SECTION III-2

ASSESSMENT SYSTEM ARCHITECTURE: Design features needed in the structure and operation of an early literacy assessment system

This chapter considers what it means to have a balanced, well-functioning assessment system in terms of (a) fundamentals of literacy assessment, (b) system architecture and design principles, and (c) steps that need to be taken to actually plan for and design such a system. The content provides some of the relevant explanation and backing for ***Principle #2*** *and associated* ***Phase I Planning and Design Recommendations.***

Phase I RECOMMENDATIONS (Principle # 2)
Principle #2: The ELAS must reflect ASSESSMENT SYSTEM DESIGN FEATURES that make it coherent, comprehensive, and continuous across time and contexts of use.
The **ELAS LEADERSHIP TEAM** should:
1.4: Develop and adopt a logic model and theory of action for the structure, functioning, and evaluation of the proposed ELAS.
1.5: Identify the educational decisions to be made, assessment information needed to support those decisions, and the stakeholder(s) who will be making the decision(s).
1.6: Construct a framework for the ELAS that includes clearly articulated relationships among the assessment tools and practices relative to a model of competency development in reading, writing, speaking, or listening.
1.7: Use the framework to conduct an audit of all existing district- and school-level assessment tools and practices currently in use to determine whether they meet criteria for inclusion and should remain a part of the system.

Introduction

The seminal publication *Knowing What Students Know: The Science and Design of Educational Assessment* (Pellegrino, Chudowsky & Glaser, 2001) crystalized the call for balanced systems of assessment:

> Assessments at all levels—from classroom to state—will work together in a system that is comprehensive, coherent, and continuous. In such a system, assessments would provide a variety of evidence to support educational decision making. Assessment at all levels would be linked back to the same underlying model of student learning and would provide indications of student growth over time (Pellegrino et al., 2001, p. 9).

Many authors since have helped advance this conceptualization of assessment systems, as well as an understanding of what constitutes a well-functioning system (e.g., Chattergoon & Marion, 2016; Conley, 2014; Council of Chief State School

Officers [CCSSO], 2015; Darling-Hammond et al., 2013; Pellegrino et al., 2014). While much has been learned about designing and implementing high-quality assessment systems over the past 20 years, there are few examples of well-functioning systems, particularly systems incorporating state summative tests and assessments at other levels of the system (e.g., district, classroom). Despite recent efforts to articulate principles of assessment systems (Deeper Learning 4 All, 2018), creating a balanced assessment system remains challenging and finding high-quality examples in practice is very rare (see e.g., Conley, 2018).

In planning for and designing an early literacy assessment systems (ELAS) for students, it will be important for a district's ELAS Leadership Team to leverage what has been learned about three things:

1. the nature of assessment
2. the assessment of literacy
3. the principles of assessment system architecture

This body of knowledge should inform an agenda for thoughtful design of an ELAS that can enhance equitable learning and life opportunities for all students. In this section we first review key conceptual issues regarding the nature of assessment, since these issues are foundational for understanding the broader principles for system design and implementation. We conclude the chapter with a discussion of the need for development of a theory of action for the assessment system and the use of a logic model to help uncover that theory of action and guide the process of system design, implementation, monitoring, and evaluation.

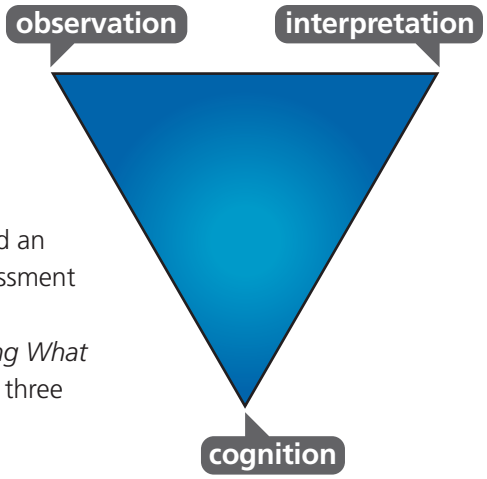
Assessment as a process of evidentiary reasoning

We assess students to make judgments about what they know and can do, but assessment does not offer a direct pipeline into a student's mind. Assessing educational outcomes for children is not as straightforward as measuring height or weight; the attributes to be measured are mental representations and processes that are not outwardly visible. Thus, an assessment is a tool designed to observe students' behavior and produce data that can be used to draw reasonable inferences about what students know and can do. Deciding what to assess and how to do so is not as simple as it might appear.

The process of collecting evidence to support inferences about what students know represents a chain of reasoning from evidence about student development and learning that characterizes all assessment, from classroom quizzes and standardized achievement tests to the conversation a student has with their teacher as they read a story or work through the meaning of a text.

The first question in the assessment reasoning process is "evidence about what?" Data become evidence in an analytic situation only when one has established their relevance to some question or concern. Data do not provide their own meaning; their value as evidence can arise only through some interpretational framework. In the present context, educational assessment provides data such as spoken or written work, marks on answer sheets, or students' explanations of their thinking. These data become evidence only with respect to understandings about how students acquire knowledge and skill.

In the *Knowing What Students Know* report, the process of reasoning from evidence was portrayed as a triad of three interconnected elements, forming an "assessment triangle." The vertices of the assessment triangle (see **Figure III.2.1**) represent the three key elements underlying any assessment: a model of student **cognition** and learning in the domain of the assessment; a set of assumptions and principles about the kinds of **observations** that will provide valid evidence of students' competencies; and an **interpretation** process for making sense of the evidence in light of the assessment purpose. The three are represented as vertices of a triangle because each is connected to and dependent on the other two. A major tenet of the *Knowing What Students Know* report is that for an assessment to be effective and valid, the three elements must be in synchrony.



● **FIGURE III.2.1**
The Assessment Triangle
Source: *Knowing What Students Know* (Pellegrino et al., 2001)

Cognition

The *cognition* corner of the triangle refers to theory, data, and practice about how students represent knowledge and develop competence in a domain (e.g., reading, writing, speaking, or listening). In any particular assessment application, a theory of development and learning in the domain is needed to identify the set of knowledge and skills that is important to assess for the intended context of use, whether that be to characterize the competencies students have acquired at some point in time to make a summative judgment, or to make formative judgments to guide subsequent instruction so as to maximize learning. A central premise is that the cognitive theory should represent the most scientifically credible understanding of typical ways in which learners represent knowledge and develop expertise in the domain of interest.

Observation

Every assessment is also based on a set of assumptions and principles about the kinds of tasks or situations that will prompt students to say, do, or create something that demonstrates important knowledge and skills. The tasks to which students are asked to respond on an assessment are not arbitrary. They must be carefully designed to provide evidence that is linked to the cognitive model of learning and to support the kinds of inferences and decisions that will be made on the basis of the assessment results.

The *observation* vertex of the assessment triangle represents a description or set of specifications for assessment tasks that will elicit illuminating responses from students. In assessment, one has the opportunity to structure some small corner of the world to make observations. The assessment designer can use this capability to maximize the value of the data collected, as seen through the lens of the underlying assumptions about how students learn in the domain.

Interpretation

Every assessment is also based on certain assumptions and models for interpreting the evidence collected from observations. The *interpretation* vertex of the triangle encompasses all the methods and tools used to reason from the observations. It expresses how the observations derived from a set of assessment tasks constitute evidence about the knowledge and skills being assessed.

In the context of some assessment, the interpretation method is based on scores that are indicative of varying levels of performance. In the context of other assessment, the interpretation can be based on an intuitive or qualitative model rather than a quantitative one. Even informally, teachers often make coordinated judgments about what aspects of students' understanding and learning are relevant, how a student has performed on one or more tasks, and what the performances mean about the student's knowledge and understanding.

A crucial point is that each of the three elements of the assessment triangle not only must make sense on its own, but also must connect to each of the other two elements in a meaningful way to lead to an effective assessment and sound inferences. Thus, to have a valid and useful assessment, all three vertices of the triangle must work together in synchrony. Central to this entire process are theories and data on how students learn and what students know as they develop competence for important aspects of a domain such as literacy.

Starting with a model of development and learning is critical, since the model suggests the most important aspects of student achievement about which one would want to draw inferences, and provides clues about the types of assessment tasks that will elicit evidence to support those inferences for whatever goal one has in mind with respect to using that information.

A system calls for multiple assessments

Any valid and useful literacy assessment will involve a process of reasoning from evidence about some key aspect of the development of reading, writing, speaking or listening. Thus, a system of literacy assessment necessarily involves multiple such assessment tools and practices. These multiple assessment tools and practices would focus on key elements of the development of early literacy and would be used by various individuals to make judgments about student progress. Sections III-3 and III-4 provide background information and guidance with respect to four key aspects of the individual and collective set of assessment tools and practices that should be chosen to make up an early literacy assessment system:

1. how they relate to knowledge from theory, research and practice about the development of components of literacy,
2. the interpretive purposes they would be intended to fulfill in promoting literacy development,
3. the types of assessment that could be used for specific components of literacy, and
4. desirable properties of such assessment in terms of validity, reliability and fairness.

All of the information provided in Sections III-3 and III-4 is predicated on the core assumption that (a) assessment of early literacy is a process of reasoning from evidence connected to theoretically and empirically sound conceptions of literacy development and (b) the assessment tools are well designed and provide high quality information for the intended interpretive use. These assumptions lie at the core of an early literacy assessment system.

In what follows, we focus on the broader criteria that need to be used in the process of selection and assembly of assessment tools and practices for them to function together, i.e., the ways they need to relate to each other to serve as a balanced "assessment system." The **Portraits** in **Section II** illustrate the multiple aspects of the development of literacy that educators are interested in assessing. They provide examples of how assessment practices and tools might reflect a rich, interconnected model of literacy development and how they can fit together across time and use context, in ways that are consistent with three important system design properties: coherence, comprehensiveness, and continuity.

Criteria for balanced assessment systems

As noted at the beginning of this Guide, *"a collection of assessments does not entail a system any more than a pile of bricks constitutes a house"* (Coladarci, 2002). Assessment systems are balanced when the various assessment tools and practices in the system:

a. are coherently linked through a clear specification of the learning targets,

b. comprehensively provide multiple sources of evidence to support educational decision making, and

c. continuously document student progress over time (Pellegrino et al., 2001).

These properties—coherence, comprehensiveness, and continuity—create a powerful image of a high-quality system of assessment, rooted in a common model of development and learning.

Coherence

By *coherence*, we mean that the models of student learning underlying the various assessments within the system should be compatible. While a large-scale state assessment might be based on a model of learning that is coarser than that underlying the assessments used in classrooms, the conceptual base for the state assessment should be a broader version of one that makes sense at the finer-grained level. In this way, the external assessment results will be consistent with the more detailed understanding of learning underlying classroom instruction and assessment.

As one moves up and down the levels of a system, from the classroom through the school, district, and state, assessments along this vertical dimension should align. As long as the underlying models of learning and development are consistent, the assessments will complement each other rather than present conflicting goals for learning.

In addition to vertical coherence among assessments that range from the classroom to the district to the state level, we should also be concerned about coherence among classroom assessments serving various purposes (e.g., grading, formative feedback). Horizontal coherence is the alignment among curriculum, instruction, and assessment with the goal of helping students develop proficiency in a content domain (Pellegrino et al., 2001).

Comprehensiveness

By *comprehensiveness*, we mean that a range of assessment approaches should be used to provide a variety of evidence to support educational decision making. In an area such as early literacy development, multiple assessments are needed to cover the depth and breadth of the many facets of literacy development that we need to evaluate. No single assessment result can be considered a definitive indicator of a student's knowledge and skill. Information from multiple assessments enhances the validity and fairness of the inferences drawn by giving students various ways and opportunities to demonstrate their competence. Multiple measures can also be used to provide evidence that improvements in test scores represent real gains in learning, as opposed to score inflation due to teaching narrowly to one particular instrument (e.g., Koretz, 2009).

Continuity

An ideal assessment system should be designed to be *continuous*. That is, assessments should measure student progress over time, akin more to a videotape record rather than to the snapshots provided by most current tests. To provide such pictures of progress, multiple sets of observations over time must be linked conceptually so that change can be observed and interpreted. Models of student progress in learning should underlie the assessment system, and assessments should be designed to provide information that maps back to the progression.

In **Section I** we provided a graphical illustration of how a multi-level assessment system might look and mentioned some of the factors that would serve to achieve balance and support these three principles. **Figure III.2.2** refers back to that illustration and highlights four critical features that make it a balanced and integrated system relative to literacy. We also note where in this Guide we elaborate on what needs to be considered for each of the four features.

An Integrated System

- Unified by common learning goals derived from learning theory, research, & content standards (Chapter III-3)
- Synchronized by unifying progress variables that map out expected trajectories of learning and development (Chapter III-3)
- Coordinated within and across system levels & purposes (Chapter III-4)
- Use of quality assessment tools aligned to specific literacy components and levels of proficiency (Chapter III-4)

● **FIGURE III.2.2**

Example of a Multi-level assessment system that illustrates coherence, comprehensiveness, and continuity.

Systems within systems

The three criteria discussed above can be used in the conceptualization, design, and/or evaluation of an ELAS. But systems do not stand alone; it is important to recognize that we also need to take into consideration the reality that systems typically reside within other systems. As conceptualized in *Systems for State Science Assessment* (Wilson & Bertenthal, 2006):

- systems are organized around a specific goal;
- systems are composed of subsystems, or parts, that each serve their own purposes but also interact with other parts in ways that help the larger system to function as intended;
- the subsystems that comprise the whole must work well both independently and together for the system to function as intended;
- the parts working together can perform functions that individual components cannot perform on their own; and
- a missing or poorly operating part may cause a system to function poorly, or not at all.

This idea of systems within systems is noted explicitly in Principle #1 and discussed in Section III-1. The ELAS must be in balance with other school, district, and state level systems related to curriculum, instruction, assessment, professional learning, and accountability. And within the assessment system there will be sub-systems that operate at different levels and serve different purposes.

Examples would be assessments designed for different purposes (see Section III-4) that operate at the classroom and/or district levels, as well as across levels of the Pre-K through 12 system.

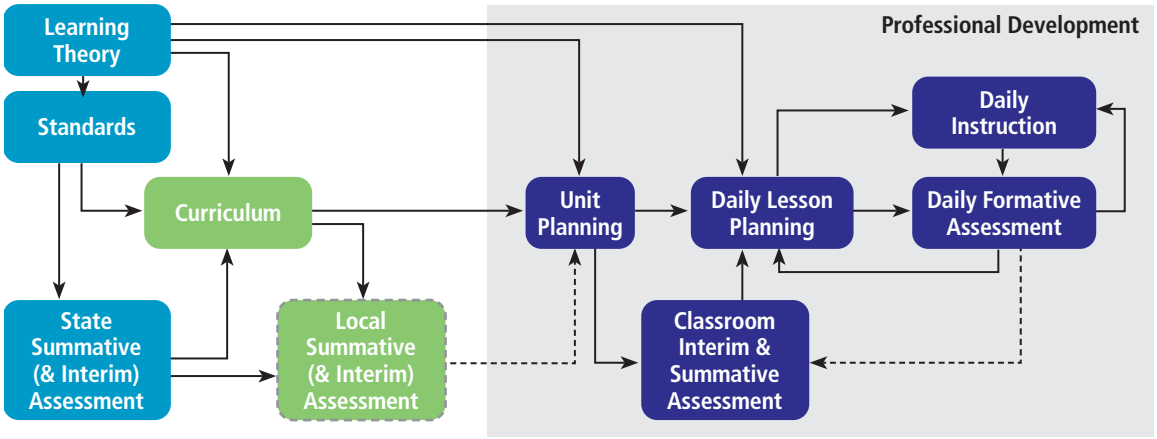
Because there can be considerable complexity associated with planning for and designing the assessment system, given the purposes it is intended to serve and the levels at which it is intended to operate, developing an ELAS theory of action and explicating a logic model for the system can be beneficial and essential in going about this process. These ideas are considered and developed below.

"The ELAS must be in balance with other school, district, and state level systems related to curriculum, instruction, assessment, professional learning, and accountability. And within the assessment system there will be sub-systems that operate at different levels and serve different purposes."

Developing a theory of action and logic model for the ELAS

A common problem across and within state, district, and classroom assessment levels is that the assessment components are not conceptually coherent—they don't align to the same conception of literacy. This can often produce conflicting results and inferences about students. Consequently, the use of these assessments doesn't lead to the desired outcome of educational improvement. It is therefore essential to make explicit one's assumptions about literacy and a "theory of action" related to the use of information derived from the system of assessments.

Figure III.2.3 shows a simplified version of the components of a standards, curriculum, instruction, and assessment system at state, district, school and classroom levels.



● **FIGURE III.2.3**

A "Simple" Theory of Action Relating Teaching, Learning, and Assessment

Notice that everything flows from theory and research on literacy development and learning. Much more would need to be articulated as part of the theory of action about how each of the elements shown above relate to each other and what each is intended to accomplish relative to the goal of supporting the development of literacy. In addition, what is included within each of the boxes and how they function would be part of the elaboration of the system model and the theory of action for how the system is supposed to work.

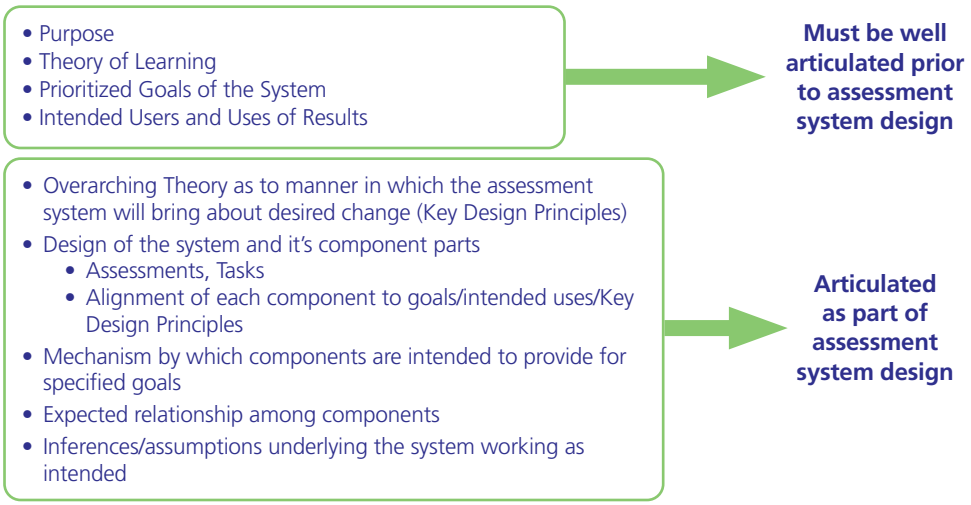
Notice also that much of the action in this representation is focused close to the classroom (area shaded gray), where coordination is needed among curriculum, instruction, and various types of assessment. The Figure III.2.3 also highlights a point made earlier in Section III-1 that effective system operation hinges on teacher expertise, including ongoing opportunities for professional learning.

A theory of action for an ELAS can be conceptualized as an empirically and logically stated argument. It can express a set of underlying assumptions about what something is supposed to do, how it is supposed to function and what is supposed to result. As such, it can serve as a set of testable hypotheses. When clearly articulated, the theory of action outlines how and why a given assessment or system, as designed, will support the achievement of specified goals. It requires specification of each component of the assessment system, the connection(s) between components, and the manner in which they jointly fulfill the requirements of the system.

To help develop and articulate a theory of action for an ELAS, it is recommended that the district's ELAS Leadership Team lay out a "logic model" for the assessment system. A logic model forces one to specify the presumed theory of action. It helps to make explicit assumptions about how particular components are supposed to work, who is to be impacted, what the expected consequences should be, and WHY.

In a complex system, it is critical that the theory of action be articulated, especially with regard to how assessment information is to be used to improve outcomes over time—who will use what information and how. Competing theories of action can be made explicit in the system design phase—choices can be made based on the quality of the evidence and/or argument in favor of adopting one theory in lieu of alternatives.

Consistent with the above, **Recommendations – 1.4, 1.5, 1.6 and 1.7** indicate that the ELAS Leadership Team should lay out a logic model for system design, implementation, and evaluation of the ELAS. The logic model development process forces attention to: a) existing conditions, b) resources, c) inputs, d) outputs, and e) proximal and distal outcomes.



● **FIGURE III.2.4**
Pieces in Articulating the Theory of Action and Logic Model for an ELAS

There is a focus not only on the elements of the system but most especially on the assumed logical and causal relationships among them. The logic model enables monitoring the building of the ELAS and its enactment. It also enables strategies for evaluation of the ELAS along the way and for adjustment and correction as needed. **Figure III.2.4** provides a glimpse of what needs to be considered in this process.

Tools/Resources for PHASE I, Principle #2:

Tools Specific to Logic Models and Theory of Action

The development of a theory of action for the assessment system and a logic model for the system components and design is a challenging task that takes time. These selected resources can guide district ELAS Leadership Teams and others through this process.

Logic Model Development Guide (W.K. Kellogg Foundation, 2004)

The W.K. Kellogg Foundation Logic Model Development Guide focuses on the development and use of the program logic model. The logic model and its processes facilitate thinking, planning, and communications about program objectives and actual accomplishments. This guide provides an orientation to the underlying principles and language of the program logic model so it can be effectively used in program planning, implementation, evaluation, and dissemination of results.

Available at: https://www.bttop.org/sites/default/files/public/W.K.%20Kellogg%20LogicModel.pdf

Logic Models for Program Design, Implementation, and Evaluation: Workshop Toolkit (Institute for Education Sciences, 2015).

This Toolkit is designed to help practitioners learn the purpose of logic models, the different elements of a logic model, and the appropriate steps for developing and using a logic model for program development and evaluation. The toolkit includes an agenda, slide deck, participant workbook, and facilitator's manual. The materials have been designed for use by schools, districts, states, and other groups serving them.

Available at https://ies.ed.gov/ncee/edlabs/projects/project.asp?projectID=401

Theories of Action Aren't Enough: An argument for Logic Models

This article by Juan D'Brot provides some helpful ways to think about and work with theories of action and logic models.

Available at https://www.nciea.org/blog/assessment/theories-action-arent-enough-argument-logic-models

Tools Specific to Assessment Audits or Inventories

The development of a theory of action for the assessment system and a logic model for the system components and design is a challenging task that takes time. These selected resources can guide district ELAS Leadership Teams and others through this process.

District Assessment System Design (DASD) Toolkit (National Center for the Improvement of Educational Assessment, 2018)

This toolkit is useful for districts to determine users of assessment, the different ways that assessment information can be used, and which assessment approaches are most valuable in meeting the assessment information needs of different assessment users in the district.

Available at: https://www.nciea.org/featured-resources

Student Assessment Inventory for School Districts (Achieve, 2014)

This toolkit guides district leaders in taking stock of how many assessments are administered throughout a school year and for what purposes they give assessments. Designed from a student perspective, the audit tool can be used by leaders to make decisions about what amount of testing is appropriate and to be more transparent with parents about the testing in schools.

Available at www.achieve.org/assessmentinventory.

A listing of all Tools and Resources mentioned in this Guide to help you develop an early literacy assessment system (ELAS) is available online at **www.MichiganAssessmentConsortium.org/ELAS.**