



Created by ETS. Delivered by Questar.

RESEARCH REPORT

The *Winsight*[®] Assessment System: Contributions to Promoting Educational Equity

E. Caroline Wylie and Catherine M. Millett
ETS, Princeton, N.J.

November 2018

Table of Contents

Introduction	4
Inequities in the Educational System.....	6
<i>Winsight</i> ® Assessment Principles to Support Educational Equity	7
The Winsight Assessment System.....	9
The Three Assessment Components: Summative, Interim, Formative	10
Teachers: The Lynchpin in the Assessment System	12
The Winsight Approach to Equity	13
A. Good Assessment Practices: Part of All ETS Design and Development Work.....	13
B. Specific Winsight Design Features.....	15
C. Future Winsight Features or Components	18
Conclusion.....	20
Figure 1. Winsight Assessment System Overview.	10
Figure 2. Winsight System Assessment Supports.	12
Figure 3. Three-tiered Approach to Addressing Traditionally Underserved Groups.	13
Table 1: Summary of Winsight System Components that Address Equity Concerns.....	19

INTRODUCTION

Educational equity means that every student has access to the educational resources and rigor they need at the right moment in their education across race, gender, ethnicity, language, disability, sexual orientation, family background, and/or family income.¹

Within the U.S. educational system, roughly 50.4 million students attend public elementary and secondary schools and 5.2 million students attend private elementary and secondary schools.² Achieving educational equity requires the collective effort of many individuals and organizations, from federal and state levels of government, to efforts at the local level. We are privileged to be at an educational organization whose mission and daily work are dedicated to working with teachers, students, school officials, governments (local, state, and national), and other stakeholders to advance educational equity.

Our view is that assessment is one mechanism that can contribute to achieving educational equity. We consider assessment to play a key role in an educational system in which the goal is to provide the appropriate resources (e.g., learning opportunities, tools and materials, teachers with appropriate training) to the right students in the moment that they need them. At the same time, students need to be provided with assessments and feedback that promote and encourage the kind of learning opportunities that both appropriately meet students' needs and are sufficiently rigorous to enable them to maximize their potential. We believe that good assessment systems should be useful for all students regardless of where they reside, what school they attend, where they are in their learning, or their demographics.

Students do not experience their education in a vacuum. Students live and learn in a multifaceted system, and outcomes reflect many inputs and factors. As a result, information is needed to understand student learning at both a micro level of detail in order to make appropriate and meaningful decisions for individual students, and at an aggregate level, with information to make curriculum, resource, and policy decisions. Information can be used to understand student performance at the aggregate level in order to inform more local decisions such as how to deploy coaches in the next quarter, or what professional development is needed. Examining student performance in aggregate and by subgroups can be used to identify disparities, monitor progress, and identify policies and resources that can improve student learning outcomes.

The *Winsight*[®] Assessment System is a new ETS investment focused on developing a balanced assessment system that incorporates summative, interim, and formative components in order to provide meaningful and interpretable information for all stakeholders (students; parents or

¹ The Aspen Education & Society Program and the Council of Chief State School Officers. 2017. *Leading for Equity: Opportunities for State Education Chiefs*. Washington, D.C.

² <https://nces.ed.gov/fastfacts/display.asp?id=372>

caregivers; teachers; principals; and district, state, and federal administrators and policymakers). The Winsight summative component is intended to provide a strong signal regarding what is valued in College- and Career-Readiness state standards, which can, in turn, inform the design of rigorous and relevant teaching and learning activities. The interim component provides additional assessment opportunities to monitor student progress, and the formative component aims to support teachers and students unpacking critical ideas central to the domain. A critical feature of the Winsight Assessment System is accessibility. The Winsight Assessment System will provide a robust set of accessibility features, accommodations, and resources across the system for all students, and in particular for students who typically have access challenges, such as students with disabilities and students who are English learners (who also may be speaking or learning two or more languages — that is, dual-language learners, multilingual learners), thus increasing the likelihood that teachers will have accurate insights into what all students really know, understand, and can do.

We also believe that, as part of the development of an assessment system, it is important to go beyond supporting the thoughtful and appropriate use of assessment data. We recognize that there are inequities in the educational system and society that are outside

our direct influence, but we can support students beyond accessible assessments that promote rigorous instruction and meaningful data and feedback, by:

- developing teacher learning through the provision of assessment models that exemplify the breadth, richness, and rigor of state standards;
- providing professional learning opportunities for teachers to deepen their understanding of how all students learn; and
- developing reports that provide actionable information, not just score or subscores.

In the sections that follow, we present a set of guiding principles that have been important to the development of the Winsight Assessment System to date and will guide future development and delivery, describe the components of the system, and outline how these principles have been enacted to date. The purpose of this paper is to present a vision for how the Winsight Assessment System promotes equity, along with both the short-term and longer-term steps that we are taking to develop this equitable assessment system. Our hope is that the sharing of this information will generate conversations, input, and guidance from critical decision makers and stakeholders that will inform the future development of this important initiative.

INEQUITIES IN THE EDUCATIONAL SYSTEM

Historically, educational opportunities have not been equally available for all students. Such opportunities have been denied to many students who are economically disadvantaged, are from single-parent families, or have parents who have lower levels of education. Educational research has well-documented findings of barriers to access and equitable opportunities for many students who are underserved. While we all love to hear the inspiring anecdotes of the few students who “are beating the odds,” the fact remains that for many students their K–12 education experiences are of repeated struggles and lack of support from the educational system. Under the federal law known as the Every Student Succeeds Act (ESSA), four groups are identified for achievement-monitoring purposes. These groups are made up of traditionally underserved students — and all of these groups are overrepresented in the population of students who are not “on track” for college and career readiness:

- students with disabilities;
- students who are English learners;
- students from major racial and ethnic groups; and
- economically disadvantaged students.

We focus on these four groups in particular because information about them is generally available for monitoring and analysis across

systems at the state and federal level. We recognize that there are other marginalized groups in schools that are treated with inequity and unfairness, discrimination, and prejudice.

By way of evidence of disadvantage, we can look to many sources, but the National Assessment of Educational Progress (NAEP), often referred to as “The Nation’s Report Card,” provides a robust and valid source. NAEP is overseen by the Commissioner of the National Center for Education Statistics in the U.S. Department of Education and applies a standard measuring stick to a nationally representative sample of U.S. students in all 50 states and territories. Its data reflect what U.S. students know and can do and can also be used to track group trends on cross-sectional data over time. While there have been shifts in performance and in achievement gaps over time, in this context it is arguably the differences in performance between groups that are most important because these group differences translate into real differences in life chances for individuals. While NAEP does not assess students beyond high school, the research is clear about the positive impact on earnings of those with a college degree compared to those with only a high school diploma,³ and shows that students of color are less likely to attain either a two-year or four-year degree compared to their White counterparts.⁴

For each of the four groups of students identified in ESSA, NAEP results by grade and for both reading and mathematics consistently tell the same story: the

³ Carnevale, A. P., Smith, N., & Strohl, J. (2010). *Help wanted: Projections of job and education requirements through 2018*. Washington, DC: Georgetown Center on Education and the Workforce.

⁴ Nettles, M. (2017). *Challenges and opportunities in achieving the national postsecondary degree attainment goals* (Research Report No. RR-17-38). Princeton, NJ: Educational Testing Service. <https://doi.org/10.1002/ets2.12141>

performance distribution for students in the four target groups is significantly different from that of those who are not — across grade levels and content areas. A greater proportion of students in the target groups are classified as “below basic” compared to students who are not in the target groups. At the upper end of the proficiency scale, the opposite is true, with a much smaller proportion of students in the target groups classified as “advanced” compared to students not in those groups. In other words, if students are poor (categorized on NAEP as “low socioeconomic status”), have a disability, are English learners, or are African American, American Indian/Native Alaskan, or Hispanic, they are more likely to be classified in the “below basic” category than their wealthier, nondisabled, English-speaking, White, and Asian-American counterparts.

WINSIGHT ASSESSMENT PRINCIPLES TO SUPPORT EDUCATIONAL EQUITY

The overarching principle we developed to guide our work to support educational equity for the Winsight Assessment System is:

Equity in assessment is not the same as equality: students with the greatest needs and least resources require significant efforts and means so that they are provided with opportunities to support their learning and to demonstrate what they know and can do. Pursuing equity in assessment requires stakeholders to simultaneously strive to improve the educational outcomes for all students, whether focusing on those with lower academic outcomes and fewer opportunities to accelerate their growth or supporting high-achieving students from traditionally underserved groups who are often overlooked.

This idea can then be further unpacked into three categories of principles that focus on Winsight as an assessment system within a larger educational system; Winsight design and development; and Winsight reporting, interpretation, and use of results.

Winsight as a Systemic Lever

1. Assessment is one mechanism to document inequities in educational outcomes in order to motivate and inform efforts to narrow or eliminate them. Use of appropriately disaggregated assessment data and reports can provide educational systems with one form of insight into their system that can be used to inform their planning.
2. Assessments should contribute to educational equity by assisting stakeholders in identifying those who face the greatest educational obstacles and providing insight into how to help those students overcome those obstacles. At the school and school district levels, this can inform allocation of resources (teachers and other supports).

3. A focus on equity requires a focus on specific traditionally underserved groups. Not all individuals in a traditionally underserved group will have had some of the disadvantages experienced by many members of that group. However, these exceptions do not negate the fact that the group as a whole has been underserved in ways that can be measured. Assessment data can inform how school districts develop and implement supports to serve the greatest number of students.
4. Equitable assessment is one part of a larger system of educational equity. It is important to identify and create related supports around the assessment components that positively impact the larger system, such as supports for teacher learning and development, parent and guardian engagement, and guidance for appropriate use of score reports.

Winsight Design and Development

5. Students' experiences with assessments should give them a sense of accomplishment and allow them to demonstrate what they know and can do. Assessment developers are tasked with fashioning assessments that provide these types of opportunities for all students.
6. A diverse pool of assessment developers, external reviewers, advisors, and pilot test takers need to be part of the assessment creation process to advance the creation of assessments that reflect the breadth of test takers' experiences and aspirations.
7. Since many students speak more than one language,⁵ an assessment system must seek to support the skills of dual- and multi-language speakers. We must take seriously the issue of language accessibility in order to support their learning.
8. To address the diversity of students in terms of cognitive, sensory, and physical abilities, an assessment system must use careful design principles, including accessibility and accommodation features, so that the assessment is not a barrier to students demonstrating what they know and can do.

Winsight Reporting, Interpretation, and Score Use

9. Assessments must support valid interpretations appropriate for their intended purposes; provide reliable score information for all students; and be fair in the sense of both reducing construct-irrelevant variance and making available a broad range of accommodations and accessibility features so that all students are able to demonstrate their knowledge and skill without hindrance.
10. Assessment reports and the support structures around reporting should inform and empower stakeholders by shining a spotlight on areas of both needs and strengths.

⁵ The U.S. Census Bureau reports that at least 350 languages are spoken in U.S. homes (<https://www.census.gov/newsroom/press-releases/2015/cb15-185.html>).

Reports should be developed so that they communicate important ideas to stakeholders without them requiring special training to interpret the reports.

We recognize that the issue of supporting each and every student cannot be addressed by a single or simple solution: guiding principles are not sufficient. The challenges and needs of students learning English are quite different from students who have visual impairments, for example. Furthermore, the approaches needed to address issues of access to the assessment are themselves quite different from approaches needed to address more complex problems faced by many students of color or low-income students whose educational performance is impacted by an intricate set of interrelated factors. The various strategies and solutions need to be used together since the categories of traditionally underserved students are not mutually exclusive, and many students fall into more than one category. In the sections that follow, we first describe the Winsight Assessment System and then the three-tier approach to equity that we are taking to operationalize these principles to meet the wide range of student needs.

THE WINSIGHT ASSESSMENT SYSTEM

The Winsight Assessment System integrates summative,⁶ interim (including both benchmark assessments and testlets), and formative assessment components, initially focused on mathematics and English language arts (ELA) in grades 3–8 and high school (see Figure 1) to provide a coherent and aligned assessment system.

The system is intended to make what students know and can do more visible to all stakeholders, including the students themselves, as well as parents, teachers, and administrators. The Winsight Assessment System is intended to use each state's standards in conjunction with research-based learning progressions and key practices not just to identify students' knowledge, skills, and abilities but also to provide additional information about students' learning trajectories and what the appropriate next instructional steps might be.

Learning progressions: a research-based, qualitative description of students' thinking that begins with early conceptions and describes how they grow into formal ideas.

For a student who does not yet have a robust understanding of a particular mathematical topic, those early levels can provide a means for describing the key concepts and ideas he/she does have. From an equity perspective, this is particularly important for lower-performing students, as it provides a way to shift away from deficit language. For more information, see www.ets.org/research/policy_research_reports/focus_on_rd/issue3.

⁶ See glossary of terms at end of document.

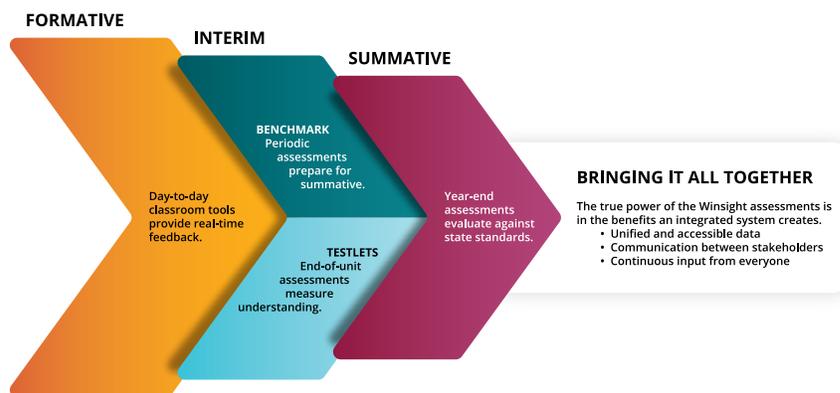


Figure 1. Winsight Assessment System overview.

The Three Assessment Components: Summative, Interim, Formative

Much of the development work for the summative and benchmark components has been completed in preparation for field testing prior to operational delivery starting in 2020. The testlets are currently in the design phase for a staggered delivery after the summative and benchmark components. The formative assessment component is at an earlier prototyping stage. As a result of the development work being completed in stages, some aspects of the work described below have been started while other aspects are in the planning stages.

The *Winsight summative assessment* is primarily intended to measure students' academic performance with respect to each state's standards in ELA and mathematics for annual accountability purposes. Individual student year-to-year growth is intended to be monitored using a vertical scale, which allows for comparisons to be made across grade levels. The assessment will provide individual student subscores as starting points for receiving classroom teachers' and administrators' follow-up. The online assessment will be built to balance efficiency of measurement with rigorous assessment of complex constructs by using *multistage adaptive testing*, a wide range of technology-enhanced items, and both human scoring and automated methods for scoring constructed-response items. The results (total scaled score, four or five subscores that represent major domains within the standards, and proficiency levels for all students and for subgroups) will be available in an online, interactive reporting system that can be used by state, district, and school administrators; teachers; and counselors to examine student progress toward college and career readiness in mathematics and ELA.

The *Winsight interim assessment* components are intended to be used by teachers for instructional design and modification and by administrators for monitoring student learning between annual summative administrations. The Winsight Assessment System separates interim assessment into two distinct components: benchmark assessments and testlets.

Interim benchmark assessments provide snapshots at the start, middle, and end of the year and are intended to be predictive of student performance on the summative assessment, recognizing that the predictive power of the benchmark assessment increases the closer in time that it is taken to the summative assessment. Benchmark assessments can be used primarily by teachers and administrators as monitoring tools because each benchmark assessment is built to the same blueprint as the summative assessment, albeit shorter.

Interim testlet assessments will be developed to target critical aspects of state standards, replacing what is often referred to in schools as “common assessments.” The testlets can be used at flexible intervals as best determined by the needs of classroom teachers as pretests for baseline information at the starts of units or after units of instruction have been completed (e.g., at the start or end of a unit focused on proportional reasoning in mathematics or argumentation in ELA). They will provide results that highlight progress on critical learning progressions and key practices. The testlet results can help teachers determine additional opportunities to reinforce ideas and concepts that students may not have fully mastered and to build on and extend areas of strength. The testlet results, mapped to targeted learning progressions, also provide a way for teachers to calibrate qualitatively their own classroom assessment judgments.

The *formative assessment component* of the Winsight Assessment System is the part that teachers and students may interact with most frequently, to support teachers’ and students’ ongoing formative assessment processes. Teachers will be able to use the formative assessment resources that model good teaching and learning practices to support ongoing instruction and to track student progress within tasks, assessing students’ competency development along a continuum of development from novice to expert. These resources will differ from the

interim assessment components in that they are intended to provide finer-grained information (compared to the testlets) for teachers and students. When the Winsight Assessment System is completed, we intend for there to be a range of formative resources, from quick checks-for-understanding to extensive scenario-based tasks that provide more scaffolded learning opportunities to inform immediate adjustments to teaching and learning. When the teacher needs to closely examine student responses to guide instructional decision making, an interface or dashboard will make the workflow seamless (e.g., show the teacher where students are within tasks, present student work that requires review, and prioritize student work to be reviewed based on their responses to automatically scored items). Winsight will use a technology platform to support student-to-student interactions (e.g., peer review, collaboration, metacognition, peer tutoring), which research has demonstrated can have powerful impacts on student learning (Black & William, 1998).⁷ Information in the formative setting will focus on helping teachers make sense of student responses in light of state standards, learning progressions, and key practices, and will provide guidance to the next instructional steps. While testlets support periodic assessment, the formative assessment resources will support teachers’ and students’ ongoing, daily assessment of learning.

⁷ Black, P., & William, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, policy & practice*, 5(1), 7–74.

Teachers: The Lynchpin in the Assessment System

Assessment systems do not operate in a vacuum. Every day, over 98,000 public schools⁸ teachers are in classrooms navigating assessment systems. ETS recognizes that merely providing the manuals and online resources that support test administrations is not sufficient. The Winsight Assessment System provides system-supporting resources around assessment literacy and should help teachers understand and make use of the learning progressions and key practices, both as part of Winsight assessment reports and to serve as teaching and learning supports (see Figure 2).

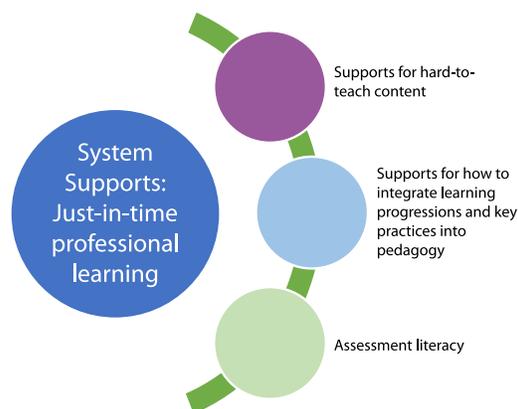


Figure 2. Winsight System Assessment supports.

The Winsight system supports are aligned with recommendations from the 2017 National Academies Press report, *Promoting the Educational Success of Children and Youth Learning English*,⁹ one of which specifically notes the importance of assessment literacy for teachers of dual-language learners/English learners (DLLs/ELs) and other relevant stakeholders.

We intend to provide rich professional learning resources for teachers. Resources are also intended to target hard-to-teach topics in College- and Career-Readiness standards, such as argumentation in ELA or mathematical modeling, and to enable teachers to make appropriate use of the learning progressions and key practices in their instruction, not just as part of the assessment system. We plan to provide resources that will model good teaching and learning practices (e.g., using rubrics as part of assessment tasks to illustrate ways to be clear about expectations to students), supported by just-in-time professional development opportunities (i.e., resources that teachers can access on demand, such as webinars, online videos, and resources for school-based professional learning communities). Use of these resources will be optional for teachers, and we will work with teachers during development to help ensure that they are meaningful and useful.

The various components within the system are intended to interact with each other in meaningful ways. The summative and benchmark assessments will use the same reporting scales and share item types. Accessibility and accommodations features will be consistent across components and, where appropriate, reports may direct stakeholders to other system components as a means of obtaining more targeted assessment information. In the next section, we describe in more detail our initial thinking about how Winsight can approach the equity challenge.

⁸ <https://nces.ed.gov/fastfacts/display.asp?id=84>

⁹ <https://www.nap.edu/catalog/24677/promoting-the-educational-success-of-children-and-youth-learning-english>

THE WINSIGHT APPROACH TO EQUITY

For the Winsight Assessment System, we have adopted a three-tiered approach to address the needs of all students, including those that are traditionally underserved (see Figure 3).

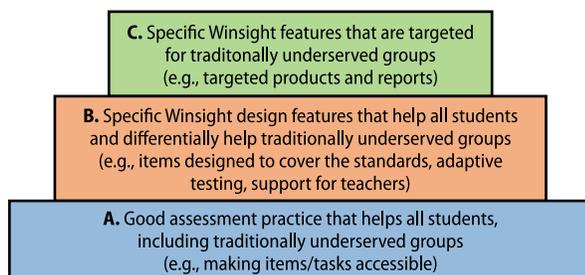


Figure 3. Three-tiered approach to addressing traditionally underserved groups.

A. Good Assessment Practices: Part of All ETS Design and Development Work

The principles discussed earlier, particularly around Winsight design and development (principles 5–9), are more specifically articulated in this section in terms of how we follow design and development practices laid out in the *ETS Standards for Fairness and Quality*¹⁰ as standard procedures that guide all assessment development — along with other guidelines that are specific to designing constructed-response items — for designing accessible tests and for keeping the needs of English learners in mind (see Further Reading for examples). In addition, new development at ETS, for assessments primarily to be delivered using online technologies, is focused on being *born accessible* — in other words, designed from the beginning to be accessible to all students including students with disabilities and students who are English learners (see bottom layer in Figure 3). This approach helps ensure that student performance is not hampered by an inability to access the assessment questions themselves. Furthermore, applying universal design processes supports all students (Hall, Sparks, Spero, & Van Kuren, 2005).¹¹ In other words, our intention, through careful design, is to support every student’s ability to demonstrate her or his best performance, which in turn gives policymakers, administrators, teachers, and students improved information about the student’s understanding of the content rather than information about his or her ability to interact with the assessment. Improved information, in turn, allows teachers to better differentiate instructional supports to the skill level of the student. Below we highlight the approaches that we follow to support the accessibility of the system and assessment components:

- **Evidence-Centered Design (ECD)** guides assessment development. For example, for the summative assessment, this approach focuses first on a clear identification of the constructs to be assessed, then on evidence for the constructs and approaches to eliciting that evidence.

¹⁰Educational Testing Service (2014). *ETS Standards for Fairness and Quality*. Princeton, NJ: Author.

¹¹Hall, T., Sparks, N., Spero, I., & Van Kuren, L. (2005). *Universal Design for Assessment*. *Soapbox Digest*, 4(1).

- **Universal Design approaches** guide assessment creation to maximize accessibility.
- **508 compliance:** Winsight follows the World Wide Web Consortium’s Web Content Accessibility Guidelines (WCAG 2.0¹²), which are becoming the definition of accessibility under section 508 of the Rehabilitation Act, and other relevant standards, including the IMS Global Learning Consortium’s Accessible Portable Item Profile (APIP). These standards ensure that Winsight will provide built-in accessible features and accommodations that are required by many states’ laws.
- **Accessibility experts** are working closely with the design and item-writing teams to identify methodologies to maximize accessibility from the outset, considering the range of cognitive, sensory, and physical abilities of students accessing these assessments.
- **Usability/cognitive load perspective** will be used to examine item/task presentation and general navigation to ensure that the presentation of the directions and the assessment are as intuitive as possible.
- **Accommodations to support ELs** will be included in the assessment components. For example, Winsight includes the option of providing mathematics translation initially in Spanish (in which students can toggle between Spanish and English versions of each item so that they can maximize their use of both languages to demonstrate their mathematics understanding) and additional language glossaries¹³ (key words explained in English or translated), as needed in state-specific contracts, will be provided. In broader research efforts, we will explore other language supports in combination with accommodations to better understand this approach to supporting ELs and the potential benefits (e.g., language supports to accompany reading texts, dual-language support, scaffolding tasks, various types of feedback).
- **Internal item reviews** are conducted for all items and tasks for accuracy, ease of reading, and fairness issues.
- **External reviews** of item specifications and item pools are provided by diverse external panels of experts, including teachers; state- and district-level stakeholders; and university faculty and researchers in the content areas who represent different views and life experiences. We also work with advisors who have specific expertise in working with students in poverty, ELs, students with disabilities, and students of color.
- **Comparability of experience:** Students will have access to a consistent set of built-in accessible features, accommodations, and resources across the system that are presented in a consistent manner. Because we provide the same accommodations across all tiers of the assessment system, students will not have to familiarize themselves with multiple systems and approaches.

¹² <https://www.w3.org/TR/WCAG20/>

¹³ In Fall 2015, Spanish was the home language of approximately 77.1 percent of all EL students in public K–12 schools. Arabic, Chinese, and Vietnamese were the next most common home languages (spoken by approximately 2.4 percent, 2.1 percent and 1.7 percent of all EL students, respectively). Source: <https://nces.ed.gov/fastfacts/display.asp?id=96>

- **Research basis for each assessment component:** Throughout the development process, each claim about a component in the Winsight theory of action¹⁴ will be supported by evidence gathered through research, pilot and field tests, special studies, and operational use. Broader research activity involves work such as exploring whether and how learning progressions — which are representations of how modal ways understanding develops — hold for all groups of students and examining ways in which we can better capture evidence of student learning and expand the supports provided for students and other stakeholders.

B. Specific Winsight Design Features

As reflected by the second category illustrated in the middle layer of Figure 3, we have undertaken a wide-ranging series of actions to create an assessment system that is more attuned to deeper student learning and, therefore, can provide better assessment information to inform teaching and learning. The innovative design efforts relate to the earlier design principles (5 through 8). Our approaches to how we report information — and support interpretation and use of the information — addresses principles 1 through 4 (Winsight as a Systemic Lever) and principles 9 and 10 (Winsight Reporting, Interpretation and Score Use). These design features will be of value to all students, but especially present the potential for a differential positive impact on traditionally underserved groups. In particular, the inclusion of the testlets and formative assessment components within the Winsight Assessment System is intended to make a significant effort to support improved teaching, learning, and assessment practices for teachers, and thus support the learning needs for all students.

Incorporating innovative design features

- **Moving beyond multiple choice** by including the broadest range of item and task types within the summative assessment (using innovations in machine scoring technology to supplement human scoring given constraints of cost and time) so that students are required to engage in complex thinking when providing responses. As a result, the summative assessment will be a more complete reflection of the standards and the information reported will be more actionable.
- **Using learning progressions** to inform the development of items that measure students' understanding across essential concepts within each grade level supports the development of assessment items that are accessible to students at a variety of levels. This approach supports providing opportunities for success for all students.¹⁵

¹⁴Wylie, E.C. (2017). *Winsight™ Assessment System: Preliminary Theory of Action*. ETS Research Report Series. doi:10.1002/ets2.12155

¹⁵Graf, E. A., van Rijn, P. W. (2015). Learning progressions as a guide for design: Recommendations based on observations from a mathematics assessment. *Handbook of test development*, Edition: 2nd, Chapter: 9, Publisher: Routledge, Editors: Suzanne Lane, Mark R. Raymond, Thomas M. Haladyna, pp.165–189.

- **Using cognitively based learning progressions, key practices, and reusable tools** as part of the summative-assessment ancillary materials and resources identifies and models important ideas and concepts in the standards through the summative assessment. These resources are intended to provide teachers and students with resources that they may find useful for classroom practice.
- **Providing interim assessment components** that balance the need for predictive assessment information (benchmark assessment) with assessments that are more reflective of local curriculum (testlets) to provide opportunities for more targeted understanding of students' strengths and weaknesses closer to the point of instruction. This enables immediate and targeted adjustments to curriculum and support.
- **Including a formative assessment component** that provides assessment resources to be integrated directly into instruction to provide opportunities for both students and teachers to assess students understanding and to provide guidance about appropriate next steps.
- **Supporting more precise measurement of students across the achievement continuum** by integrating the progressions that span grade levels into the assessment design, along with a multistage adaptive test delivery approach to routing students through the summative assessment. Based on performance on an initial block of items, students will be routed to a more targeted second item block, which will lead to more meaningful information at the extremes of the achievement continuum. Providing a better measure of students at the extremes will lead to better information about those who may need additional support and for those gifted and talented students who may need further challenge.
- **After initial launch of the summative assessment, future research will explore innovative designs**, such as routing students who perform very poorly on early stages of the ELA assessment to a more diagnostic reading measure or to a portion of the assessment that uses read-aloud to provide better information on students who have decoding-based disabilities. Similar routing options can also be explored for mathematics using computational and representational fluency measures to provide additional insights into low performance.

Providing information to students, parents, teachers, and other stakeholders sooner

- **Reporting to parents:** Providing reports to parents in multiple formats (e.g., paper and online reports with accompanying explanatory information and options for languages other than English) to increase accessibility to the information and provide a springboard for parent-teacher-student conversations.
- **Providing meaningful information earlier:** Using the progressions and key practices to report on student progress in a way that informs summer professional development and summer school course design, and informs instructional planning for the start of the school year by each grade's receiving teacher(s).

- **Providing clear information to all stakeholders:** Offering explanations about appropriate and inappropriate uses of assessment results.
- **Offering just-in-time teacher professional development supports:** Providing access to online learning communities, materials for school-based learning communities, or short Khan Academy®-like videos that unpack critical issues in College- and Career-Readiness standards (e.g., optimal use of mathematical representations, mathematical argumentation/modeling, issues that ELs have with specific mathematical notation or language, use of rubrics to support ELA instruction), along with resources built to support deepening teacher content knowledge for teaching using the learning progressions and key practices. We will also explore developing supports for paraprofessionals and other school-based personnel.
- **Offering assessment literacy resources:** Created for a wide range of stakeholders to support meaningful parent-teacher conversations and appropriate uses of assessment information at all levels, along with clear guidance about inappropriate uses of assessment information. Resources will include materials for professional learning communities to help teachers reflect on assessment information to plan for future instruction, reflect on their own practices, and use the data to evaluate curricular resources.

Informing the education conversation at multiple system levels

- **Exploring data representation and disaggregation** in ways that might help schools and districts more meaningfully use the data — for example, supporting comparisons of results to “schools/districts like you.” In addition, as the full Winsight Assessment System is built out, we can support schools’ and districts’ ability to link data from multiple sources by helping them better use the data in meaningful ways.
- **Creating resources to support state-, district-, and school-level discussions around student results** that aim to shift the focus from individual student results to a broader conversation about how districts and schools are performing across grade levels and subgroups and using the Winsight assessment data as a starting point for such discussions. Motivated by the NRC (1993)¹⁶ comment “Without equity in resources there can be no equity in assessment” (NRC, 1993), we are exploring how we can provide contextual additional data to support the interpretation of school and district assessment results, such as teacher turnover, teacher experience, student transience, and funding disparities.

¹⁶ National Research Council (1993). *Measuring what counts: A conceptual guide for mathematics assessment*. Washington, DC: The National Academies Press. doi: 10.17226/2235

C. Future Winsight Features or Components

The final category of effort, illustrated in the top layer of Figure 3, will focus on future, longer-term opportunities to develop system components and features that are targeted to traditionally underserved groups and their teachers. While we have a road map for future system improvements, this is an area in which we are looking for input to shape our work to meet the gaps in our equity agenda.

Winsight team members will be actively working with classroom teachers and school administrators in its planned user-based product development (UBPD). We will develop a co-creation network for the formative components and future revisions of the Winsight Assessment System. In this network, ETS plans to develop these Winsight elements using a design-based research methodology — collaborating with districts, states, and national organizations — as well as undertaking asset acquisitions with partner organizations to develop innovative and teacher-tested products that address the needs of future-looking districts and states. In this setting, the Winsight team will have the opportunity to continue to hear the needs of teachers and administrators firsthand and add to what we learned from the initial Winsight market research, with a specific focus on addressing the needs of traditionally underserved students and the educators working with them. In UBPD, we can listen to the educators who work with these students, ask the right questions about their needs, apply our research in resolving these questions, and collaborate with these educators to identify a methodology or product that supports their work. The outcomes of this work would serve as the first steps in developing additional products and features in the Winsight Assessment System that are specifically designed for traditionally underserved populations.

Future development efforts could target improved accessibility and accommodations features, reporting options/features such as opportunities to provide unique views of performance and help distill the information for traditionally underserved populations, the inclusion of additional diagnostic assessment tools such as diagnostic reading assessments, or an EL screener as part of the suite of integrated options that schools and districts can use. Additionally, we will look for opportunities and explore whether and how to integrate English Language Proficiency (ELP) assessments and ELA assessments that support ELs.

As described above, we are employing the best of the assessment design approaches and processes for each component of the Winsight Assessment System and considering ways in which equity considerations move us beyond the assessment itself to consider the broader supports around each component of the system, including reporting and teacher supports. Table 1 summarizes the key approaches that have been or will be employed for the Winsight Assessment System.

Table 1. Summary of Winsight System Components that Address Equity Concerns

THREE-TIERED DESIGN APPROACH	SPECIFIC ELEMENTS
ETS design and development work adheres to good assessment practices	<ul style="list-style-type: none"> • Evidence-Centered design • Universal design processes • 508 compliant • Usability and cognitive load issues • Accessibility and accommodations to support students with disabilities and English learners • Internal item reviews for accuracy, reading load, and fairness • External reviews at each stage of design process • Research basis for each assessment component
Winsight-specific features	DESIGN
	<ul style="list-style-type: none"> • Broad range of item types to assess standards • Use of learning progressions and key practices to identify and model important ideas and concepts in the standards through the summative assessment • Distinguish between purposes of interim assessment by creating benchmarks and testlets • Formative assessment component
Winsight-specific features	SUPPORT
	<ul style="list-style-type: none"> • More precise measurement of low-performing students • Meaningful information at the start of the school year • Accessible reports for parents • Just-in-time professional development supports for teachers • Supports around assessment literacy for a range of stakeholders, including guidance on appropriate and inappropriate uses of the assessment information • Exploring connecting assessment results with other data sources • Meaningful information provided earlier • Multiple approaches to communicating with parents • Assessment literacy resources
Future Winsight features or components	<ul style="list-style-type: none"> • Continued prototyping around formative assessment supports • Ongoing research to improve later iterations

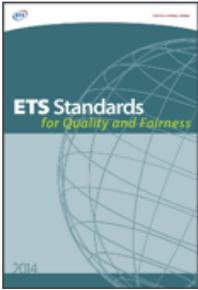
CONCLUSION

In this paper, we have described critical principles for a comprehensive assessment system that serves all students, including those from traditionally underserved groups. We focused on the three components (summative, interim, and formative) that form the Winsight Assessment System, through which stakeholders at all levels can access information to support their decision making — whether the allocation of professional learning resources is within a state or district or used to shape the next day's lesson to best support student learning. We presented a set of principles for equity that are central to this work. We also described the ways in which development, delivery, and reporting within each component are being considered through an equity lens so that we can meet our goals of:

- creating accessible assessments that promote rigorous instruction and meaningful data and feedback;
- supporting teacher learning through the provision of assessment models that exemplify the breadth, richness, and rigor of state standards;
- presenting reports that provide actionable information; and
- supporting professional learning opportunities for teachers to deepen their understanding of how all students learn.

This paper is not intended to be just a presentation of information, but rather a stimulus for discussion and learning that can lead to further innovation for the Winsight system in particular and for assessment efforts in K-12 assessment more broadly.

For Further Reading



ETS Standards for Quality and Fairness

The ETS Office of Professional Standards Compliance performs audits of all ETS tests to help ensure that they meet the exacting criteria of the *ETS Standards for Quality and Fairness*. These standards reflect our commitment to produce fair, valid, and reliable tests.



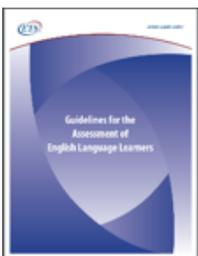
Guidelines for Constructed-Response and Other Performance Assessments

These guidelines are designed to assist staff in accumulating validity evidence for developing and scoring constructed-response and other performance assessments. They supplement the *ETS Standards for Quality and Fairness* by identifying standards that relate specifically to performance assessments and by offering guidance in interpreting and meeting those standards.



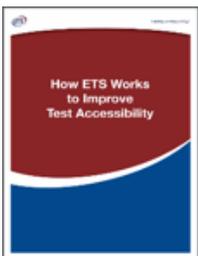
ETS Guidelines for Fair Tests and Communications

This publication identifies aspects of test questions that might hinder people in various groups from performing at optimal levels and helps enhance the overall effectiveness of communications. Fairness reviews are conducted by specially trained reviewers.



Guidelines for the Assessment of English Language Learners

These guidelines were designed for test developers, testing program administrators, psychometricians, and educational agencies. They were developed to ensure that assessments are fair and valid for English language learners. They focus on large-scale content area assessments administered in the United States to students in grades K–12. Many of these approaches can also be applied to other populations and assessments.



How ETS Works to Improve Test Accessibility

This publication describes to various stakeholders the work done at ETS to enhance the accessibility of our assessments and related products. It provides practical guidance about how, given their constructs, assessments can be made as accessible as possible to most test takers, including those with disabilities who do not need alternate test formats (ATFs). The document also describes how questions and assessments can be made more amenable to adaptation as ATFs.

Glossary of Terms

- *Balanced assessment system*: A combination of assessments that work together to provide information to the range of stakeholders within the system, at the right grain size and right frequency to meaningfully impact the decisions that those stakeholders must make.
- *Evidence-Centered Design*: Evidence-Centered Design¹⁷ is a framework for assessment development that begins with a clear identification of the constructs to be assessed (domain model), evidence for the constructs (evidence model), and approaches to eliciting that evidence (task model).
- *Formative assessment*: This type of assessment focuses on eliciting evidence of student learning during instruction to shape learning as it is happening, rather than to measure learning at the end of a unit or period of time. It has recently been defined as the following: A planned, ongoing process used by all students and teachers during learning and teaching to elicit and use evidence of student learning to improve student understanding of intended disciplinary learning outcomes, and support students to become more self-directed learners (CCSSO, 2018)¹⁸.
- *Interim assessment*: This type of assessment provides data at specific points during a school year or at the end of an instructional unit that can be aggregated across teachers, schools, or the district and used to understand what students have learned.
 - *Interim benchmark assessments* in the Winsight Assessment System reflect the summative assessment blueprint and are intended to be used a few times per year to provide information about future performance on the summative assessment.
 - *Testlet assessments* in the Winsight Assessment System are tied to specific learning progressions or key practices and are intended to provide common assessments that could be used at the start or end of an instructional unit.
- *Learning progressions and key practices*: These are descriptions of theories of students' competency development that articulate how learning develops from a novice understanding or practice of a competency to a more sophisticated understanding or practice. Winsight mathematics testlets rely on one or two learning progressions for each testlet to support development and reporting. The ELA testlets use key practices to support development and reporting.
 - *Learning progressions* describe conceptual thinking and understanding at increasing levels of sophistication, often identifying major shifts in understanding. Learning progressions outline common ways in which student understanding matures, although they may not apply to all students and should not be taken as prescriptive.

¹⁷ Mislevy, R. J. & Haertel, G. D. (2006). Implications of Evidence Centered Design for Educational Testing. *Educational Measurement: Issues and Practice*, 25(4), pp 6–20.

¹⁸ Council of Chief State School Officers (2018). *Revising the definition of formative assessment*. Washington, DC: Council of Chief State School Officers.

- *Key practices* for ELA were articulated as a way of organizing the information from multiple learning progressions into a set of closely related tasks and interactions that serve a common purpose; exercise a common set of skills (in reading, writing, and/or critical thinking); and must be mastered to achieve college and career readiness, such as the key practice of building and sharing knowledge.
- *Multistage adaptive testing*: This approach is used for computer-adaptive tests in which sets of items are selected for each test taker by an algorithm based on student performance on an initial set of items.
- *Summative accountability assessment*: This assessment is primarily intended to measure students' proficiency with respect to each state's standards to inform educational stakeholders and the public of the status of the school or system.
- *Universal design*: Universal design for learning¹⁹ principles and more specific guidelines for large-scale assessments²⁰ provide guidance that includes attending to multiple means of representation, multiple means of action and expression, and multiple means of engagement.

¹⁹Rose, D. H., & Meyer, A. (2002). *Teaching Every Student in the Digital Age: Universal Design for Learning*. Alexandria, Virginia: Association for Supervision and Curriculum Development.

²⁰Thompson, S. J., Johnstone, C. J., & Thurlow, M. L. (2002). *Universal design applied to large scale assessments* (NCEO Synthesis Report No. 44). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.



Created by **ETS**. Delivered by **Questar**.

ETS — *Measuring the Power of Learning.*[®]

Copyright © 2018 by Educational Testing Service. All rights reserved. ETS, the ETS logo, MEASURING THE POWER OF LEARNING and WINSIGHT are registered trademarks of Educational Testing Service (ETS). The Winsight logo is a trademark of ETS. All other trademarks are property of their respective owners. 40893