

Achieving Balance in Districtwide Arts Assessment Systems: A Review of Issues and Promising Approaches



Suggested citation: Stites, R., Bland, J., & Campbell, A. (2009). *Achieving balance in districtwide arts assessment systems: A review of issues and promising approaches*. Menlo Park, CA: SRI International.

CONTENTS

ACKNOWLEDGEMENTS	iii
INTRODUCTION	1
Notes on Methods.....	3
Background	3
Overview of the Review	5
DISTRICT ARTS ASSESSMENT SYSTEM MODELS.....	7
Relative Strengths and Weaknesses.....	8
DESIGN CHALLENGES AND LESSONS FROM PRACTICE	11
Issues of Comprehensiveness	11
Issues of Coherence.....	14
Issues of Continuity.....	15
Issues of Feasibility	16
Issues of Visibility.....	18
INTEGRATION WITH ACCOUNTABILITY AND IMPROVEMENT SYSTEMS.....	21
FINDINGS AND IMPLICATIONS	25
REFERENCES	29
APPENDIX: INTERVIEW PROTOCOLS.....	31
Topics for First Round Telephone Interviews.....	31
Topics for Second Round of Telephone Interviews.....	32

EXHIBITS

Exhibit 1	Selected Findings from the <i>An Unfinished Canvas</i> Reports	4
Exhibit 2	Assessment System Model Ratings.....	10
Exhibit 3	Assessment System Model Strengths and Weaknesses	20

ACKNOWLEDGEMENTS

We thank Deb Brzoska of the John F. Kennedy Center for the Performing Arts; Erica Halverson of the Department of Educational Psychology, University of Wisconsin, Madison; and Patty Taylor, California County Superintendents Educational Services Association (CCSESA) Arts Consultant for their careful reading and helpful suggestions on a draft of the report. We also thank the expert informants who kindly consented to talk with us during the first and second round of our telephone interviews. We also wish to thank our SRI International colleagues who helped in the production of the report. For guidance and research support, we thank Katie Woodworth; for editorial assistance, we thank Merdith Ittner; for graphics and production assistance, we thank Eileen Behr.

Finally, we are grateful to the California Alliance for Arts Education and to the California County Superintendents Educational Services Association for commissioning this report. We are especially grateful to Laurie Schell of the California Alliance for Arts education and Sarah Anderberg of CCSESA for providing sound guidance and strong support throughout the process of this work. Funding support for this project was provided by the William and Flora Hewlett Foundation.

INTRODUCTION

In this paper, we examine issues and promising practices in districtwide systems for assessing student performance in standards-based, sequential K-12 arts education programs. The 2007 report of the California Arts Education Strategic Task Force (California County Superintendents Educational Services Association, 2007) recommended that the state encourage district assessment of student learning in the arts and that the state fund a pilot district-level arts assessment program. Our review describes three models for the design of large-scale assessment of student learning in the arts at the district level in California. Our primary interest is to identify issues and describe promising approaches for achieving balance in districtwide assessment systems for arts education.

Arts educators in California can draw upon a substantial body of existing work to provide guidance to teachers, schools, and districts in developing and implementing quality assessments of student learning in the visual and performing arts.¹ For this reason, our review does not focus on issues related to the design of arts assessments for the purpose of monitoring student learning and guiding instruction. Instead, our focus is on quality criteria and promising practices in district-level assessment systems for the purpose of collecting and communicating information about student learning and instructional program quality to audiences beyond the classroom and school. It is the need to convey information to external audiences that most clearly distinguishes large-scale assessments—whether at the district, state, national, or international levels—from small-scale assessments which are not used to communicate information about learning and achievement beyond the classroom. The strengths and weakness of particular types of assessment of arts learning are also beyond the scope of this review. We assume that any strong districtwide arts assessment system will include assessments of a variety of types and formats, including selected response formats (such as multiple-choice tests), performance assessments (such as multistep tasks and performances scored on a rubric), and/or portfolio assessments (collection and ongoing evaluation of student work products and performance evaluations over time).

Quality criteria for large-scale arts assessment systems at the district level overlap with those for the state level and other large-scale arts assessment systems. The current review builds on the findings of *An Unfinished Canvas: A Review of Large-scale Assessment in K-12 Arts Education* (Stites & Malin, 2008). In that review, the focus was on quality criteria for the design of state-level arts assessment systems. Here, we turn our attention to issues in the design of arts assessment systems at the district level. Like other large-scale assessment systems, the key to quality in districtwide arts assessments systems is balance. The *An Unfinished Canvas* review applied three criteria for balance in large-scale assessment systems—comprehensiveness, coherence, and continuity—identified and described in the National Research Council Report (2001), *Knowing What Students Know*.²

¹ See, for example, San Diego County Office of Education, *Arts Assessment Resource Guide*. (Sacramento, CA: California County Superintendents Educational Services Association, 2008) and other resources available from the CCSESA Arts Initiative at <http://www.ccsesaarts.org/content/home.asp>.

² The 2001 report from the National Research Council, *Knowing What Students Know: The Science And Design of Educational Assessment*, is a review and synthesis of implications of advances in the cognitive and measurement sciences for reshaping educational assessment. The full report is available online at <http://www.nap.edu/openbook.php?isbn=0309072727>.

A balanced assessment system is one that combines features of comprehensiveness (covers all important standards and learning goals), coherence (reflects a common understanding of learning in the discipline that links assessment to curriculum and instruction and guides classroom-based as well as district-level or state-level assessments), and continuity (enables monitoring of learning progress over time) (National Research Council, 2001; Stites & Malin, 2008, p. 2)

Applying this vision of balance in large-scale arts assessment systems at the district level has required some modification of definitions of each of the existing criteria as well as additional criteria. In the context of district-level assessment systems:

- **Comprehensiveness** is defined in terms of balancing *fidelity* (focus and specificity of measures; ability to capture the most important aspects of student learning in each of the four arts subjects—visual arts, dance, music, and theatre) with *bandwidth* (breadth of coverage of measures; ability to cover the range of student learning in all four arts subject across all levels of schooling—elementary, middle school, and high school).
- **Coherence** includes *alignment* of arts assessments *to standards* and *to curriculum and instruction*.
- **Continuity** is having assessments that define and measure *benchmarks* for student performance *articulated across the grade levels* in each arts subject area.

The two added criteria are feasibility and visibility.

- **Feasibility** is defined as a balance between *external supports* in the form of state (or regional) leadership, resources, and technical assistance and *internal capacity* in terms of district and school leadership, resources, buy-in, and professional development.
- **Visibility** is defined as the ability of the district level assessment system to generate and effectively *communicate* information on student learning and instructional quality *to external audiences* (including parents, policymakers, and the general public) as well as *to internal audiences* (including students, teachers, and administrators).

The introduction to the earlier *An Unfinished Canvas* (Stites & Malin, 2008, p.2) review of promising approaches for large-scale arts assessment at the state level concluded with the following statement:

No existing assessment model or system strikes a perfect balance in terms of these ideal features. Compromise and trade-offs are inevitable. The vision of a balanced approach to large-scale arts assessment—one that captures evidence of important aspects of student learning in the arts and that also supports delivery of strong, sequential, standards-based arts instruction—is not yet a reality.

Our review of promising approaches for large-scale arts assessment at the district level has led us to the same conclusion—we have not found any district that meets all the criteria for a fully comprehensive, coherent, continuous, feasible, and visible district-level arts assessment system. However, we have found many examples of promising practices in district-level assessment of student learning in the arts. Given the current state of the California economy and school finances, the compromises and tradeoffs that are inevitable in the design and implementation of a district arts assessment system in any case are more pronounced. The purpose of this review is not to set the standard for quality in district-level arts assessment beyond the reach of any school district in California, but rather to clarify the decision points where compromise is possible while maintaining a degree of overall balance in the quality and effectiveness of the assessment system.

NOTES ON METHODS

The summary of design issues and quality criteria for districtwide assessment systems for arts education is based on a review of documents and a series of interviews with people who have experience and expertise relevant to the topic (see Appendix). The initial search and review of documents (including a variety of research, policy documents, practice guides, and program descriptions) allowed us to refine our research questions, identify and refine topics for further inquiry, identify sources of information (documents and people), and create templates for summarizing information related to the research questions. We then conducted two rounds of telephone interviews. In the first round, we completed interviews with nine experts on arts education policy and practice. Our questions focused on identifying general issues in the development, implementation, and current status of district-level arts assessment. We also asked for help in identifying districts that would be good candidates for case studies. After concluding the first round of document reviews and telephone interviews, we met with representatives of the California Alliance for Arts Education, the California County Superintendents of Education Services Association, the California Department of Education, and others. At that meeting, we presented and discussed our initial findings—including the finding that no districts suitable for full-blown case study analysis had yet been identified—and were advised to broaden our study in two ways. First, we concluded that developing a composite description of promising practices from many districts would be a more productive approach for our study than the case studies of exemplary districtwide assessment systems we had originally planned to do. Second, we decided to expand the scope of our review to include consideration of the integration of districtwide assessment systems within broader arts education accountability and instructional improvement systems. To pursue these ends we conducted additional reviews of literature and a second round of interviews to gather information on promising approaches for achieving balance in districtwide systems for assessing student learning and improving program quality in K-12 arts education.

BACKGROUND

In early 2007, SRI International published *An Unfinished Canvas. Arts Education in California: Taking Stock of Policy and Practice*, a statewide study on the status of arts education in California (Woodworth et al., 2007). That study's findings served as the impetus for a series of follow-up studies, including a review of current models for large-scale arts assessment and state arts assessment systems. Several findings from the subsequent *An Unfinished Canvas* studies (Gallagher, et al., 2008; Guha, Woodworth, Kim, Malin, & Park, 2008; Woodworth, Campbell, Bland, & Mayes, 2009; Woodworth, Peterson, Kim, & Tse, 2009) highlight the importance of strengthening arts assessment and accountability systems in California school districts (for an overview of key findings from the studies see Exhibit 1).

Exhibit 1
Selected Findings from the *An Unfinished Canvas* Reports

- 89% of California K-12 schools fail to offer a standards-based course of study in all four disciplines—music, visual arts, theatre, and dance—and thus fall short of state goals for arts education.
- Standards alignment, assessment, and accountability practices are uneven in arts education and often not present at all.
- Inadequate elementary arts education provides a weak foundation for more advanced arts courses in the upper grades.
- Secondary arts education is more intense and substantial than elementary arts education, but participation is limited.
- Inadequate state funding for education is a top barrier to the provision of arts education, and reliance on outside funding sources, such as parent groups, creates inequities.
- Pressure to improve test scores in other content areas is another top barrier to arts education.
- Districts and counties can play a strong role in arts education, but few do.
- Many districts lack the infrastructure—including strategic arts plans, district arts committees, and arts coordinators—to support arts education programs and build towards implementation of state standards.
- Higher-capacity districts are more likely to take a systemic approach to arts education.

Policy recommendations related to California arts education assessment and accountability stemming from the *An Unfinished Canvas* reports included the following:

At the State Level

- Strengthen accountability in arts education by requiring districts to report on the arts instruction provided, student learning in the arts, and providers of arts instruction, and by supporting the development of appropriate standards-aligned assessments for use at the state and district levels.

At the School and District Level

- Signal to teachers, parents, and students that the arts are a core subject by providing professional development for teachers and establishing assessment and accountability systems for arts education.

Strengthening district assessment of K-12 arts education program quality is also a priority at the federal level. For example, all projects funded under the National Endowment for the Arts *Learning in the Arts for Children and Youth* grants must include assessment of students “according to national or state arts education standards” and “(w)here appropriate, ...multiple forms of assessment, including pre- and post-testing.”³

³ From National Endowment for the Arts, Grants for Arts Projects *Learning in the Arts for Children and Youth*, July 22, 2009, from <http://arts.endow.gov/grants/apply/GAP10/LITA.html>. Applicants for NEA grants must also choose between one the following outcomes: (1) Children and youth demonstrate increased skills, knowledge, and/or understanding of the arts, consistent with national or state arts education standards; (2) Teachers, artists, and other educators demonstrate increased knowledge and skills necessary to engage children and youth in arts learning, consistent with national or state arts education standards. To assist applicants in developing plans for evaluation of one of these two outcomes, the NEA provides an online guide to “Outcome-Based Evaluation: A Working Model for Arts Projects” (see <http://www.arts.gov/grants/apply/out/index-out.html>). The online guide to outcome-based evaluation (written in 2004) does not include results on standards-aligned assessment of students as an

OVERVIEW OF THE REVIEW

This review of issues and promising practices in districtwide arts education assessment systems is presented in three parts. The first section describes and outlines the general areas of strength and weakness of three basic models for district-level arts assessment systems. The second section takes a closer look at challenges and lessons from practice in the design of balanced districtwide assessment systems for arts education that meet the quality criteria of comprehensiveness, coherence, continuity, feasibility, and visibility. The third section discusses the integration of districtwide assessment systems within broader arts education accountability and program improvement systems.

The purpose of this review is to clarify issues and decision points that California school districts will face in efforts to achieve balance in their own district systems for assessing student learning, for improving program quality, and for enhancing reciprocal accountability for arts education.⁴ Reciprocal accountability systems not only hold schools and teachers responsible for student performance but also hold federal, state, and local educational agencies and stakeholders accountable for ensuring that schools have adequate capacity and resources to provide strong instruction to all students. Following the discussions of each model's potential areas of strength and weakness in supporting effective assessment and accountability systems, design challenges related to each quality criterion, and issues in the integration of district assessment, accountability, and improvement systems, the paper concludes with recommendations for achieving balance in the design of districtwide arts education assessment systems for California school districts.

“essential element” of the evaluation plans required from NEA applicants. However, the review criteria for FY2010 applicants for Learning in the Arts for Children and Youth grants does include “plans for assessment according to national or state arts education standards” (see *NEA Grants for Arts Projects: Application Review: Review Criteria*, retrieved July 22, 2009, from <http://arts.endow.gov/grants/apply/GAP10/ApplicationReview.html>).

⁴ A detailed description of reciprocal accountability for arts education can be found in J. Landon and D. P. Russell, *Accountability in Arts Education: Building a Statewide System of Reciprocity* (Pasadena, CA: California Alliance for Arts Education 2008) available at http://www.artsed411.org/advocate/docs/CAAllianceforArtsEd_AccountabilityArtsEd_2008.pdf.

DISTRICT ARTS ASSESSMENT SYSTEM MODELS

In our initial review of documents and interviews, we gathered information on large-scale arts assessment practices in more than 20 school districts across the country. In the second round of data collection, we followed up with telephone interviews to gather more detailed information on district-level arts assessment practices in a dozen school districts in seven states. Looking across the information we gathered on promising practices in district level arts assessment, we found that we could categorize districts into three groups based on the source of the assessment tasks used in each district.

The first group is comprised of districts that have adopted state-developed assessments. We will call this assessment system the **Adoption Model**. All of our examples of districts employing the Adoption Model are in the State of Washington where districts are required to use the state-developed Classroom-based Performance Assessments (CBPA) in their districtwide arts assessment systems. No other state has mandated use of a common set of arts assessments in all districts, but this does not preclude using the Adoption Model as the basis for the design of a districtwide arts assessment system in other states. Many states, including California,⁵ have developed prototype arts assessments at the state or regional level that are available for selection and voluntary use by local districts.

The second group is comprised of districts that have developed their own set of common assessment tasks for districtwide assessment of student learning in the arts. We will call this system the **Development Model**. We found several examples of districts that have developed their own set of common arts assessment tasks for districtwide use in a number of northeastern states.⁶ For example, in Pennsylvania, the state education code calls for schools to develop local assessments of state standards (including arts standards) not included in the state assessment

⁵ The California Arts Assessment Network (CAAN) was formed by the California Department of Education (CDE) in partnership with The California Arts Project (TCAP) in 1998 and since then has been working with several counties and districts to develop and disseminate arts assessment tools. CAAN provides resources to help California schools and districts develop local arts assessment systems. It provides its affiliate schools and districts with assessment samples, models for training school and district staff on use of the assessments, and a sample item pool. In partnership with TCAP, CAAN has a website, "Student Work Online," where teachers can post student work and obtain assessment feedback from other members. Recently, The California Arts Project (TCAP) has been developing a prototype eighth grade music assessment.

⁶ See, for example, The New England Arts Assessment Network, *Arts Assessment in New England: Narrative Accounts of Existing Projects*, (June 2006), for descriptions of district arts assessment systems in Connecticut, New Hampshire, Maine, Rhode Island, and Vermont. The New England report describes district arts assessments that blend elements of all three models described here. For example, the two district arts assessment systems in Connecticut described follow the Development Model in developing common performance tasks for districtwide use but modeled their assessments after Washington State's Classroom-based Performance Assessments. The district arts assessment systems described in the other New England states combine elements of utilizing state-level resources and models (Adoption Model) with district-level development of common assessment tasks (Development Model) and district-level rubrics (Benchmark Model).

system. To support local arts assessment development, the Pennsylvania Department of Education developed guidelines and models for districtwide assessment of student learning in the arts, and some school districts in Pennsylvania developed and implemented their own common assessment tasks for districtwide assessment of student learning in the arts.

Districts in many states have developed common benchmarks and scoring rubrics for districtwide use but have allowed teachers and schools within the district to develop their own assessment tasks. We will call this system the **Benchmark Model**. Our examples of districts using the Benchmark Model come primarily from Arizona. It should be noted, however, that some of the districts we studied in Arizona are moving toward development of common, districtwide arts assessments more in line with the Development Model.

The three models for district level arts assessment systems we have identified are distinguished by the source and nature of the assessment tasks used in each. Examples of district-level assessment systems for arts education of the type described by each model can be found in many states and are not restricted to districts and states included in our review. School districts in California can choose any one of these three models as the foundation for the design of a districtwide assessment system for arts education.

- **Adoption Model** (District-adopted assessments). In this model, a district selects and implements assessments of student learning in the arts developed by the state or county educational agency, or by another external organization.
- **Development Model** (District-developed assessments). In this model, a district develops its own common set of assessments of student learning in the arts and implements them districtwide.
- **Benchmark Model** (District-developed benchmarks). In this model, a district develops and implements common benchmarks and/or scoring criteria (rubrics) that define districtwide expectations for student learning in the arts. No common set of assessment tasks is used districtwide in this model.

RELATIVE STRENGTHS AND WEAKNESSES

Each of these three models has a distinct pattern of potential strengths and weaknesses with respect to the five quality criteria for balance in a districtwide arts assessment system. One of the key strengths of the Adoption Model, from the district perspective, is that the burden of developing arts education assessments is borne primarily by the state. This is not to say that this type of arts assessment system is without costs to the district. Districts in Washington, for example, must take on responsibility for selecting state-developed Classroom-based Performance Assessments (CBPA) aligned with district curricula and must train teachers to properly administer, score, and report results on the selected CBPA. In terms of criteria for balance in districtwide arts assessment systems, one of the greatest potential strengths for an Adoption Model assessment system is in the area of comprehensiveness and especially bandwidth. Economies of scale in developing assessments tasks at the state level make it far more likely that tasks will be created to measure proficiencies on a broad range of standards for all four visual and performing arts. For the same reason, continuity—having access to sequential assessments covering proficiencies at the elementary, middle school, and high school levels for all four arts is also a key potential strength of an Adoption Model assessment system. State-developed assessments are more likely to be closely aligned with state arts standards than they are with district curriculum and instruction—though such assessments may push district curriculum and instruction toward alignment with state standards. The quality of the arts assessments available to districts and the way selection and implementation of state-developed assessments pushes

districts and, at the same time, supports districts to build capacity for greater alignment of district curricula and instruction with state arts standards gives the Adoption Model high potential for feasibility. The model also has high potential for visibility because of the way that statewide use of common assessments facilitates communications of arts learning results to the broadest possible range of external and internal audiences.

The Development Model can be a relative high cost option for a district, and as such, when applied in practice, rarely results in a comprehensive districtwide assessment system with coverage of all four arts subjects at the elementary, middle school, and high school levels. The key potential strengths of the Development Model are close alignment of assessments with district arts curriculum and instruction and strong potential for supporting district capacity building and arts education program improvement. Good continuity in the form of sequential expectations for arts learning articulated across grade levels is possible within the Development Model, though, in practice, the cost of developing assessments typically limits the range of grade levels and arts subjects covered by the assessment system. The Development Model's potential for feasibility is strong in some areas and weak in others. Although district development of common assessments can be a strong foundation for building capacity within the district (in the form of leadership and expertise) to improve delivery of standards-based sequential arts instruction, few districts can afford the investment in human capital needed to fully develop and implement this model. Visibility of learning results in the arts is relatively strong within the Development Model because results are reported on assessments that are easily understood within the district context. However, the potential for making arts learning results visible to external audiences such as other schools and districts, policymakers, and the general public is relatively weak in this model.

The Benchmark Model can be a relatively low cost option for a district, but requires a considerable commitment of time from individual teachers and schools to be fully developed and implemented. The fact that individual teachers or teams of teachers at the school or departmental level must take on primary responsibility for developing the assessment tasks that will be scored according to the districtwide benchmarks and rubrics can be an area of strength or of weakness in the Benchmark Model. It will be an area of strength when teachers have the opportunity, ability, and support to become fully engaged in the task of developing and implementing assessments of student learning in the arts. However, assigning responsibility for assessment task development to individual teachers, departments, and schools can be a serious weakness when there is a lack of professional development opportunities, a lack of common planning time, and low or uneven levels of teacher, departmental, and school commitment to using district benchmarks. For this reason, the Benchmark Model has relatively low potential for supporting high levels of comprehensiveness, coherence, and continuity. Strong visibility may be particularly hard to achieve with this model because, in the absence of common assessments, there is very limited potential for reporting learning results in ways that will be easily understood and correctly interpreted beyond the classroom.

Exhibit 2 below summarizes the relative strengths and weaknesses of each district assessment system model. A more detailed discussion of the rationale for the ratings assigned to each system model can be found in the section on design challenges and lessons from practice that follows.

Exhibit 2
Assessment System Model Ratings

	Adoption Model	Development Model	Benchmark Model
Comprehensiveness	●	○	●
Fidelity	○	○	●
Bandwidth	●	○	○
Coherence	●	●	○
State standards	●	○	○
Local curricula	○	●	○
Continuity	●	●	○
Feasibility	○	●	●
State support	●	○	○
District capacity	○	●	○
School capacity	○	○	●
Visibility	●	●	○
External	●	○	○
Internal	●	●	○

KEY: ● = Strong
 ○ = Moderate
 ○ = Weak

DESIGN CHALLENGES AND LESSONS FROM PRACTICE

The discussion in this section highlights some of the practical challenges that districts face in designing districtwide assessment systems. Beyond constraints stemming from state and local policies and from limited funding and human resources, district choices in the design of their assessment systems are also constrained by the depth and breadth of the existing arts education program in the district. Wherever possible, the discussion in the following section is grounded in descriptions of practices and lessons learned in the district arts education assessment systems we documented in our second round of interviews. Particular strengths and weaknesses of each assessment system model relative to the criteria for balance in an assessment system are summarized in Exhibit 3 at the end of this section.

ISSUES OF COMPREHENSIVENESS

No district arts assessment system can include the number and variety of assessments that would be needed to cover every important aspect of knowledge and skills that is taught and learned in all four arts subjects and at all levels from kindergarten through high school graduation. Choices must be made about which arts subjects will be assessed at what levels. Typically, districts choose to assess student learning in the arts at no more than three to five grade levels (for example, at grades 5, 8, and 10, or at grades 3, 5, 7, plus end-of-course tests in high school). This is true across all three models. The choice of which arts subjects to include in the districtwide assessment system at each grade level will be constrained by the extent to which curriculum and instruction in each subject are available to students at each level. For example, implementing an elementary level dance assessment will not be meaningful if the district offers no dance instruction in elementary school.

However, a districtwide assessment system may also be designed to push the extent to which standards-based arts instruction is available at each level in the district. For example, our interviewees from Adoption Model districts told us that implementing districtwide assessments provided encouragement and guidance for curriculum development. From one such district we heard the following:

The assessment is driving instruction. It's driving instruction because there wasn't anything there before. There just aren't that many theatre curricula to choose from—so teachers develop their own. Now they have a target.

From another Adoption Model district we heard the following comment:

Right now, the district does not have arts curriculum adopted for grades 5 through 8. Our department is working on that, but the department is only 2 years old. Currently teachers are developing their own aligned, sequenced programs.

The question of how comprehensive a districtwide arts assessment system can and should be also involves consideration of the tradeoffs between how deeply (fidelity) and how broadly (bandwidth) assessments in the system will cover the knowledge and skills in the district arts

curriculum. The key question to ask in evaluating the fidelity of any assessment is how well the assessment has captured the learning objectives embedded in curriculum and instruction. Under ideal implementation conditions, the Benchmark Model offers stronger potential for achieving high fidelity in a districtwide assessment system than either the Adoption or Development Models. When the district assessment system consists solely of common benchmarks for student performance in the arts, assessment tasks can be more closely tailored to align with particular course curricula and instruction. For example, in one of the Benchmark Model districts we studied, students are required to earn at least one credit in fine arts (including band, choir, mariachi, percussion, theater, dance, visual arts, and digital arts) to graduate from high school. Because end-of-course summative assessments are required by the state, teachers first looked at goals for what students should be able to do in the arts at the end of high school, and then developed criterion-referenced tests (CRTs)⁷ aligned closely with the instructional content of particular courses. Using end-of-course CRTs in this way can result in assessment tasks with a high level of fidelity to particular arts course curricula and instruction. By contrast, the common assessment tasks developed at the state level in the Adoption Model and at the district level in the Development Model are aligned with standards and broader arts curricula covering a greater variety of content and learning expectations. For this reason, the potential for a high level of fidelity—focused assessment of particular aspects of learning in individual arts courses—may not be quite as good in these two models as in the Benchmark model. The tradeoff for lower fidelity in the Adoption and Development models is better, long-term potential for coherence as teachers adapt and align instruction with the learning expectations embedded in standards-based and curriculum-based assessment tasks.

Achieving a high level of fidelity in assessments may present somewhat different challenges in different arts subject areas. For example, individual performance ability seems to be relatively easily assessed with high fidelity in the visual arts but is much more challenging to assess in music, theatre, and dance. In the performing arts, ensemble performance is often the focus for instruction, and there may be little time to evaluate evidence of individual performance ability outside of the context of a group performance. Interviewees from districts representing all three models mentioned the particular difficulty of achieving fidelity in assessment of individual performance skills in music. Washington’s CBPA for music are designed as individual assessments and our interviewees from Washington (Adoption Model) districts told us that this has forced music teachers to move away from the assessment of group performances and shift toward the less familiar practice of assessing individual performance skills. Interviewees from Development and Benchmark Model districts also told us that music was the most difficult of the four arts subjects to assess. The reasons they gave for the difficulty included the fact that teachers had no time for individual student assessment in music and that they were not able to assess an individual student’s musical technique in the context of an ensemble music performance. An interviewee from a Benchmark Model district offered the following comments:

Visual arts assessments have been relatively easy and music the hardest to do well. Theatre and dance, like music, have the advantage of performance being a natural part of the curriculum...(but now) career and college readiness has helped shift the focus to individual assessment. Music is used to concerts, so they’ve had some trouble focusing at the student level because there’s no time for individual assessment...Focusing on individual assessment in music performance has been hard.

⁷ A criterion-referenced test (CRT) is a test that evaluates student performance against specified criteria for mastery of knowledge, skills, and abilities as opposed to a norm-referenced test (NRT) that evaluates a student’s performance by comparing it to performances on the test by other students (norming sample). Standards-based tests are a type of CRT (the reference criteria are standards).

We also heard the following comment on some tradeoffs in assessing music performance at the individual level as opposed to the group level from an interviewee from a Development Model district:

Seventh grade music is probably the hardest to assess, because students are in one of four courses; they are either in band, orchestra, chorus, or general music. As a result, the (district common) assessment has to be centered around the lowest common denominator. The teachers have been having discussions about the validity of this sort of assessment, since it doesn't assess individual student technique. It assesses things like sight reading ability, listening skills, things like that. We have such strong performance ensembles, and (the teachers) are not sure the assessment really gets at that.

These two sets of comments from our interviews raise a number of issues related to the difficulty of assessing individual performance ability in music. Time available for individual performance and assessment in music (and probably also in other performing arts) seems to be a significant limiting factor. The absence of effective assessment tools, strategies, and skills to evaluate individual ability in the context of group performances is another limiting factor. Finally, the second set of comments raises the challenge of developing and using common assessments that are valid measures of musical skills across the variety of musical subjects and performance forms that students study. It also notes the problem of overemphasizing individual student performance in assessments at the expense of evaluating the quality of ensemble work.

Achieving good bandwidth (breadth of coverage of arts subjects across all grade levels) in a districtwide assessment system may be somewhat easier in an Adoption Model district than in either a Development Model district (where the cost of developing common assessments for every arts subject area may be prohibitive) or in a Benchmark Model district (where reliance on individual arts teachers and departments to develop assessment tasks may limit development of assessments to some but not all arts subjects or to only the high school level because of the lack of arts specialists, especially at the elementary and middle school levels). Yet, even within the Adoption Model, the district may need to supplement state-developed assessments to achieve good bandwidth in the districtwide assessment system. We heard from an interviewee from an Adoption Model district the following comment:

The (state) assessments we are using are fairly global across standards, but there are some areas we have found...that need to be filled in. We need to fill in some areas around career and technical education where there is crossover between arts and career classes. The (state) assessments aren't covering the career component—this is in areas like photography, ceramics, and graphic arts—so our district created some new assessments and is sending them to the state as an example.

Although state-developed assessments may not be available to cover every facet of the district arts curriculum, coverage of state standards is likely to be relatively comprehensive and assessments equally available in all four subject areas. This gives districts in the Adoption Model some flexibility in selecting the breadth of coverage of assessments across arts subjects. An interviewee from one Washington (Adoption Model) district told us the following:

CBPA [state assessments] are equally available in all four arts areas, but because theatre and dance are not taught as often in the district, we don't need as many CBPA to cover those subjects.

The Adoption Model also provides districts with opportunities to easily expand their districtwide assessment system as they expand the range of their district arts curriculum. An interviewee from

another Washington (Adoption Model) district described the way that state-developed assessments can support expansion of the arts curriculum.

Right now we have [CBPA] assessments in visual arts, music, and theatre...dance is coming along. That will be done through a community district partnership because we can't hire in this fiscal climate. The district has developed some partnerships with local dance companies, to teach dance one day per week. They will use CBPA to track progress.

The existence and availability of assessments to cover expansion of the arts curriculum into areas not previously taught in the district is an important strength of the Adoption Model.

ISSUES OF COHERENCE

Just as the comprehensiveness of a districtwide assessment system for arts education is constrained by the comprehensiveness of arts curriculum and instruction in the district, so too is the coherence of an arts assessment system constrained by the degree to which there is a coherent (aligned with state and district standards) arts instructional program in the district. Coherence in a districtwide assessment system should be considered at two levels: alignment of the assessment to state arts education standards and alignment of assessments to district arts curriculum and instruction. In general, the Adoption Model offers the strongest potential for alignment of assessments to state standards because state standards are the primary source of specifications for the development of state assessments. The Development Model offers the strongest potential for alignment of assessments to district curriculum and instruction with the tradeoff being that local standards and instructional content will play a larger role in shaping the targets for assessment than will state standards. The degree of flexibility given to teachers in designing assessments and the resulting variability in the assessments within the Benchmark Model makes it difficult to achieve consistent alignment of assessments with either state standards or district curriculum.

Well-written standards and strong alignment of district curriculum and instruction to state standards can provide strong support for achieving coherence in a districtwide assessment system for arts education. We heard the following from a Development Model district:

Our [state] standards were pretty well-written, and there's good taxonomy across them. As a result, it was a pretty natural process to align curricula and standards and assessments.

Variability in the quality of arts standards may make coherence easier to achieve in some arts subject areas than others. An interviewee from a Benchmark Model district told us that achieving alignment of visual arts assessments to standards was made more difficult by the fact that the visual arts standards focused on concepts and vocabulary that are more abstract and more difficult to assess than performing arts standards.

In reality, districts do not teach to all district (or state) standards. Determining how many arts standards and prioritizing which arts standards should be the focus for instruction is itself a challenging and contentious issue. Establishing which standards will be assessed is a key part of this challenge. In practice, a coherent district arts assessment need only be aligned with a small number of key standards taught in the district arts curriculum. The first comments below are from an interviewee in a Benchmark Model district, the second from an interviewee in an Adoption Model district.

There are too many standards to teach in one year. The state standards came out in 2006 and we've worked for about 2 years to get people to understand those standards. Because

of the training we did in the last 2 years, developing course expectations was strongly tied to standards. We cover about 75 percent of standards.

There is no way a district can cover all standards. So in our district we take the grade-level expectations and hone them down into a document that is manageable. If we were to teach all standards in all subjects, students would have to be in school for 39.5 years. Our district has chosen which standards to focus on so students can get more depth.

Among the standards selected to be taught in a district, an even smaller number will be selected to be assessed. Giving teachers a leading role in the selection of assessments that are best aligned to curriculum is a common strategy for achieving coherence. The following is from an interviewee in a Development Model district.

I spoke with teachers to find the most important assessment standards. I let the teachers make a lot of the decisions, since it was in their classrooms. We've revised the assessments every year. Faculty helped design them at each step.

Giving teachers a leading role in selecting assessments to be included in the districtwide system can also have the effect of encouraging them to make adjustments to their instruction to bring it into better alignment with standards. The potential for curriculum realignment with standards is greatest when the assessments that teachers are selecting are themselves well aligned with standards, as is illustrated in the comments below from an interviewee from a Washington (Adoption Model) district.

We have asked teachers to choose the CBPA that best fits their curriculum...to select CBPA that match what they're doing. At times, they had to adjust their teaching. They looked at the CBPA and realized that they weren't emphasizing the right things. It helps them adjust their focus for the year to refine instruction and to align with state standards.

Asking teachers who teach the same arts subjects at the same grade level to use the same assessment tasks for the districtwide assessment can also move a district toward greater alignment of standards, curriculum, and assessments.

Collaboration between districts and the state in an ongoing process of standards and assessment revisions may also be needed to move toward greater coherence in district arts assessment systems. Interviewees in Washington districts told us that the state department of education had been receptive to feedback from teachers on ways to better align the CBPA to district arts curricula and that the state was open to making corrections.

ISSUES OF CONTINUITY

Continuity in a district arts assessment system is also constrained by the extent of continuity in the districts arts education program and, as is true for comprehensiveness and coherence, improving the continuity of the districts arts assessment system can be part of an ongoing process of improving continuity of arts curricula and instruction across grade levels. The Adoption Model offers the greatest potential for achieving a system of assessments that measure clearly articulated sequences of arts skills development across the grade levels. This is so because state standards, by design, are more likely to include articulated sequences of skills development in each arts subject across elementary, middle school, and high school levels than are district curricula or benchmarks that may not extend across all grade levels. If the districtwide assessment system clearly embodies a sequence of skills development across grade levels, it can help districts identify and remove gaps in skills development in the curriculum. For example, one of our interviewees from an Adoption Model district made the following points:

The assessments are connected, but teaching and learning have gaps. The secondary level requires notation in music, but the elementary curriculum doesn't provide that instruction. The assessments are now making us look at the overall system.

Continuity of districtwide assessments in a Development Model district may vary from one arts subject area to another. Because the common assessments developed for districtwide use are typically not comprehensive across all arts subjects at all levels, continuity may be only possible for some arts subjects and not for others or for only limited spans of grade levels in some arts subjects. For example, one interviewee from a Development Model district described different sequences of assessments for arts subjects in middle grades and in high school.

Districtwide assessments are now given for third- and seventh-grade music and for fifth- and eighth-grade visual art—these are the disciplines that are taught to all K-8 students. High school arts are all elective-based—there's no graduation requirement—but there's a common end-of-course assessment for each course—music, visual arts, dance, and theatre, as well as music technology and video production. There is often only one teacher for each of these disciplines, so there's not a lot of need for cross-collaboration the way there is at the elementary and middle school levels.

Continuity may be difficult to achieve in a Benchmark Model assessment system because articulation of skills across assessments at different levels may be limited by a lack of arts specialists at the middle school and elementary levels and limited opportunities for collaboration with high school arts specialists. State standards and district arts curricula that are clearly articulated across grade levels are essential ingredients for achieving continuity in a districtwide arts assessment system. But, as noted above, the Benchmark Model provides relatively weak support for alignment of district arts curricula and state arts standards, and this weak support further limits the potential for achieving continuity in the district arts assessment system. We heard the following from one Benchmark Model district interviewee:

Our curriculum does not build from one year to the next. Philosophically, we are not there. I believe in teacher autonomy but what I want teachers to think about is the gap between what we want kids to know at the end of the semester and where the kids are.

ISSUES OF FEASIBILITY

Any districtwide assessment system for arts education—or a state or district assessment system in any other core subject area—requires a large investment of time and resources to develop fully and implement effectively. Quality in a districtwide assessment system will not come cheaply or easily in any case, but each of the three general models for districtwide assessment systems imposes somewhat different demands on external supports and internal capacity. The feasibility of choosing the Adoption Model, the Development Model, or the Benchmark Model will depend upon where a district is most likely to find the leadership, capacity, and resources to support the design of a districtwide assessment system. The Adoption Model requires strong leadership and investment of resources at the state or regional level to operate effectively. In this model, much of the heavy lifting—design and development of standards-aligned assessments—happens outside the district, though as noted above, districts must also invest considerable time and resources to ensure that assessments are well chosen and well used. In the Development Model, the district must itself have sufficient capacity to design and implement common assessments, though the state can also play a vital supporting role by ensuring that state standards are well written and by providing technical assistance for district-level assessment design. In the Benchmark Model, the success of the system rests primarily in the hands of teams of teachers in individual schools and departments with both the state and the district playing supporting roles in the design of assessments.

In all three assessment system models, getting teachers and administrators actively engaged in the design and implementation of the assessment system is an important component of success. Active engagement in the design of the assessment system at any level will help build understanding and acceptance of the legitimacy and value of the system. An interviewee from a Benchmark Model district offered the following comments on the importance of involving as many teachers as possible in the design and implementation process.

We had all teachers work on it to strengthen buy-in. Then we focused on standardization and now are turning to how they are grading so it becomes more standardized. Next we need to implement professional development. It's a big challenge to get everyone on the same page. There is a big difference between people who get it and those who don't. [Fortunately,] a couple years ago we had an influx of new teachers who are really excited about doing this.

Engaging administrative staff can also be important to successful design and implementation as noted by an interviewee from a Washington (Adoption Model) district.

We really needed to get principal buy-in because it is not a funded mandate. We used some time during the principal summer institute to educate them about the CBPA, put them through a theatre CBPA and scoring. We also got CBPA on the assessment calendar so principals knew when data was due and knew to remind teachers.

The need for professional development to support the design and implementation of a districtwide assessment system goes beyond the need to obtain teacher and administrator buy-in. Many aspects of a districtwide assessment system will be new to teachers and administrators and therefore building district capacity to effectively design and implement the arts assessment system through robust and varied technical assistance and professional development opportunities is an important part of achieving feasibility. The comments below, the first from a Development Model district interview, and the second from a Washington (Adoption Model) district interview illustrate the range and scope of professional development needs.

My elementary art teachers are all retirement age. A rubric was completely foreign to them. So that was a lot of hard work there...The first thing we do now with the arts teachers is talk about anchors and look at the rubric.

There is a need for initial use training. For that, you need to take teachers through the CBPA, where to find them, and how to use online reporting (which is new for elementary teachers). We offer training classes to improve scoring reliability. We also created a FAQ sheet with the basics and developed a PowerPoint presentation to educate teachers ...There is also a need for training on technology—how to document arts using cameras, etc. ...Early release days for professional development are helpful. We also have a wonderful music coach who mentors music teachers.

The comments below from a Development Model district interview highlight the need for professional development for district coordinators and others who will take on leadership roles in the design and implementation of the district arts assessment system as well as the potential for finding expert assistance from colleagues in other disciplines in the district and from arts professionals outside the district.

I (the district coordinator) did not initially have the knowledge and skills to put this all into place and it was very helpful to speak with counterparts in other disciplines about the common assessments they'd developed. They really helped me to develop the scoring process and keep it as valid and reliable as possible. My assistant superintendent, my direct supervisor, is completely behind this 100 percent. Every time I had questions

or needed resources, she's been right there. I also got a great deal of professional support from colleagues in the local arts community.

ISSUES OF VISIBILITY

At a time when the existence of arts education programs in K-12 education is threatened by pressures to reduce spending on public education and simultaneously raise achievement levels in core academic subjects included in statewide testing and accountability systems, there is a pressing need to make the value and effectiveness of arts education more visible. Districtwide assessment systems have an important part to play in increasingly the visibility of learning results from district arts education programs. The Adoption Model has the strongest potential to support both internal visibility (communication of arts learning results across schools, levels, and departments within the district) and external visibility (communication of arts learning results to audiences of local, regional, and statewide stakeholders including parents and community stakeholders, other school districts, the general public, and policymakers). The statewide use of a common set of assessments aligned to state standards creates the strongest potential for a common understanding of arts learning results from school to school and district to district. The external and internal visibility of arts learning results is much more limited in both the Development Model (where results may not communicate well beyond the local community and district) and the Benchmark Model (where results may not communicate well to any external audiences and may not communicate beyond the classroom or department level within the district).

Visibility in any assessment system model is partly a matter of tailoring and reporting information gained from assessments in ways that communicate well to particular audiences. In general, information for internal audiences needs to be reported in greater detail than information reported to external audiences. The following comments are from an Adoption Model district interview.

This year for the first time, teachers have to report back the percentage of students who have met the standards. The state just requires that we do the test. I'm requiring that our district collect information on percentage of students meeting proficiency. I developed a reporting form and have a spreadsheet that calculates percentage of students meeting standards. So in addition to the state report, we have a district report. It's much more quantitative. We need to have the detailed information to change instruction.

A strong districtwide assessment system with well-crafted reporting of results has the potential to change perceptions of the quality and the value of arts learning among external and internal audiences. As noted in the comments below from an Adoption Model district interview, communicating district assessment results effectively to external audiences can enhance perceptions of the value of arts education.

I present all the time...The district needs someone who will be an advocate...Data should be used to promote the positive aspects of a program, rather than punitive. Because of the assessments and using data, I have been able to talk to stakeholders about the arts being equivalent to a core curriculum. Data changes perception. This year, of all the districts that are making cuts, our district did not cut any arts program—we are actually growing our arts program.

Effective internal communications can also change perceptions of the importance and quality of arts instruction as part of the core curriculum. An interviewee from a Development Model district observed that the advent of arts assessments had not had much impact on community perceptions of the importance of arts education—the community was already supportive—but there had been a big change in the ability to communicate the importance of arts education to her colleagues and

others directly involved in district decisionmaking and resource allocation (for example, school board members).

The one thing I always say to my teachers is that this is what central office people speak, this language of assessment and data. When I started speaking their language, they started to take me a lot more seriously. I was speaking the same data language to them, talking about student achievement in these measurable terms. That gave what I was saying more validity.

Teaching arts educators to speak “the same data language” as their colleagues in other disciplines may have other benefits as noted in the comments below from an Adoption Model district interview.

The math department is learning from arts teachers about how to make standards more plain spoken to parents and others. There has been growth in the professionalization of arts teachers—they are now speaking the same language as science and math—teaching to standards to reach measurable learning goals. This has led to more respect from other teachers, because we have standards and assessments, and has put arts teachers on the same professional level as other core subjects.

Internal communication of arts assessment system data can also be instrumental in instructional improvement efforts by professional learning communities of arts educators within a district, as noted by an interviewee from an Adoption Model district.

We are now taking the next step and using the assessment data to develop priority standards which are basically power standards where you take the state standards and work with them in professional learning communities.

The use of assessment data to support professional development is an important part of the integration of the arts assessment system with broader systems of accountability and instructional improvement described in the next section.

Exhibit 3 below provides a descriptive summary of the potential strengths and weaknesses of each of the three assessment system models relative to the five criteria for balance in a districtwide arts assessment system.

**Exhibit 3
District Assessment System Model Strengths and Weaknesses**

	Comprehensiveness Balance of fidelity and bandwidth	Coherence Alignment to standards	Continuity Articulation across levels	Feasibility Capacity alignment	Visibility External & internal communication
<p>Adoption Model</p> <p>District selects and implements assessments of arts learning developed at state, regional, or county level.</p>	<p>Moderate fidelity to content of district arts instruction.</p> <p>Strong bandwidth; all four arts covered at all levels.</p>	<p>Strong alignment to state arts standards in all four arts subjects.</p> <p>Moderate to weak alignment to district curriculum and instruction.</p>	<p>Strong articulation of benchmarks for performance across all levels in each of the four arts subjects.</p>	<p>Strong state leadership, resources and support needed.</p> <p>Supports state, district- and school-level capacity building.</p>	<p>Strong external and internal visibility.</p> <p>Student learning in the arts reported to state, districts, and schools on common assessments aligned with state standards.</p>
<p>Development Model</p> <p>District develops and implements common assessments of arts learning for use districtwide.</p>	<p>Moderate fidelity to content of district arts instruction.</p> <p>Moderate to weak bandwidth; all four arts may not be covered at all levels.</p>	<p>Moderate to weak alignment to state arts standards.</p> <p>Strong to moderate alignment with district arts curriculum and instruction.</p>	<p>Moderate articulation of benchmarks for performance across all levels; some arts subjects may not be articulated across levels.</p>	<p>Strong district leadership, resources and support needed.</p> <p>Supports district- and school-level capacity building.</p>	<p>Moderate external visibility; strong internal visibility.</p> <p>Student learning in the arts reported to district and schools on common assessments aligned with district arts curriculum.</p>
<p>Benchmark Model</p> <p>District develops and implements common benchmarks and/or rubrics for teachers to use in developing assessments of student arts learning.</p>	<p>Strong fidelity to content of district arts instruction.</p> <p>Weak bandwidth; coverage of arts subjects may be limited or absent below high school.</p>	<p>Moderate to weak alignment to state arts standards.</p> <p>Moderate to weak alignment with district arts curriculum and instruction.</p>	<p>Weak articulation of benchmarks for performance across levels; may be limited articulation across grades or courses within level.</p>	<p>Strong school or department leadership, resources, and support needed.</p> <p>Supports school level or department level capacity building.</p>	<p>Weak external visibility; moderate internal visibility.</p> <p>Student learning in the arts reported to district and schools on benchmarks aligned with arts course content.</p>

INTEGRATION WITH ACCOUNTABILITY AND IMPROVEMENT SYSTEMS

The 2007 report from the California Arts Education Strategic Task Force recommends voluntary district-level arts assessment programs that “help school districts hold themselves accountable for the quality of standards-based arts education programs...help teachers adapt their instruction to enhance student achievement...[and] help administrators make effective decisions about instruction, personnel, and resources needed for arts programs” (California County Superintendent Educational Services Association, 2007, p. 9). More recently, the 2008 report from the California Alliance for Arts Education, *Accountability in Arts Education: Building a Statewide System of Reciprocity*, calls for a commitment to full implementation of existing state arts education policies in California and a “broader accountability for arts learning—a coordinated system of synergistic commitments, policies, and practices that motivate responsibility and responsiveness among individuals and institutions” (p. 21). The arts education accountability infrastructure envisioned by the report includes policies to ensure adequate state and local funding, effective reporting mechanisms, content standards and instructional materials adoption, qualified teachers, and equitable scheduling and enrollment. The envisioned infrastructure also includes key practices: needs assessment and strategic planning; quality curricula, professional development and instruction; and effective assessment of arts learning. It is within this broader context of accountability that design decisions and plans should be made for districtwide assessment systems for arts education in California school districts.

Creating a reciprocal accountability system for arts education in California will require coordinated action by the state, districts, and supporting agencies at the state, regional, and local levels. The Adoption Model for a districtwide arts education assessment system seems to offer the greatest potential for a balanced distribution of responsibility and accountability for arts education at the district, regional, and state levels. In this model, the state takes on primary responsibility and can be held accountable for the development of prototype arts assessments aligned with state standards and for the technical assistance and professional development resources and materials needed to guide district-level design and implementation of arts assessment systems. The Development Model shifts the primary locus of responsibility and accountability for quality in a sequential, standards-based K-12 arts education program to the district level. As a result, the capacity and likelihood of designing and implementing a reasonably comprehensive, coherent, continuous, feasible, and visible districtwide arts education assessment system will vary greatly depending upon the level of resources and support available within the district. In this model, an arts education “Matthew Effect”⁸ may occur in which resource-rich districts use arts assessment data to support continuous improvement and expansion while resource-poor districts, lacking comparable data on student performance in the district, may experience a decline in the scope and quality of their arts education programs (see Woodworth, Campbell, Bland, & Mayes, 2009). Finally, the Benchmark Model shifts the locus of responsibility and accountability for the quality of the arts education program to the school, department, or individual teacher level. In this model, the “Matthew Effect” may operate within

⁸ Referring to the biblical idea that the rich get richer and poor get poorer. Reading researcher Keith Stanovich used the term “Matthew Effect” to describe the cumulative effects of poor reading proficiency in elementary school causing students to fall further and further behind in later grades.

the district as pockets of excellence in arts education at individual schools or in particular arts departments use assessment data to continuously improve while weaker schools and departments within the district grow weaker still.

Though an Adoption Model assessment system may be relatively less likely to exacerbate existing inequities in the quality and resources available for arts education across districts and within districts, no districtwide assessment system model can, by itself, expand access to high-quality standards-based sequential arts instruction. Districtwide assessment systems of all three types have the greatest potential to support expanding and improving quality of a district arts education if they are well integrated within broader systems for arts education program planning, program quality monitoring and improvement, and professional development.

A well-designed districtwide assessment system for arts education can provide key data and support to accountability and program improvement systems such as comprehensive arts education program audits,⁹ teaching quality and professional development appraisals,¹⁰ and professional learning communities.¹¹

- **Program Audits** systems based use comprehensive program quality indicators to evaluate the quality of an arts education program across a broad range of domains (such as arts curriculum, instruction, professional development, administrative support, and so on), with scoring criteria to evaluate levels of quality of the arts education program in each domain.

Evaluation of resources and capacity for assessing student learning in the arts should be an integral part of any comprehensive audit of a district arts education program. Such audits should, as far as possible, incorporate evaluation of the comprehensiveness, cohesiveness, continuity, feasibility, and visibility of the districtwide arts assessment system. Conversely, a strong and balanced districtwide arts assessment system should provide timely reports on student performance in the arts that can be easily fed into and combined with other evidence of arts program quality in a comprehensive program auditing process. Results of districtwide assessments can be used to pinpoint areas of strength and weakness in district arts instruction and to evaluate the degree of alignment of the district curriculum with arts standards.

- **Teaching Quality Appraisal** systems use indicators of instructional quality and teacher professional development to evaluate district arts instruction and identify professional development needs by assessing opportunities for teacher professional development, alignment of arts teaching with arts standards, and other indicators of teaching quality.

⁹ See, for example, M. Burt, E. Lindsley, & D. P. Russell, *Insider's Guide to Arts Education Planning, 2nd Edition* (2009), California Alliance for Arts Education, available at <http://www.artsed411.org>. Also see, Kennedy Center Alliance for Arts Education Network, *A Community Audit for Arts Education* (2007), available at <http://www.kennedy-center.org/education/kcaen/resources/CAudit6-9.pdf>.

¹⁰ See, for example, guidelines for evaluating elementary, middle school, and high school fine arts teachers in the Texas Professional Development Appraisal System, available online at <http://www.cedfa.org/growing/index.php?file=pdas.php>.

¹¹ See, for example, Gail Burnaford's (2006) report on action research in professional learning communities supported by the Chicago Arts Partnerships in Education (CAPE), *Moving Toward a Culture of Evidence: Documentation and Action Research Inside CAPE Veteran Partnerships* (2006), available at http://www.capeweb.org/cape_research/gb_vet.pdf.

Evaluation of teachers' understandings and abilities to administer and interpret districtwide assessments of student learning in the arts should be an important component of any teacher quality and professional development needs assessment. Districtwide assessment results can provide evidence of teachers' ability to teach to arts standards and can assist in planning for professional development and district allocation of materials and resources to support instructional improvement.

- **Professional Learning Communities** are systems of enhanced professional communications, often supported by technology to support sharing of resources, collaboration, and development of professional learning communities among arts educators (within and beyond the district) often through applications of information technology resources to support digital documentation and networking.

Collecting and reporting information on student learning in the arts through a districtwide assessment system provides information that is meaningful and of high value to teachers throughout the district. If the districtwide assessments are well aligned to state and/or national standards, the professional community within which results on the district assessments are meaningful can be large and diverse. Giving teachers the ability to share information on instructional practices and evidence of student learning among a group of professional peers is of tremendous value in facilitating teacher learning and instructional improvement. Digital media and online communications have opened up many new possibilities for recording and sharing arts instructional processes and results.¹²

¹² See, for example, CAPE's multimedia teacher portfolios system at <http://www.capeweb.org/rexamples.html>. Also, see examples of digital portfolios at Digication (<http://www.digication.com/>) and Richer Picture (<http://www.richerpicture.com/>).

FINDINGS AND IMPLICATIONS

General findings and implications of this review of potential models for the design and implementation of districtwide arts assessment systems in California include the following.

Comprehensiveness: Capturing Evidence of Important Arts Learning

Comprehensiveness in districtwide assessment systems for arts education can help to ensure that California school children are being well prepared for college and careers in the 21st century. Design features that maximize comprehensiveness in a districtwide assessment system include the following:

- State or regional development of prototype assessments aligned with standards in all arts areas (including emerging areas of arts learning aligned with 21st-century college and career readiness).
- State or regional agencies provide technical assistance to districts on design or selection and on implementation of performance assessments in all arts areas, including techniques and tools for assessing individual performance ability in the context of group performance (particularly in the performing arts—music, dance, and theatre).
- District participation in state or regional level collaboration for ongoing development, review, and revision of arts assessments.

Coherence: Enhancing Standards-Based Arts Instruction

Coherence in an arts assessment system can help California school districts monitor and maintain standards-based arts education curriculum and instruction throughout the district's schools and classrooms. Design features that maximize coherence in a districtwide assessment system include the following:

- State standards for the arts are well written to facilitate alignment of curricula, standards, and assessments.
- State standards for the arts are updated and revised to reflect emerging understandings of artistic knowledge, skills, abilities, and habits of mind needed for success in 21st-century careers (including, for example, design skills, studio thinking,¹³ and competencies in using emerging media and technologies).
- Teachers participate in district selection of standards to be assessed and in the selection or design of assessments best aligned with district curricula and instruction.

Continuity: Enhancing Sequential Arts Learning

Continuity in an arts assessment system can provide key support for standards-based, sequential arts learning across grade levels in California school districts. Design features that maximize continuity in a districtwide assessment system include the following:

¹³ See L. Hetland, E. Winner, S. Veenema, and K. M. Sheridan, *Studio Thinking: The Real Benefits of Visual Arts Education*. (New York, NY: Teachers College Press, 2007).

- Districtwide assessments are well aligned with articulated sequences of skills development and learning in arts subjects across grade levels from elementary to middle school to high school.
- Districtwide assessment development and implementation is used to identify and close gaps in arts teaching and learning across grade levels.

Feasibility: Building Capacity for Arts Education

In considering the feasibility criterion, California districts should not focus solely on constraints imposed by existing capacity (limited human and material resources) within the district to implement an arts assessment system, but rather should consider the potential for building district capacity for arts education through assessment of student learning and program quality. A strong districtwide assessment system coupled with support for program improvement are the essential ingredients of a reciprocal accountability system that over time can strengthen the capacity of California districts to deliver high-quality arts education programs. Design features that maximize feasibility (capacity alignment) include:

- Teachers and administrators are actively engaged in the design and implementation of the districtwide arts assessment system.
- Districtwide arts assessment results are used to identify areas of need for professional development and technical assistance.
- Design and implementation of districtwide arts assessments facilitate the development of professional learning communities among arts educators.

Visibility: Communicating the Value of Arts Education

California school districts will want to give particular attention to the visibility criterion in designing a districtwide arts assessment system. To foster and preserve public support for arts education, California districts will need to take steps to effectively communicate the quality and value of standards-based sequential arts education to a broad range of stakeholders. Design features that maximize visibility include the following:

- Results from the districtwide assessment system are reported to both external (state, parents, general public) and internal (teachers, administrators, school boards) audiences.
- Assessment results are communicated in ways that enhance external and internal audiences' perceptions of the quality of arts education programs and the value of arts learning.

Findings from our earlier *An Unfinished Canvas* review (Stites & Malin, 2008) of large-scale arts assessment systems also have important implications for the design and implementation of districtwide arts assessment systems. One key finding of our earlier review of state-level assessment systems was that the purposes of accountability and instructional improvement are best served by a broad range of evidence of student learning in the arts. Using multiple measures of student learning is a common standard for quality in any assessment system. Including a range of assessment formats, from selected-response formats (for example, a multiple-choice test on musical notation) to complex, performance-based assessment tasks (for example, an ensemble musical performance) to portfolio formats (for example, a digital recording of a series of individual musical performances) can greatly enhance the comprehensiveness, coherence, and continuity of any districtwide arts assessment system.

Finally, when we put the findings our earlier *An Unfinished Canvas* review of approaches to state-scale assessment of arts learning together with the findings from our current review of

district-level models we see clearly the critically important roles of strong teacher professional development and adequate external supports and technical assistance in the successful design and implementation of any districtwide arts assessment system.

Assessment systems that are well integrated with teacher professional development and instructional improvement systems have the strongest potential for increasing district capacity to implement high-quality standards-based sequential arts instruction.

A districtwide arts education assessment system should produce evidence of student learning in the arts that is in accessible formats and can provide meaningful guidance for planning ongoing teacher professional development and informing a continuous improvement process for the arts education program. Alignment of districtwide assessments with important standards and key curriculum content will make reporting and interpretation of assessment results more accessible and more meaningful for these purposes.

Adequate external support (in the form of well-written state standards, arts assessment prototypes, professional development resources, and technical assistance) from the state and from regional and county agencies and organizations will be critical for success in the design and implementation of district arts assessment and accountability systems.

The importance of support from the state and from regional or county-level educational agencies in building quality in districtwide arts assessment systems cannot be overemphasized. Well-written state standards that are subjected to ongoing revision and improvement are critically important to the success of each of the three models for the design of districtwide assessment systems described in this report. Not only must state arts standards be easily translated into measurable learning goals, but they must also be continuously updated to reflect developing understandings of learning and achievement in each arts discipline as well as the knowledge, skills, and abilities needed for success in the 21st century workplace and economy.¹⁴ Developing good measures of standards and student learning in the arts poses technical challenges and costs that are beyond the capacity of all but the largest and most resource-rich districts. Economies of scale favor state-level or regional-level development of professional development and technical assistance resources and infrastructure.

¹⁴ See the Partnership for 21st Century Skills framework at <http://www.21stcenturyskills.org/index.php>.

REFERENCES

- Burnaford, G. (2006). *Moving toward a culture of evidence: Documentation and action research inside CAPE Veteran Partnerships: Veteran Partnership Reports 2005-2006* Chicago, IL: Gail Burnaford and CAPE. Available at http://www.capeweb.org/cape_research/gb_vet.pdf
- Burt, M., Lindsley, E., & Russell, D. P. (2009). *Insider's guide to arts education planning (2nd edition)*. Pasadena, CA: California Alliance for Arts Education. Available at <http://artsed411.org/insidersguide/index.aspx>
- California County Superintendents Educational Services Association. (2007). *California Arts Education Strategic Task Force report: Findings and recommendations 2007*. Sacramento, CA: Author.
- Gallagher, H. A., Campbell, A. Z., Esch, C. E., Malin, H., Mayes, N., & Woodworth, K. R. (2008). *An unfinished canvas. Allocating funding and instructional time for elementary arts education*. Menlo Park, CA: SRI International. Available at <http://policyweb.sri.com/cep/projects/displayProject.jsp?Nick=artsed>
- Guha, R., Woodworth, K. R., Kim, D., Malin, H., & Park, J. (2008). *An unfinished canvas. Teacher preparation, instructional delivery, and professional development in the arts*. Menlo Park, CA: SRI International. Available at <http://policyweb.sri.com/cep/projects/displayProject.jsp?Nick=artsed>
- Hetland, L., Winner, E., Veenema, S., & Sheridan, K. M. (2007). *Studio thinking: The real benefits of visual arts education*. New York, NY: Teachers College Press.
- Kennedy Center Alliance for Arts Education Network. (2007). *A community audit for arts education*. Available at <http://www.kennedy-center.org/education/kcaaen/resources/CAudit6-9.pdf>
- Landon, J., & Russell, D. P. (2008). *Accountability in arts education: Building a statewide system of reciprocity*. Pasadena, CA: California Alliance for Arts Education. Available at http://www.artsed411.org/advocate/docs/CAAllianceforArtsEd_AccountabilityArtsEd_2008.pdf
- National Endowment for the Arts. (2004). *Outcome-based evaluation: A working model for arts projects*. Washington, DC: National Endowment for the Arts. Available at <http://www.arts.gov/grants/apply/out/index-out.html>
- National Research Council. (2001). *Knowing what students know: The science and design of educational assessment*. Committee on the Foundations of Assessment, J. W. Pellegrino, N. Chudowsky, & R. Glaser (Eds.), Center for Education. Division on Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

- The New England Arts Assessment Network. (June, 2006). *Arts assessment in New England: Narrative accounts of existing projects*. Available at http://www.ed.state.nh.us/EDUCATION/doe/organization/curriculum/Arts/Arts_Assessment_in_New_England_-_Aug_2006.pdf
- Partnership for 21st Century Skills. (2009). *P21 framework definitions document*. Washington, DC: Partnership for 21st Century Skills. Available at http://www.21stcenturyskills.org/documents/p21_framework_definitions_052909.pdf
- San Diego County Office of Education (2008). *Arts assessment resource guide*. Sacramento, CA: California County Superintendents Educational Services Association. Available at http://www.ccsesaarts.org/content/assessment_guide.asp
- Stites, R., & Malin, H. (2008). *An unfinished canvas: A review of large-scale assessment in K-12 arts education*. Menlo Park, CA: SRI International. Available at <http://policyweb.sri.com/cep/projects/displayProject.jsp?Nick=artsed>
- Texas Education Agency. (2002). *The professional development and appraisal system: Fine arts teachers*. Austin, TX: Texas Education Agency. Available at <http://www.cedfa.org/growing/index.php?file=pdas.php>
- Woodworth, K. R., Campbell, A. Z., Bland, J. A., & Mayes, N. L. (2009). *An unfinished canvas. District capacity and the use of new funds for arts education in California*. Menlo Park, CA: SRI International. Available at <http://policyweb.sri.com/cep/projects/displayProject.jsp?Nick=artsed>
- Woodworth, K. R., Gallagher, H. A., Guha, R., Campbell, A. Z., Lopez-Torkos, A. M., & Kim, D. (2007). *An unfinished canvas. Arts education in California: Taking stock of policies and practices. Summary report*. Menlo Park, CA: SRI International. Available at <http://policyweb.sri.com/cep/projects/displayProject.jsp?Nick=artsed>
- Woodworth, K. R., Petersen, D. M., Kim, D. H., & Tse, V. (2009). *An unfinished canvas. Local partnerships in support of arts education in California*. Menlo Park, CA: SRI International. Available at <http://policyweb.sri.com/cep/projects/displayProject.jsp?Nick=artsed>

INTERVIEW PROTOCOLS

TOPICS FOR FIRST ROUND TELEPHONE INTERVIEWS

The purpose of the initial round of telephone interviews was to gather information on three topics: identifying districts with exemplary arts assessment systems, locating research and descriptive literature on assessment of student learning in the arts that is relevant to district level arts assessment systems, and identifying key issues and challenges in developing and implementing district arts assessments.

We contacted each interviewee by e-mail to arrange a time for a brief telephone interview (30 to 45 minutes with follow-up by phone and/or e-mail as needed). A short description of the purpose of the review and topics for the interview was included in the initial e-mail.

Topic 1: Exemplary District Arts Assessment Systems

- Where would we find the best existing examples of district-level assessment of student learning in the arts?
 - Which states?
 - Which districts?
 - Which arts subjects are assessed (visual arts, dance, music, theater, integrated)?
 - What purposes does assessment serve (formative/inform instruction, benchmark/monitor learning progress, summative/measure achievement)?

Topic 2: Information on Promising Models and Approaches

- Where would we find information (research or documentation) on promising models or approaches (not yet implemented) to district-level arts assessment?
 - Authors, organizations?
 - Journals, websites, etc.?
 - Models for assessing which arts subjects (visual arts, dance, music, theater, integrated)?
 - Models for which assessment purpose (formative/inform instruction, benchmark/monitor learning progress, summative/measure achievement)?

Topic 3: Key Issues and Challenges for District Arts Assessment

- What are the most pressing issues that need to be addressed to help districts in California develop and implement district assessments of student learning in the arts?
 - Policy
 - Teacher assessment skills and knowledge
 - District staffing
 - External support (regional, state)
 - Resources (funding, quality products, IT, other)

TOPICS FOR SECOND ROUND OF TELEPHONE INTERVIEWS

Topic 1: General Background Information

Topic 2: Comprehensiveness

- Are all four arts disciplines equally represented?
- Does implementation differ from one arts area to another? How? Does it differ by grade level? How?
- How well focused are assessments to align with the most important aspects of arts learning in each discipline?
- How well covered are the full range of arts standards?

Topic 3: Coherence

- What is the state of alignment of assessments to curriculum and instruction?

Topic 4: Continuity

- How well connected are assessments to a sequential, standards-based curriculum across grade levels?

Topic 5: Feasibility

- What issues in district/school/teacher capacity to implement assessments have arisen? What has been done to address these?

Topic 6: Visibility

- How well have results of assessments been communicated to external audiences? What has been done to increase visibility? Use of technology?