planning and Assessment System for discussion at the next meeting. The council concurred.

Next Steps: AIR will (1) revise the comparison table to include the criteria requested by the members and locate this information for each assessment; (2) request information or presentations from either experts, testing company representatives, or districts currently using the assessments to provide PEAC with more information on the technical quality of the assessments and possible measurement models for each assessment; (3) research other assessments to ensure the list includes the best possible assessments, even if they are not widely used in Illinois; (4) provide information on EPAS as an option for high school grades and subjects.

III. Topics in Stage 1 (cont.)

> **Guidance**

B. Operating Guidelines for Student Growth (Model, Guidance)

Larry Stanton provided an overview of key areas for consideration for PEAC in creating guidance on operational rules for student growth measures. Specifically, the proposed areas for consideration that were discussed were

- Attributing students to particular teachers
- Excluding student data
- Linking student growth to specific subjects or courses
- Considering prior achievement

One member requested clarification on how the operational rules related to last month's discussion on the percentage of students included in a student learning objective (SLO) with a Type III assessment. Mr. Stanton clarified that the operating guidelines were a continuation of that discussion but also included student growth for Type I and II assessments. Council members noted that the central tension with this particular topic area is (1) avoiding situations where, without meaning to, the rules make it excusable to leave more challenging students out of student growth measures and (2) holding teachers accountable for things they actually can influence (e.g., it is difficult to affect student learning if the student is not in class regularly). One member noted that if PEAC were going to establish cut-scores for student attendance, they might consider giving teachers credit for working more intensely with students who are infrequent attendees.

Council members suggested that the next discussion include how to handle teachers who have accepted preservice teachers, how to handle multilevel classrooms, and how to address teachers who have multiple preparations.

Council members also requested that PEAC's work on this topic be coordinated with Peter Goddard's state-level work on the state longitudinal data system. ISBE representatives noted that they were in communication with the Goddard group and would ensure continual collaboration on the topic. One council member noted that the data systems being addressed by the two groups were quite different in that the attribution and roster verification processes were more straightforward at the district level than they are in the state-level data systems being addressed by the longitudinal data group.

AIR will add the topics suggested by the committee to the areas for consideration and return with a detailed presentation and set of options for PEAC to react to at the November meeting.

III. Topics in Stage 1: Introduce (cont.)

> Guidance

C. Growth for First-Year Principals and Assistant Principals

Jenni Fetters provided the PEAC committee with a handout and overview summarizing a national scan that AIR completed on evaluation policies for first-year principals and assistant principals. The scan revealed that few states differentiate new and experienced principals with respect to how student growth is included in an evaluation. Many states differentiate these two statuses in other aspects of the evaluation, such as formative feedback cycles, observation or visitation frequency, and summative ratings. The overview found that a few states had exempted first-year principals from a summative evaluation altogether, or linked induction and a tiered certification system with a formative evaluation for first-year principals. Similarly, for assistant principals, very few states differentiated their evaluation process explicitly. DC Impact is the only system that has a fully differentiated evaluation process for assistant principals—the key modification for student growth involves the assistant principal setting a student achievement goal that is aligned with their professional scope of work. The handout provided several options for PEAC to consider in making recommendations on this topic.

One council member asked about the possibility of making adjustments to the rules for first-year principals. Vicki Phillips noted that it is not possible to exempt first-year principals but it would be possible to consider how to give better rules for assistant principals. Another council member noted that some districts have already struggled with this topic and made decisions about it on the basis of the existing rules—with some schools taking very different approaches. Other members noted they were aware of several other districts that had information as well. PEAC requested that AIR collect information from as many of these districts as possible to inform the council's discussion at the next meeting.

AIR will follow up with council members to locate districts with relevant information on first-year and assistant principal evaluation designs and identify areas where existing rules might need to change to give more flexibility. PEAC members were encouraged to send additional thoughts, feedback, and ideas on the topic to Jenni Fetters.

V. Topics in Stage 2: Consider

A. Summative Rating Definitions (Model)

Larry Stanton provided a brief presentation on three summative rating approaches for PEAC to consider for the state model. Mr. Stanton noted that a decision to select a summative rating approach rests on several assumptions: (1) teachers are receiving feedback at multiple points during the evaluation cycle;

(2) teacher performance should be rated only on a representative, balanced sample of evidence from multiple sources; (3) the purpose is not to eliminate professional judgment from the rating process but to create a consistent, transparent approach that structures judgment appropriately. The presentation included three options for summative ratings: (1) a numerical approach (weighted average), (2) a profile approach (decision rules or matrix), and (3) a holistic approach (professional judgment based on evidence). Mr. Stanton reviewed the advantages and disadvantages of each option and provided examples of each. PEAC was then asked to consider four questions:

1. Which approach should we adopt for the state model?
2. How many ratings should there be for student growth? Are they required at each level of student growth rating?
3. If there are four ratings, should they be the same as the summative ratings?
4. What's the anticipated distribution of ratings? How difficult should it be to receive an unsatisfactory or excellent rating?

One council member raised a question about how a profile approach fits with the legislative requirements for the state model, specifically the 50 percent weighting requirements for practice and student growth. Another council member asked how the low, medium, and high growth ratings were arrived at in the profile approach. Several council members expressed a preference for the profile approach over a numerical approach because it avoids reducing teacher performance to a single score (which lends itself to ranking) and because it leaves more room for professional judgment. Two council members noted that they preferred the numerical approach because it provided more information about how a teacher has performed over time and provided a more detailed scale of performance. A profile approach can result in information loss by placing all teachers in four categories. Steve Ponisciak noted that from a measurement perspective, it would be better to have a five-category approach for the profile method. The four-category approach is problematic because it is difficult to know how far apart teachers are within the four buckets (e.g., a teacher who lands in the low growth/unsatisfactory box could be right on the border with moderate growth or at the very bottom). The challenge with a five-level approach is that SLOs data may not have enough information to discriminate among five levels. Noting these views and the limitations raised, the council agreed that there appeared to be a consensus on a profile approach, possibly with four levels.

One council member noted that a profile approach already had been adopted in the principal evaluation model and suggested the council include it as one possible option to model the teacher profile after in order to create alignment and consistency between the models. Two council members wanted to also consider whether certain things in the teacher practice model should be weighted more heavily in the profile, noting that this had been done in the principal model's profile. One council member noted that it might be possible to design an approach that takes into account how a teacher scores over multiple years. For example, if a teacher scores very low for two years in a row, in year three a different summative profile might be used that details different outcomes.

In addition, several council members passed out suggested profile designs for PEAC consideration. Council members were encouraged to send any suggested profile designs to AIR for consideration at the November meeting.

<u>**Next Steps:**</u>   AIR will compile suggested profile approaches and develop a set of options for PEAC to decide from among at the November meeting. PEAC members are to send suggested profiles to AIR.

VI. Topics in Stage 3: Decide

A. SLOs for Type III Assessments (Model, Guidance)

Council members were provided with a summary of the state model requirements discussed at the September meeting and asked whether it accurately reflected the PEAC's final recommendations. Council members noted that the requirements for weighting were incorrect—the SLO weights should be determined by the district, not by the teacher and evaluators. AIR agreed to correct the error. AIR noted that, in their review of the proposed requirements, the state model does not currently require an approval process for the SLOs and that this was significantly different from other states and districts across the country that typically require at least an administrator or evaluator review and approval of the SLOs to ensure quality and comparability of rigor and attainability. Council members noted that the purpose of Type III assessments and SLOs was to provide teachers, who know the classroom and students best, with more control over how student growth is assessed. A core concern was that by establishing an approval process at the district level, reviewers would not have the knowledge or understanding of the specific classroom context necessary to evaluate the SLOs. In addition, it would likely be a significant burden on district staff members, as well as teachers and administrators. Other council members noted that a building-level approach to approval might be more appropriate, but concerns over burden remained. Another member noted that the review and approval process might be a good place for teacher leaders to take a new role.

Council members agreed to approve the SLO recommendations (with the SLO weighting correction) and asked that AIR provide several options at the November meeting for including a recommendation on an SLO approval process.

Because of time constraints, council members were asked to provide AIR with feedback on the draft SLO guidebook either after the meeting or via e-mail.

<u>**Next Steps:**</u>   AIR will revise the documents to reflect the correction to the weighting language and will send council members revised versions of both documents for the November meeting. In addition, AIR will prepare several options for approval processes for PEAC to consider.

VI. Topics in Stage 3: Decide (cont.)

B. Teacher Evaluation—Nonobservable components (Guidance)

Cassie Meyer asked the council for feedback and comments on the updated

guidance document for the Danielson domains that are not assessed through observation. One council member requested that the language be revised to more directly reference district and union collaboration on the decision points raised in the document. Another member requested that the reference to Danielson resources be removed because it would require districts to purchase the materials. Instead, the text should say there are extensive resources online.

One member asked whether teacher self-assessment or self-report should be added as a possible measure. Council members noted that self-reflection and self-assessment are expected parts of teacher practice in all domains on the rubric, not only for the nonobservable domains. Noting it only on these domains could send the wrong message. Council members noted that it will be important to include self-reflection throughout the state model for teacher evaluation, perhaps by highlighting its role in pre- and postobservation conferences.

Next Steps: AIR will revise the language in the document to reflect the feedback PEAC provided.

VII. Topics in Stage 4: Give Guidance

A. Guidance on the Use of Survey Data in Principal Evaluation 2.0
One council member asked that the introduction to the document note that all schools will be required to implement the survey. Another asked that information be included on whether districts could run additional surveys alongside the 5Essentials survey. Larry Stanton noted that the scope of this document was quite narrow and that UIC had already developed more extensive communication documents that addressed these questions. Several council members noted that the language stating that it could not be used as part of principal evaluation did not accurately reflect the council's August discussion, which had focused on ensuring that districts did not use the 5Essentials survey as an *outcomes* measure in principal evaluation. Another member noted that the rules explicitly allow school climate surveys to be used in principal evaluation. Council members requested that the language in the August meeting notes be incorporated into the revised version.

Next Steps: AIR will revise the language in the document to reflect the feedback PEAC provided.

VII. Topics in Stage 4: Give Guidance (cont.)

B. District Decision-Making Guide for Teacher Evaluation
Several council members noted that in the first section on communication the language should be revised to focus on the district and union collaborating and communicating jointly and the reference should be to the joint committee rather than the district alone in order to highlight the importance of labor management collaboration. Council members also noted that the document needs to align with other planning resources developed by other organizations in the state. Cassie Meyer indicated that AIR reviewed those documents and checked for alignment but welcomed addition review from PEAC members to ensure the alignment was

as strong as possible. One council recommended that an explicit reference or link to the appropriate rules or legislation be included alongside each topic area so districts can quickly and easily locate the related language.

Next Steps: AIR will revise the language in the document to reflect the feedback PEAC provided.

The next meeting will take place November 9, 2012, at the Alumni Center in Normal, Illinois.

VII. Public Comment

Amy Aslop (Illinois Federation of Teachers) reminded the council that it was important, as topics emerge that are more complex and challenging than originally anticipated in the timeline, that the council take sufficient time to fully consider the manner.

Jack Janezic (Illinois Education Association) thanked the council for their robust discussion.

VIII. Adjournment:

The meeting was adjourned at 2:45 p.m.

SYMBOL KEY