DISTANCE EDUCATION FOR TEACHER TRAINING:

Modes, Models, and Methods

Mary Burns
Education Development Center, Inc.
Washington, DC

EDC Education Development Center

Acknowledgments

Thank you to the following colleagues both near and far for providing information, resources, and fact-checking on this guide.

Sarwat Alam, Director of Learning Systems and Pedagogy, USAID Pre-STEP Project (2013), Pakistan

Dr. Catherine Margaret Beukes-Amiss, Director, Centre for Innovation in Learning and Teaching (CILT), University of Namibia, Windhoek, Namibia

Arjana Blazic, Teacher Trainer, and Course Designer, EduDigiCon, Zagreb, Croatia

Alisa Buchstab, Junior Policy Advisor in the Sector Program Education,Deutsche Gesellschaft für Internationale Zusammenarbeit GmbH (GIZ), Germany

Dr. Eduardo C. Cascallar, University of Leuven, Belgium; Managing Director, Assessment Group International (Europe and USA)

Will Clurman, CEO, and co-founder, eKitabu, Nairobi, Kenya

Valeria Cruz Gomes, Head of Training and Support, ProFuturo, Madrid, Spain

Dr. Robyn A. Defelice, Learning Strategist and Consultant, Bloomsburg, Pennsylvania, USA

Dr. Nathalia Edisherashvili, Researcher, Institute of Education, University of Tartu, Estonia

Concepción Gallego Garcia, Expert on Global Partnerships and Institutional Relations, ProFuturo, Madrid, Spain

Dr. Sophia Gorgodze, Director National Assessment and Examinations Center, Ministry of Education and Science of Georgia

Dr. Sara Hennessy, Professor of Teacher Development and Pedagogical Innovation, Faculty of Education, University of Cambridge, and Research Director, EdTech Hub, Cambridge, England

Shane Ives, Serious gamer, solar electrician, Albuquerque, New Mexico, USA

Eilean von Lautz-Cauzanet, Policy Advisor in the Sector Program, Deutsche Gesellschaft für Internationale Zusammenarbeit GmbH (GIZ), Germany

Stephen McDonald, app developer, Somerville, Massachusetts, USA

Dr. Mary Mendenhall, Teachers College, Columbia University, New York City, New York, USA

Dr. Barbara Moser-Mercer, Visiting Professor, University of Nairobi, Coordinator African Higher Education in Emergencies Network (AHEEN), Nairobi, Kenya

Denise-Henry Orndorff, Instructional Technology Coach, Frederick County Public Schools, Virginia, USA

Ee-Reh Owo, Schools Director, Justice Rising, Democratic Republic of the Congo

Aroa Sánchez Rodríguez, Educational Specialist, Innovation and Products, ProFuturo, Madrid, Spain

Jodi Sansone, Instructional designer, and eLearning specialist, Jodisansone.com, USA

Dr. Beverly Shirley, Academic Program Officer for University of the West Indies Open Campus British Overseas Territories, (Cayman Islands, Anguilla, British Virgin Islands, Monserrat, Turks & Caicos, Bermuda), George Town, Grand Cayman, West Indies

Dr. Carmen Strigel, Director, Education Technology, RTI International, North Carolina, USA

Dr. Torrey Trust, Associate Professor, Learning Technology, College of Education University of Massachusetts Amherst, USA

Freda Wolfenden, Professor of Education and International Development, School of Education, Childhood, Youth and Sports Studies, The Faculty of Wellbeing, Education and Language Studies, United Kingdom Open University, Milton Keynes, England

Dr. Diana Woolis, CEO, Sustainable Learning Strategies, New York City, New York, USA

Nicole M. Zumpano, Director of Instructional Technology Coaching, The Learning Technology Center (LTC) of Illinois, USA

My thanks to Education Development Center (EDC) colleagues for sharing program information and for providing resources, review, and feedback on chapters or content.

Helen Boyle, Vice President, Director of Program Strategy

Susan Bruckner, Senior International Technical Advisor

Nancy Meaker Chervin, International Technical Advisor

Rachel Christina, Director, International Basic Education

Leslie Goodyear, Distinguished Scholar/ Principal Evaluation Director

Nevin Katz, Web and App Developer

Stephanie Knutson, International Accreditors for Continuing Education and Training Compliance Manager

Nora Nunn, International Technical Associate

Shelley Pasnik, Senior Vice President

Gerald Sanders, Facilities Administrative Manager

Tamara Vitolo, Research Associate, Center for Children and Technology

Katherine Yasin, Principal International Technical Advisor, Director of English for Latin America.

Special thanks to Bronwyn Taggart for her careful editing.

My deepest gratitude to Mary Hooker, International Technical Advisor, Education Development Center, for review, feedback, and wonderful insights on multiple chapters in Section II of this guide.

My final thanks to those who fund, develop, research, evaluate, design, teach and participate in distance learning programs for teachers across the globe. This guide draws on your work. Thank you as well to EDC's Digital Design Group.

About the Author

Mary Burns is a senior technology and teacher professional development specialist at EDC. A former 10-year teacher in the United States, México, and Jamaica, she has worked in the area of technology-enabled professional development since 1997, instructing, designing, and evaluating both distance-based and face-to-face professional development for teachers, teacher educators, and instructional coaches. She has authored peer-reviewed papers, books, articles, and blog posts about teacher professional development, distance learning, and teaching with technology. She works in Asia, Africa, the Middle East, Latin America, the Caribbean, Europe, and the United States.

Preferred Citation

Burns, M. (2023). *Distance Education for Teacher Training: Modes, Models and Methods.* (2nd Edition). Washington, DC: Education Development Center.

Section II. Chapter 17

ASSESSING DISTANCE LEARNERS

Table of Contents

Best Practice: Successful distance learning programs use a range of formative and summative assessments to improve instruction and to accurately measure teachers' knowledge, skills, competencies, and attitudes.

17.1 Overview

At its core, assessment is the relationship between the "instruction learners receive and what they actually learn as a result" (Black & William, 2018, p. 20). As such, assessment is a critical component of any distance education program. However, in some cases, assessment may be the weakest component of a distance program.

Figure 17.1
Assessment Versus Evaluation

"Assessment" and "evaluation" are often used synonymously, but they are different. Assessment *in this guide* refers to individuals, whereas evaluation refers to programs (although that rule does not apply in real life—individuals can be evaluated, and programs can be assessed).

Assessment refers to any of a variety of procedures used to obtain information. It includes distinct types of measures of knowledge, skills, and performance, usually in the service of learning. Assessment may have an evaluative component —a summative assessment, such as a final exam— that places a value or judgment on performance.

Evaluation, the focus of the next chapter, is a set of procedures for determining the value or overall worth of a program. It essentially examines impact or outcomes based on predefined criteria.

Because of the "open" and "distant" nature of Open and Distance Learning, these programs face a particular dilemma in assessing pre- and in-service teacher matriculation: how to measure teacher-candidates' process of learning, the products of their learning, and their progress in learning; how to assess their "fitness" as teachers and how to do so in ways that contain direct, observable evidence of "teaching in action;" and how to validate or authenticate the written work of learners they may never see (Letseka & Pitsoe, 2013).

Many distance-based continuing education programs may not be required to summatively assess whether and what teachers have learned as a result of the program. Or they may use standardized tests that measure out-of-date skills—focusing on declarative knowledge versus procedural knowledge or conceptual and epistemological knowledge (Niess, 2011).[1]

Additionally, constrained financial resources, a lack of access to adequate technologies, logistics, and a lack of qualified test-design specialists and trained assessors may make it difficult to support more valid and realistic performance-based assessments, such as in-class observations of teacher performance, personalized assessments, or digital portfolios of teacher work. Finally, many entities may not wish to assess teacher learning; their aim simply may be to get teachers and

[1]This categorization is often described as *knowing that* (declarative knowledge); *knowing how* (procedural knowledge—application of skills); and *knowing why and wherefore* (conceptual knowledge—deep understanding, and methods of knowledge acquisition) (Niess, 2011).

teacher-candidates in and out of the distance education system as effortlessly as possible.

This chapter focuses on *assessing* teacher learning and performance in a distance education program. It discusses how successful distance education programs have overcome many of the above issues by using a range of assessment methods as appropriate. It advocates that assessment be designed to gain evidence about learners' capabilities, and this evidence be used to adjust instruction to better meet learners' needs (Black & Wiliam, 2018; Heritage, 2010).

While this chapter focuses on assessment, Chapter 18 will discuss *evaluating* distance education programs. Because assessment and evaluation are so closely linked, Figure 17.1 explains the distinction.

17.2 Assessment in Teacher Education Programs

Assessment is essentially a vehicle to gather some idea of what a pre-service teaching candidate or in-service teacher knows, can do, values, and believes (Letseka & Pitsoe, 2013). It is typically categorized as *formative*—a low-stakes assessment, which is not scored, and the information from which is used to improve teaching and learning processes—or *summative*, which is a high-stakes assessment, the aim of which is to produce a product that is then evaluated or judged using some form of a grade or mark (Black & Wiliam, 2018).

Assessments within distance education programs generally adhere to this formative-summative dyad and in so doing, generally serve the following purposes:

- Assessment to support learning. Formative assessment is typically used to support learning by informing the instructor how well, or poorly, learners understand content so that the instructor can reteach information or change the course of instruction.
- Assessment for certification. Assessment can certify that a teacher candidate has attained a defined set of benchmarks that govern professional entry into teaching. For example, aggregated learner results from exams and projects or national and international benchmarks and tests of teachers' knowledge and skills (such as the Praxis in the U.S., the Licensure Exam for Teachers in the Philippines, or the National Teacher Qualification Test in South Korea) provide evidence of teacher candidates' attainment of a set of learning outcomes or a professional body of knowledge. This certification also has an accountability dimension.
- Assessment for sorting and selection. Assessment in this category can be used for choosing, sorting, or screening teacher candidates into or across particular positions, programs, career tracks, or awards based on assessment results. For example, they can be used to determine promotion to another level of teaching as part of a career ladder; reward or acknowledge performance; or transfer a teacher to a more prestigious school (Archer, 2017).

Thus, the inferences made about the purposes of assessment often drive the assessment deployed, and the purpose of each is specific to an intended outcome. There are numerous assessment methods that can be utilized for these three distinct purposes of assessment, as Figure 17.2 outlines.

As seen from Figure 17.2, many of the above higher-order thinking assessment activities will require well-designed rubrics. A rubric is a scoring guide that assesses open-ended projects, performances, and tasks that focus on higher-level thinking skills or social-emotional skills. It lists criteria or "what counts" for a piece of work as well as gradations of quality. There are essentially two types of rubrics: *holistic* and *analytic*. A *holistic* rubric requires the teacher to score the overall process or product as a whole, without judging each part separately.

Figure 17.2
Common Types of Assessments and Their Advantages and Disadvantages (Adapted from Downing, 2006; also see Commonwealth of Learning and Asian Development Bank, 2008, pp. 4–14)

Assessment	What It Assesses	Advantages	Considerations
Tests: Select response (e.g., multiple choice, true/false, matching) Learners select the correct response among a series of options.	• Facts • Understanding of ideas • Application of principles	• They are a direct measure of the learner's knowledge of a domain. • With the exception of multiple-choice tests, they are easy to construct and easy to grade. • They can assess many more topics broadly. • They facilitate faster return of exam results to learners. • They produce accurate, objective, and reproducible scores.	• It is difficult to create good multiple-choice tests (especially with good "distractors"). • They measure a very narrow range of knowledge (identification and recall of information) rather than skills or procedural knowledge (how to do something) or higher-level thinking. • Poorly developed tests make guessing easy.
Tests: Constructed response (e.g., fill-in-the-blank, cloze, short answer, sentence completion) Learners "construct" or supply their own response. (Also includes essays—see below.)	• Facts • Understanding of ideas • Application of principles	• They are a direct measure of the learner's knowledge of a domain. • They can assess many more topics broadly. • They allow for ease of partial credit scoring. • With constructed-response tests, there is less of a predilection to guess, as on a select-response test.	• Some (e.g., Fill-in-the Blank) measure a limited depth of knowledge (identification and recall of information). • Others (e.g., Short Answer) can assess logic, reasoning, and problem solving.
Essays	• Understanding of ideas • Ability to organize information • Ability to develop an argument, support it with ideas and evidence, and formulate a conclusion based on arguments and evidence	• They are a direct measure of learner's higher-order thinking (logic, reasoning, problem solving, thinking skills, procedural and conceptual skills) as well as written communication skills.	• They require scoring rubrics, otherwise grades are unreliable. • Much depends on the quality of the essay prompt itself: its length, specificity versus generality, clarity, and its focus on specific learning outcomes.

Assessment	What It Assesses	Advantages	Considerations
Essays *(continued)*	• Ability to communicate in a variety of written forms • Written fluency in the language of instruction	• They allow learners to express their knowledge in a less constrained, more open format than a test (such as multiple-choice, short answer, fill-in-the-blank).	• Reliability and validity depend on well-trained raters, scoring scales, and well-developed prompts. • Essays take longer to grade. • Without well-designed rubrics, the subjectivity of grader is always a major concern. • Essays require writing to be taught in curriculum to master rhetorical, mechanical, and grammatical conventions. • Large language models, such as GPT or Bloom, can generate essays for learners.
Oral Assessments	• Oral fluency in general and in the language of instruction • Reasoning, problem solving, interpersonal skills • Speaking, poise, thought processes	• They are a direct measure of learner's higher-order thinking (logic, reasoning, problem solving, thinking skills, procedural and conceptual skills) as well as oral communication skills. • The give-and-take nature of conversation/oral communication may be more natural to the learner. • They can be used to confirm other assessments.	• Oral assessments require a well-developed rubric. • They may be time-consuming to mark, particularly if recorded. • It may be difficult to standardize questions. • There is the possible introduction of bias due to the personal nature of the assessment. • Cultural considerations can impede their effectiveness: learner shyness, hierarchy, lack of comfort of female learner being orally assessed by a male instructor. • Performance anxiety may be greater here since oral assessments are typically conducted in-person.

Assessment	What It Assesses	Advantages	Considerations
Classroom observations	• Practice-based skills (e.g., teaching practicum, micro-teaching, or simulated teaching) • Cumulative body of knowledge in action • The ability to transfer information and principles to novel and authentic situations • The capacity to deal with "real life" classroom situations *in situ*	• An observation is a direct measure of teaching ability and is thus performance-based. • Observations are a more authentic and direct form of assessment. • Observations when done well are empirical and objective.	• Observations require an observation form or tool and well-trained observers. • The tool should be high-inference to capture nuances in instructional quality and to measure teacher progress over time. • They require established protocols demarcating the length of observation, the position of the observer, and pre- and post-observation meetings. • Assessing teachers' practice requires ongoing observations to capture breadth of teachers' skills. • Observations will be examined in greater depth in Section 17.4.
Projects, Theses, Capstone Projects	• Reveal depth of procedural and conceptual knowledge and mastery of a particular topic • Creativity and organization of information • Theses: Writing, knowledge, and thinking skills	• Projects/theses or capstones mainly assesses cumulative procedural and conceptual knowledge of the teacher candidate or teacher. • They can assess harder-to-measure constructs such as affect, creativity, behaviors, aptitudes. • They can assess learner's depth of knowledge, understanding of theories, and research • They combine instruction and assessment—the teacher candidate learns as she creates her own assessment product. • They assess the totality of a learner's work.	• They call for more open-ended assessments (journals, portfolio, video, examples of learner work). • They require an analytic rubric. • They are time-consuming to grade. • With projects, there is potential subjectivity and a lack of validity and reliability without very well-developed guidelines and a reliable rubric. • Capstones, in particular, are often open-ended and requirements may differ according to programs or instructors (e.g., product might be a short video versus a paper versus a performance).

Assessment	What It Assesses	Advantages	Considerations
Portfolios	• Same as points above • Multiple levels of assessment (knowledge of facts, analysis and evaluation of information, and self-reflection)	• Portfolios encourage learners to display knowledge and understanding in multiple formats, especially digital or Web-based ones. • They assesses the learner's participation, processes of learning, progress, as well as final product of their learning.	• Without video or audio, capturing teacher behavior or practice, portfolios are not a direct measure of teacher performance. • They require an analytic rubric. • They are time-consuming to grade. • They suffer from potential subjectivity and lack of validity and reliability without a very well-developed analytic rubric.

An *analytic* rubric scores separate, individual parts of the product or performance first, then adds the individual scores to obtain a total score. They are matrix-like, with performance levels that show delineations on quality and with specific and measurable descriptors (e.g., 1–4, "emerging" to "proficient"). Analytic rubrics are usually preferred when a fairly focused type of response is required (i.e., for performance tasks in which there may be one or two acceptable responses and creativity is not an essential feature of the learners' responses.)

The above table also outlines both the potential and the conundrum of assessment. Almost every method discussed in Figure 17.2 can be utilized for both formative and summative purposes. Thus, assessment is less about the *methods* that distance programs use, and more about the *inferences* they draw about learners from these assessment outcomes (Black & Wiliam, 2018). Where inferences relate to the status of the learner or concern their future potential, then the assessment is functioning *summatively*. Where the inferences relate to the kinds of actions that would best help the student learn, then the assessment is functioning *formatively* (Black & Wiliam, 2018, p. 3).

17.3 Technology-Based Assessment

Technology holds numerous benefits for assessment in distance education programs. This section discusses some of the key benefits of computer-based assessments, computer adaptive testing, and technologies to support formative assessment.

17.3.1 Computer-Based Assessments

Technology has completely transformed assessment from the era of Margaret, the distance learner encountered in the Foreword of this guide, who waited months for her assessment results. Current computer-based assessments (CBAs) provide versatility, flexibility, and automation in terms of what can be measured within any distance education system. CBAs can:

- **Allow for multiple-test administrations.** Learners can take multiple, short, reliable assessments administered during the academic year. The data gathered from these assessments can be correlated with national standards so that teacher-learners can be measured on these standards (Reville et al., 2005).
- **Use learning analytics for personalized support.** Information derived from individual

learner characteristics, learner choices, and assessment data can be collected, measured, and analyzed. Learning analytics can be used to better understand learner needs; optimize course offerings, design, and instruction; provide distance education programs with the information needed to support learner progression; and enable personalized, rich learning (Tempelaar et al., 2015).

- **Improve the testing experience.** Game-based assessment apps such as *Kahoot!* and *Quizlet* and the quizzing and rewards features of tools such as *Duolingo* make assessment more fun and engaging, less "test-like" and thus less stressful for learners (Wyatt-Smith et al., 2019).
- **Provide a fuller picture of learner achievement and capabilities.** Through their ability to create dynamic and individualized assessments, CBAs can produce a more rounded and complete picture of a learner's achievements and capabilities to help instructors identify interventions, supports, and personalized pathways for learning (Wyatt-Smith et al., 2019).
- **Provide immediate and varied feedback to learners.** Computers can score tests in real time, allowing distance instructors to make real-time instructional changes based on assessment evidence and providing learners with real-time information about their progress and performance (Black & Wiliam, 2018; Wyatt-Smith et al., 2019). There are essentially two types of feedback: verification and elaboration. Verification indicates whether an answer is correct, and elaboration provides information to guide the learner toward the correct answer (Kulhavy & Stock, 1989). This feedback can be part of an overall "knowledge building cycle" (Timperley et al., 2007) where, depending on the computer application, instructors can program the system to provide learners with more elaborate feedback and with just-in-time help to resolve gaps in learner understanding (Myung et al., 2020). Tools powered by artificial intelligence (AI), such as *ChatGPT, Gradescope, Grammarly,* or *Cognii* can provide learners with verification feedback while adapting assignments according to assessment results. As natural language models, like *ChatGPT*, continue to evolve, they may also be able to offer elaboration feedback to learners.
- **Vertically align tests.** Tests can be anchored to assess the same core knowledge at increasing levels of difficulty (criterion-based testing) (Reville et al., 2005).
- **Horizontally align tests.** Tests can be scored in such a way that learners can be compared against one another (norm-referenced), which may be critical for sorting and choosing pre-service teacher candidates for teaching posts, scholarships, or further education. Raw test scores could be given phase-wise or as a total. Learners could receive a letter grade or percentile score to determine their relative position *vis-à-vis* other learners (Reville et al., 2005).
- **Include ipsative or growth measures.** Tests measure individual growth over time, so programs are able to benchmark where learners should be at the end of a course of study based on tests from the beginning of a course of study (Reville et al., 2005).
- **Help learners with disabilities.** Technology tools such as screen readers, magnification tools, and text-to-voice or voice-to-text applications can help learners with visual, auditory, and motor impairments; learners with dyslexia; and learners who simply need more time to complete a test. Using AI-powered voice assistants, visually impaired learners can use voice commands to have text read aloud to them. AI and augmented reality applications can help deaf and hearing-impaired learners read by translating texts into sign languages (Burns, 2021).
- **Streamline and automate marking.** Tools to support grading and marking, such as *RubiStar*, *GradeAssist*, *OrangeSlice*, *Hot Potatoes*, and *eMarking Assistant,* have long been popular supports for instructors wishing to eliminate the tedium of marking learners' work, especially open-ended assignments.

Increasingly AI is being harnessed to power formative and summative assessment tools based on instructor specifications. For example, *IntelliMetric* and *e-rater* use past evaluations of long-form essays to create a rubric and framework for evaluating new assignments, which it then uses to assess learners' work and offer feedback on learner work. *E-rater* weighs key features of the learner's writing skills and provides feedback. Peer-to-peer assessment can be made easier via tools such as *CrowdGrader*, *peerScholar*, *Cocertify*, and *PeerWise* (Contact North | Contact Nord, 2020).

- Automate large-scale assessments. Computer-based assessments can reduce manual labor so that certain tasks are performed more efficiently, at a higher volume, and at scale without creating undue burdens on instructors, undue expense for an education system, and unduly delayed results for learners (Burns, 2021).
- Generate big data for policy and planning. Beyond distance education, large-scale computer-based assessments, particularly cross-national ones, generate vast quantities of longitudinal data that have been analyzed and used by governments for international benchmarking of learners, analyses of countries' educational conditions, the formulation of education policies related to teaching and learning, and instructional improvement. Some of the more familiar cross-national exams are Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ), Program for the Analysis of Education Systems (PASEC), the Program for International Student Assessment (PISA) and PISA for Development (PISA-D), Trends in International Mathematics and Science Study (TIMMS), and Progress in International Reading Literacy Study (PIRLS).

17.3.2 Computer Adaptive Testing

One powerful category of computer based assessment is Computer Adaptive Tests (CATs). Computer Adaptive Tests can be used with a number of distance-based modes, such as multimedia learning, online learning, and mobile learning (see Chapters 4–6 of this guide). CATs are technology-based or online testing systems created by content specialists, psychometricians, programmers, and Web designers. They contain large banks of test items and use item response theory (IRT) for pre-calibration, determination of the item parameters (discrimination and difficulty), establishing learners' performance levels and scoring, as well as other functions.

CATs "adapt" questions to be easier or more difficult based on the learner response to previous questions. This calibration or adaptation allows for a precise and quicker measurement of a learner's knowledge using a smaller number of test items. CATs can be diagnostic (identifying learners in need of more academic support), formative (informing instructors of the learner's grasp of material), or summative (as a final determination of learning). They are typically administered online, so scoring is computerized, and results disseminated quickly. They allow for accurate measurement and a shorter and more targeted testing experience for learners (E. Cascallar, personal communication, August 29, 2022).

17.3.3 Technology-Supported Alternative Assessments

CBA is not simply about CATs or computer-based testing. As Chapters 4, 5, and 6 discuss, interactive technologies allow for multiple and flexible types of alternative (i.e., non-test) assessments—discussions, projects, and performance-based assessments (e.g., mixed reality, digital learning games, and simulations)—that provide a wealth of authentic, engaging, and holistic assessment opportunities for distance learners, both synchronous and asynchronous, Web-based and non-Web-based, and formative and summative.

Using *traditional assessments (computer-based tests)* and *immersive ones (augmented reality, virtual reality, mixed reality, extended reality, and simulations)* can create more responsive instruction for learners; assess teachers' instructional skills in authentic, or nearly

authentic, environments; measure higher-level skills; and ensure teachers' quality, readiness, and fitness to teach (Black & Wiliam, 2018; Gee & Shaffer, 2010; Timmis et al., 2015; Wyatt-Smith et al., 2019). *Digital learning games, virtual worlds, simulations,* and *immersive environments* can provide a developmental sequence of challenges that gradually increase in difficulty so that learners are working at their highest abilities. These tools can also be used to assess the learner's ability to collaborate, problem-solve, and employ systems thinking (Gee & Shaffer, 2010; Buckley et al., 2021).

In addition to these multimedia applications, *eye-tracking and facial recognition software* are increasingly used to assess difficult-to-measure skills, such as higher-level thinking skills, the social-emotional skills of persistence, creativity, and self-regulation, and learners' affective states, such as engagement or frustration (Buckley et al., 2021). These assessments could be enhanced via the use of "think-aloud" protocols so learners could explain their decisions (e.g., in a digital learning game), their rationale for such decisions, and their attitudes and affective states while wrestling with difficult concepts or tasks (Gee & Shaffer, 2010, p. 14; see also Wyatt-Smith et al., 2019).

Electronic portfolios, digital representations, projects, and *digital collections of their work* allow learners to showcase their processes of learning and progress in a distance-based course. *Online discussions*—whether they are synchronous and video-, audio-, or text-based, or asynchronous and text-based, as on an LMS discussion board—provide evidence to assess emerging understandings of concepts and theories, as well as skills such as reasoning, evaluation, and argumentation (Myung et al., 2020).

Extensive writing via *word processing or a digital writing tool*—versus handwriting—where learners put forth a thesis statement, support their idea with evidence, and arrive at a conclusion—has been shown to improve writing scores *if* learners go through the complete writing cycle of drafting, editing, revising, and rewriting (Warschauer, 2009).

Developing *blogs* and multimedia presentations, particularly with curated hyperlinked resources, can demonstrate learners' understanding of an issue, their appreciation of its complexity, and their knowledge of appropriate resources that address the issue.

Audio- and Web-conferencing tools allow learners to present information to one another and the instructor and to engage in debates about a particular teaching-related or content-based issue.

Using *mobile phones*, teacher-learners can be assessed on national language abilities (Hindi, Arabic, Urdu, Swahili) or participate in oral assessments, and their scores can be immediately tabulated and returned in real time. Similarly, learners can use the texting features of mobile phones and quickly send answers to a multiple-choice or closed-response quiz or test, which can be analyzed and tallied, with the score returned via text messaging (Morris et al., 2021).

Finally, *back-end data from LMSs*—the number of logins, time on task, and number of discussion posts— can be linked to hard assessment data, such as examinations or performance-based data, to provide a fuller assessment of a learner's effort and progress in an online course.

17.3.4 Improving Technology-Based Assessments

While exciting and promising, assessment via technology has numerous issues, not least of which are privacy and data integrity issues. For the purposes of this chapter, however, we focus on two issues that can enhance or undermine technology-based assessments as part of distance learning courses: good test design and academic dishonesty (cheating).

Improving testing

While tests often are critiqued as imperfect measures of teachers' skills, they are a staple in the assessment repertoire, and select- and constructed-response tests offer numerous

advantages, as outlined in Figure 17.2. Their greatest advantages may be twofold: First, if tests are designed well, examinees who answer correctly should be higher performers than those examinees who do not—and their answers should be based on expertise versus random guessing.

Second, quizzes and tests support "retrieval practice," which consolidates new learning (Roediger III & Butler, 2011). Retrieval practice is more commonly known by its original moniker, the "testing effect" (See Figure 11.5 in Chapter 11). When learners know they will be assessed on material, they learn it better and retain it longer than if they just study the material *without* a test (Batsell Jr. et al., 2017; Brame & Biel, 2015). The "(mere) presence" of a quiz enhances learning in part because it contributes to reduced "mind wandering," improved self-regulation, increased task-relevant behaviors such as note taking, and enhanced calibration predicted and actual performance (Haagsman et al., 2020, p. 722; Schacter & Szpunar, 2015, p. 64).

But tests only confer benefits when designed well. Poor test design is particularly problematic in distance programs that lack skilled psychometricians or assessment specialists. The biggest perpetrators of poor test design are the most frequently administered type of tests—multiple-choice tests—because they often fail to challenge learners to reason or analyze rather than simply memorizing information. One U.S.-based national study of test-bank questions from 77 university-level introductory biology courses estimated that 93% of the questions tested levels 1 and 2 on Bloom's Cognitive Domains of Learning (Knowledge and Comprehension) (Momsen et al., 2017). Bloom's Taxonomy is outlined in Figure 17.7.

There are ways to improve multiple-choice tests, as Figure 17.3 outlines.

Figure 17.3
Improving the Design of Multiple-Choice Tests (Brame, 2013; Burns, 2018)

Parts of the Question	Design Guidelines
The stem	• Should be clear, relevant, and brief. • Should be a question or partial statement. • Should be directly linked to the curriculum and the most important topics taught.
Alternatives (responses, answers)	• Create clear, concise, direct alternatives. • Should be mutually exclusive, homogeneous, and presented in logical order. • Avoid double negatives. • Avoid the use of "All of the above" or "None of the above." Those alternatives reward learners who don't know the answers. If such options must be used, do so with caution and ensure that it is the correct response approximately 1 in 4 times. • All answers—the correct answer and the distractors (the incorrect answers)—should be consistent in length, style, and construction. Learners should not be able to guess the right answer because it looks different from the wrong answers. • Increase the plausibility of distractors by choosing distractors based on common learner errors.

Other guidelines	• Ensure that tests are valid and measuring what they are supposed to and what's been taught. • Weight test content according to the amount of time spent on a particular topic. • The learner should not be able to guess the correct answer from the way the response is written. • Pay attention to language. Avoid grammar, spelling, and mechanics errors which may make it difficult for learners to even understand the question and possible alternatives • Avoid categorical terms, such "always," or "never." There's no such thing as "always" or "never," and this is a giveaway. • As long as all alternatives are plausible, the number of alternatives can vary among items. (There is a minor difference in difficulty, discrimination, and test score reliability among items containing two, three, and four distractors.) • Avoid complex multiple-choice problems (i.e., alternatives such as 1 and 2; 2 and 3; 1 and 3; 1, 2, and 3) • Each question should stand alone and be unrelated to or disconnected from other questions. The point is to avoid "double jeopardy," where if a learner answers one question incorrectly, another answer also will be incorrect. • Pilot the test when finished creating it—either with a non-test taking learner or using AI-driven chatbots which, if prompted correctly, can provide feedback on test construction, clarity of directions and distractors.

In addition to multiple-choice tests, specifically, research suggests that testing in general also can be improved by employing the following strategies:

- **Vary test structures.** Avoid assessing learners using the same type of test repeatedly. For example, instructors can administer an open-ended question for one test and short answer-questions for another or use fill-in-the-blanks questions instead of multiple-choice, since fill-in-the-blank questions require learners to retrieve information rather than just recognize the correct answer (Hultberg et al., 2018). A variety of types of tests allows instructors to gather different types of data about a variety of learners' knowledge and skills (Tropman, 2014).
- **Use pre-class quizzes as part of a flipped approach.** Flipped classrooms, as discussed in *Chapter 5: Online Learning*, typically assign learners to read text or watch a video before their in-person or synchronous class. By administering a quiz after this pre-class reading and before the actual class, assessment can help instructors diagnose learners' levels of understanding; determine who has done required readings; guide instructors on what they should focus on in a face-to-face or online session; and provide ongoing opportunities for retrieval practice (Brame & Biel, 2015; Tropman, 2014).
- **Employ collaborative tests and quizzes.** Test anxiety is a real phenomenon for many learners, with measurable physiological and psychological impacts that hinder performance (Heissel et al., 2021; Pandey & Kapitanoff, 2011). Collaborative quizzes can alleviate this anxiety for learners with higher levels of test anxiety. Learners can take an individual poll or quiz, then discuss with peers and retake the quiz. They each can take the quiz individually and decide among the group whose quiz to submit, or they can take the quiz together at the same time (Pandey & Kapitanoff, 2011).
- **Use low stakes and ungraded quizzes.** Not every assessment requires a grade, and

grading everything can have the unintended effect of causing learners to focus on a number or letter result versus actual learning. Distance instructors can diversify quizzes by including non-graded ones. With ungraded quizzes or assignments, learners can wrestle with a concept or practice a skill without fear of being judged or evaluated. They can receive credit with a checkmark, thereby acknowledging completion or effort instead.

- **Use assessment experts to design tests, especially summative ones.** Correct test design is complex and involves many requirements about the characteristics and number of items in the item bank, correct calibration of the items, and application to the correct population of learners (E.C. Cascallar, personal communication, August 29, 2022). While *Chapter 13: Preparing Distance Instructors* advocates for distance instructors to learn assessment practices, high-stakes test design involves skills best left to an assessment expert.
- **Be conscious of stereotype threat.** Testing can induce documented disproportionate psychological stressors on some learners—often ethnic, religious, and racial minorities; low-income learners; and females—thus heightening the risk of "stereotype threat." Stereotype threat is the fear of being "viewed through the lens of a negative stereotype, or the fear of doing something that would inadvertently confirm that stereotype" and has been shown to increase the test-taker's anxiety, reduce self-efficacy, and hinder actual performance (Steele, 1999). Stereotype threat is a critical, although under-explored, issue in teacher education. But it is highly germane since many current and future teachers are female and non-White, and since it occurs most commonly in STEM subjects—science, technology, engineering, and mathematics—content areas in which teacher shortages are most acute (Heissel et al., 2021; See, 2022; Spencer et al., 2016; Steele, 1999).

Technology can help to mitigate some of the effects of stereotype threat. For example, in some cases, traditionally underserved learners, who are more likely to be adversely affected by stereotype threat, may do better in online classes of traditionally "difficult" subjects, such as Algebra I, versus in face-to-face classrooms (Heissel, 2016). Computer-adaptive assessments that tailor content to learners' skill level in a tested domain can improve the reliability of the results and reduce learner frustration, mitigate stereotype threat, and potentially increase motivation (Burns, in press).

Most important, however, are non-technology solutions to stereotype threat. One is having distance instructors who reflect the diversity of their learners. A second involves instructors cultivating a personal, though professional, relationship with their learners. Third is the necessity of employing gender-responsive pedagogies. For example, university-level female learners experience significant improvement in self-efficacy relative to their male counterparts when instructors include activities that promote "social persuasion" (encouragement to adopt an idea, attitude, or stance), such as teamwork, group discussion, and collaborative projects (Espinosa et al., 2019). Active learning strategies, particularly those focused on inquiry, appear to correlate positively with female academic performance in mathematics (Johnson et al., 2020).

Finally, some instructors may have to recognize their own biases as well as fixed notions of intelligence and the purposes of assessment (Audisio et al., 2022; Black & Wiliam, 2018; Dweck, 1999, as cited in Heritage, 2010). They should thus exercise caution in communicating the purposes of test-taking—for example, avoid telling learners that the purpose of the test is judgment or evaluation of intelligence or ability, or telling or intimating to learners that one gender does better or worse than another in math, science, or reading (Heissel et al., 2021; Steele, 1999).

Addressing academic dishonesty

In addition to poor test design and practices, computer-based testing in particular, and assessment in general, is affected by a second weakness—cheating. The ease of finding information online also makes it easier for examinees to cheat, plagiarize, and "game" the technology system, manipulating the technology features of games, Intelligent Tutoring System or Computer Aided Instruction to arrive at the correct answers instead of actually wrestling with content (Baker et al., 2010).

The extent of cheating online is unknown. In some surveys, 93% of instructors and 95% of learners say that learners are *likely* to cheat, plagiarize, and copy-and-paste from the World Wide Web without attribution—but real data are hard to come by, and the perception may be far from reality (Wiley Publishing, 2020). What is known is that learners are more likely to cheat when they are under pressure; when they feel alone or unsupported and are unmotivated; when they feel the rewards of cheating outweigh the risks; and when they have a poor or non-existent relationship with the instructor (Lederman, 2020; Maeda, 2019). Whatever the root cause, issues of plagiarism and cheating pose the most fundamental threat to assessment in teacher distance education programs, raising existential questions about the purpose and validity of teacher assessments (Letseka & Pitsoe, 2013).

Digital citizenship and appropriate online communications were discussed in *Chapter 14: Preparing Distance Learners*. But the rubric of "digital citizenship" also encompasses academic integrity and ethical uses of technology (Wiley Publishing, 2020). Digital citizenship is not simply for children and adolescents but is an important mindset and skill for their teachers as well. While many education systems stress digital citizenship for adult educators—not just students—many others do not.

Different distance education systems will have different approaches to issues of online cheating, plagiarism, and violation of copyright by pre- and in-service teachers. Some may have no policies or proscriptions against such behaviors. Some may be high-trust programs that focus on educating teacher-learners about academic integrity (or not), and trust that the executive functioning and moral values of such teachers will deter them from academic dishonesty. Some systems may adopt zero-trust policies and control technology to such an extent that cheating may be extremely difficult.

The best approach to minimizing academic dishonesty may be a combination of educating and trusting teacher candidates, strong academic policies, good pedagogy, and careful design of assessments. Figure 17.4 enumerates approaches for potentially reducing cheating, copyright violations, and plagiarism in CBAs and online courses.

Figure 17.4
Approaches to Reduce Computer-based Cheating and Plagiarism

Approach	Suggested Actions
Education	• Educate learners about the importance of plagiarism, fair use, copyright, and academic integrity.
Policies	• Ensure that every online course has an honor code that explicitly details what constitutes cheating and the repercussions of cheating. Have learners co-design and validate this honor code. • Create and enforce strong academic honesty policies and acceptable use policies. • Ban cell phones and other devices in examination rooms (Wiley Publishing, 2020).

Approach	Suggested Actions
Instruction	• "Teach better:" More learner-centered, collaborative, higher-order, personalized instruction versus traditional, rote-based learning to make cheating harder (Lederman, 2020). • Provide learners with low-stakes quizzes and scaffolded assessments to practice their skills, ability, and knowledge without worrying about grades (Lederman, 2020).
Assessment design	• Create collaborative assessments so learners feel supported and are less likely to cheat for fear of hurting their peers, and because it is simply harder to do so (Lederman, 2020). • Create "open-book" tests (Lederman, 2020). • Provide learners with choice in terms of their assessments. • Design exams with conditional branching—where an exam moves to a different question based on a certain answer or condition being met. • Pose questions that relate to specific and unique course events, as opposed to general concepts, as deterrents to plagiarism. • Use a range of assessment formats—for example, computer-based, performance-based, and face-to-face.
Assessment administration	• Stagger the time of assessments and impose time limits (Wiley Publishing, 2020). • Randomly sequence exam questions, provide learners with different essay questions, or provide the same assessments but with components that vary among learners (Wiley Publishing, 2020). • Assign different examination questions to different learners. • Employ paper-based tests and in-person individual oral assessments. • For high stakes assessments, administer them in a central location with proctors and invigilators.
Technical solutions	• Copy and paste essays into a Web-based search engine to determine authorship; use plagiarism detection tools such as *Turnitin*, *GPTZero, Plagiarism Checker X*, and Microsoft *Word*'s "Similarity" feature (under "Editor"), to detect plagiarism.
Technical solutions *(continued)*	• Employ a computer-based virtual proctoring system that installs a proctor (a camera) at each computer workstation to monitor that learner throughout the exam. The room also can be outfitted with cameras that provide a bird's-eye view. Once there is evidence that a learner has cheated, the computer-based exam locks down and remains that way until video recordings are examined and a decision is reached. • Lock down browsers during online exams. • Design and develop basic password certificates based on authentication methods (Chirumamilla & Sindre, 2019). • Use sophisticated biometrics to identify users, so that one friend cannot take an exam for another (Chirumamilla & Sindre, 2019). • Use improved facial recognition software tools to help authenticate the identity of the test taker (Chirumamilla & Sindre, 2019).

17.4 Better Assessment Within a Distance Education System

There are several strategies for developing both formative and summative assessment of learners within any distance education model. We discuss some of the major ones here.

17.4.1 Develop Standards as Determinants of Success

The myriad skills and behaviors associated with good teaching often make measuring teacher quality or assessing the fitness of a pre-service candidate or an in-service teacher difficult—hence the reliance on grades and examination scores. Perhaps the most critical component of assessing teachers' readiness, fitness, or quality is to design standards for performance, instruct teachers according to these standards, and then measure teacher performance against them.

Standards can be *normative* (comparing one learner's performance with that of another); *criterion-based* (comparing a learner's performance with an empirically derived level of proficiency, such as a cut score that determines whether a learner has mastered a particular skill); or *ipsative* ("growth" model standards that involve using the learner's prior performance as the basis for comparison with his or her current performance) (Hosp, 2010, p. 5). Understanding the different types of standards is critical for test design, administration, and interpretation of assessment results.

17.4.2 Do Assessment *with* Teachers Not *to* Them

Within distance education programs, the teacher-learners who are being assessed have themselves often been missing from participation in the conceptualization, design, and administration of assessment. Since teaching and learning involve both instructors and learners, assessment should be a collaborative endeavor between "both parties in order to produce the best performance in teaching and learning" to create a shared "understanding of the criteria and standards by which quality learning will be assessed" (Letseka & Pitsoe, 2013, p. 204). As is often advocated for student assessment, teacher-learners must be involved in the assessment process itself.

"Flexible assessment" advocates that teachers be given voice in choosing the types of assessments that best represent their learning progression and that yield the best possible information and insights to improve teaching effectiveness and learning quality. As part of teacher education programs, teachers should learn how to design assessments for their own students, how to analyze and interpret assessment data, and how to implement strategies for leveraging data to adapt and modify instruction (Letseka & Pitsoe, 2013). This gap between the intent and implementation of assessment, and between instructors and learners, could begin to be bridged if instructors and distance program designers planned assessments as if learning—*their* learning—mattered most (Letseka & Pitsoe, 2013).

17.4.3 Treat Summative Assessments as Opportunities for Teaching and Learning

The distinctions between formative and summative assessment are often confusing. Homework may be the classic example of such confusion—it is thought of as formative (to assess for learning and to inform instruction). But the very act of grading it and using these grades to make a final determination about learner performance is summative (Black & Wiliam, 2018).

Black & Wiliam (2018) argue that summative assessments can, and should, be used, not just summatively but *formatively* as well. They can inform changes in instructor planning and implementation of courses or how best to work with future learners. They also may be formative for learners themselves, helping them reflect on the "strengths and weaknesses of their achievements in ways that might help them re-direct their energies in future work" (Black & Wiliam, p. 12).

17.4.4 Use Formative Assessment to Support Mastery Learning

Traditional instruction in a distance program often involves organizing the curriculum into chronological units or sequential blocks and assessing learners' understanding of the material at the end of each unit (Guskey, 2010, p. 53; see also Heritage, 2010). Yet, learning theory informs us that learners, such as pre-and in-service teachers, move through stages of learning from acquisition to fluency at different paces and that these stages often are unconnected to the sequence of topics (Hosp, 2010). Similarly, *assessment theory* states that we learn best when assessment is part of, not separate from, instruction. Thus, rather than assessing teacher-learners at one final level as a summative exercise independent from instruction, distance education courses should integrate assessment into instruction and use formative assessment to support teacher-learners at each stage of their learning. Bloom (1971) referred to this approach as *mastery learning*,[2] a process that involves the following steps:

1. **Diagnostic pre-assessment with pre-teaching.** Instructors administer a short pre-assessment to learners before instruction to determine whether they have the prerequisite knowledge and skills for success in the content they are about to study.
2. **Initial instruction.** The instructor then provides high-quality group instruction that is research-based, adapted to local conditions, and is differentiated to help learners at various stages of the learning process.
3. **Progress monitoring through regular formative assessment.** Following the initial instruction, the distance instructor administers a quick test that assesses learners' understanding and reinforces the most important learning objectives.
4. **Corrective instruction ("reteaching").** Following the formative assessment, the instructor provides corrective instruction or reteaching of the skills and concepts in which learners demonstrated difficulty. Reteaching involves making accommodations in the types of materials used and differentiating instruction—for example, by offering one-to-one tutoring for some learners, "think aloud" protocols with another, or having other learners engage in peer tutoring.
5. **A second formative assessment.** Following the above corrective activities, learners are given a second, similar type of formative assessment that helps determine the effectiveness of the corrective instruction, allows them to demonstrate proficiency in the concept, and provides a more reliable measure of learners' competencies than one, singly administered assessment.
6. **Enrichment or extension activities.** Mastery learning offers enrichment activities to provide challenging learning experiences to learners who do not need corrective instruction. This form of differentiated instruction allows learners who have easily grasped content to immerse themselves in more challenging learning situations, while the distance instructor offers remedial and corrective instruction to those who need it (Guskey, 2010, pp. 54–57).

While a distance instructor might typically lead this mastery learning for pre-service teachers, it may also be part of a coaching program for in-service teachers.

17.4.5 Measure Teacher Performance—Not Simply Knowledge

Ultimately, teacher learning at its foundation is about applied learning in an authentic context. While it is important to assess teacher knowledge *about* or *of* teaching, more important is a performance-based assessment measuring a teacher's ability to teach in a real place of practice—a classroom—as part of a teaching practicum (as discussed in Chapter

[2] Like many concepts in education, mastery learning has evolved and developed a slightly different but related meaning that now focuses on competencies that students, particularly in higher education, must master according to standards or national qualification frameworks (NQFs). The African Continental Qualifications Framework (ACQF) focuses on employability skills and examines the comparability, quality, and transparency of qualifications and supports lifelong learning: https://acqf.africa/. The European Union's National Qualifications Framework allows users to examine and compare NQFs across the E.U: https://europa.eu/europass/en/compare-qualifications

13), as part of an induction program (as outlined in Chapter 16), or as part of some kind of professional in-service program. The most common and direct measure is a classroom observation instrument. Classroom observation instruments are often rubric-like in their design and can be classified as either "low-inference" or "high-inference" in nature (see Figure 17.5).

A low-inference instrument, or "a category instrument," may be a checklist of observable indicators of teacher practice (Rosenshine, 1970, p. 281). These tools are easy to complete and can be administered by either less experienced or well-trained classroom observers. However, they measure only the presence or frequency of a behavior—not the quality—nor do they capture the complexity, breadth, and depth of teacher classroom practice.

High-inference tools, or rating systems, incorporate descriptive information or "constructs" of classroom practice and rate these along some sort of scoring scale (such as a Likert scale, from 1–5). With high-inference classroom observation tools, the observer must infer the constructs to be rated—such as the clarity of presentation or organization of learning—recording the frequency through such scales as "consistently," "sometimes," or "always" (Rosenshine, 1970). Because they involve a high degree of interpretation and inference, these observation forms should be used by well-trained observers who understand the purpose of the assessment, who have clear expectations of what each performance level looks like in practice, and who have undergone reliability training. Although they are more demanding to use, high-inference classroom observations, if used well, yield information that is both reliable and valid, better capturing the quality, complexity, and intricacies of classroom instruction (Rosenshine, 1970).

Distance education programs—specifically those with a unique focus, such as a particular reading program or instructional approach—may develop their own classroom observation tools to measure fidelity of implementation or transfer of learning. Developing reliable, valid, sensitive high-inference classroom observation tools that accurately measure specific and empirical constructs is particularly challenging. If not

Figure 17.5
Characteristics of High- Versus Low-Inference Classroom Observation Systems (Adapted from Rosenshine, 1970)

Characteristics	Low-Inference Observations	High-Inference Observations
General description	Descriptive	Inferential
Recording procedures	Categories	Signs and scale
Items	Low-inference (observation)	High-inference (judgment and interpretation)
Format	Checklist/binary (yes/no)	Likert scale or some other continuum
Coding	Simple coding	Multiple coding
Focus	Frequency	Quality
Observer skill required	Low	High
Reliability	Low	High

designed well, they may be not simply ineffective, but even harmful. Figure 17.6 outlines the purpose, appropriateness, strengths, and weaknesses of classroom observation tools. It is critical to remember that observations are a snapshot in time—one measure of potentially years of teaching—and should be combined with other forms of data before making determinations about an individual teacher's quality or ability.

Given the challenge of developing high-quality high-inference observation tools that reliably assess nuances of behavioral changes, distance education programs may want to avail themselves of the many established classroom observation tools that can be used to assess teacher performance. Examples include the Stallings Classroom Snapshot, the Marzano Teacher Evaluation Model, the

Figure 17.6
Classroom Observation Tools

Purpose
Directly assess the actual classroom practices of teachers. Unlike a survey, which is almost a secondary source of information, this measure is a primary source—direct and empirical.
Appropriateness
• Assess evidence of instructional changes, content knowledge, improved professional competencies. • A classroom observation form can assess only measurable and visible outcomes. • Teachers can be assessed along a continuum (low to high) or based on a checklist (yes/no). • Tools can be quantitative and qualitative.
Strengths
• The assessor is directly observing practice, so there is no "interference," as with surveys where teachers can hide true opinions. • Its performance-based nature makes it more objective, empirical, and valid than other types of measurement tools. • It records a fixed set of teacher behaviors, lending itself (in the best-case scenario) to focusing on discrete areas of teacher behavior that can be targeted for improvement.
Weaknesses
• It is surprisingly hard to do good observations—issues of observer bias, observer boredom, confirmation bias, Hawthorne effect (people often perform better when being observed), "halo" effect (judging a certain teacher "high" based on prior positive impressions), performance bias (people rehearse for observer), and indeterminacy are common. Therefore, observer training is a must. • High-inference observation forms demand that an observer be very well trained and able to differentiate among performance levels (using 4- and 5-point scales).

Cost Considerations

- High: Site selection, sufficient sample size (for a large-scale program, this would be in the hundreds); training for observers; transportation for observers; joint agreement by observers (filling out one observation protocol between them).
- Transcribing and analyzing qualitative information, quantitative data analysis, and report writing all add to cost.

Classroom Assessment Scoring System (CLASS), the Classroom Observation Toolkit for Early Grade Reading Improvement, and the Danielsen Framework. The classroom observation tool selected, adapted, or created by a distance program should be fit for purpose; include clear, observable, and measurable expectations based on standards of instructional excellence; utilize a standardized observation and scoring protocol; have an evidence base that proves its reliability and validity; and use multiple ratings (i.e., be high-inference) and multiple measures (The New Teacher Project, 2011, p. 3).

17.4.6 Integrate Feedback into All Stages of Assessment

Chapter 9: Professional Development focused on the importance of feedback in teachers' professional development, and *Chapter 13: Preparing Distance Instructors* emphasized that timely, actionable feedback is associated with online learners' satisfaction with an online course. The most fundamental benefit of feedback may be that its significant, measurable effects on learner performance are at the core of good assessment.

Feedback is a highly diverse construct with multiple:

- *types* of feedback: intrinsic and extrinsic, concurrent and terminal, immediate and delayed, and separate and accumulated;
- *recipients* for feedback (a class, small group, an individual learner);
- *deliverers* of feedback (instructor, learners, a software program, the learner);
- feedback *inflection points* (stepwise, answer-wise, the final results of an assessment, or as part of an actual task); and,
- *purposes* of feedback (motivation, knowledge of information, knowledge about one's performance) (Druckman & Bjork, 1994, p. 50).

Feedback helps learners close the gap between where they are in the learning process and where they should be. It consists of the following four stages that constantly "loop back" or form a cycle:

1. **Evidence.** The data or information about performance should be measured and stored.
2. **Communication.** Information is conveyed to the individual, not as raw data but in a format that makes it emotionally resonant and relevant to the person.
3. **Consequence.** The information must illuminate a specific path forward.
4. **Action.** The individual recalibrates behavior, makes choices, and acts on them (Goetz, 2011, p. 130). The cycle/loop begins again, ideally with each loop becoming shorter and more narrowly focused (Bandura, 1986).

Feedback, like assessment, is essential to continuous improvement. Giving individuals a clear goal and the means of evaluating their progress toward that goal increases the likelihood that they will attain their goal (Bandura, 1986).

17.4.7 Design Assessments That Measure Higher-Order Thinking

It is easier, faster, and less expensive to design assessments that measure learners' recall of discrete and decontextualized facts versus higher-order thinking skills. Yet as every professional knows, beyond the academic environment, we are rarely measured on our ability to furnish declarative knowledge (facts). Rather, we are judged on our professional skills, conceptual knowledge, procedural knowledge, aptitudes, and disposition.

Although there are numerous types of knowledge-based taxonomies (e.g., Marzano's taxonomy of educational objectives or Sternberg's triarchic theory of intelligence), most enduring is Bloom's taxonomy of the cognitive domains of learning. For Bloom, learning occupied a continuum—from "lower level" or "lower order" thinking skills such as *knowledge* and *comprehension* to "higher level" or "higher order" learning or thinking skills *such as application, analysis, synthesis,* and *evaluation.* All of these are outlined in Figure 17.7.

Figure 17.7
Cognitive Domains of Learning (Bloom, 1956)

Levels (Orders) of Learning/ Thinking Skills	Competence	Skills Demonstrated
Lower-order learning/ thinking skills	**Knowledge**	• Observation and recall of information • Knowledge of dates, events, places • Knowledge of major ideas • Proficiency in subject matter *Question cues:* list, define, tell, describe, identify, show, label, collect, examine, tabulate, quote, name, who, when, where
	Comprehension	• Understand information • Grasp meaning • Translate knowledge into new context • Interpret facts, compare, contrast • Order, group, infer causes • Predict consequences *Question cues:* summarize, describe, interpret, contrast, predict, associate, distinguish, estimate, differentiate, discuss, extend

Levels (Orders) of Learning/ Thinking Skills	Competence	Skills Demonstrated
Higher-order learning/ thinking skills	**Application**	• Use information • Use methods, concepts, theories in new situations • Solve problems using required skills or knowledge *Question cues:* apply, demonstrate, calculate, complete, illustrate, show, solve, examine, modify, relate, change, classify, experiment, discover
	Analysis	• See patterns • Organize parts • Recognize hidden meanings • Identify components *Question cues:* analyze, separate, order, explain, connect, classify, arrange, divide, compare, select, infer
	Synthesis	• Use old ideas to create new ones • Generalize from given facts • Relate knowledge from several areas • Predict, draw conclusions *Question cues:* combine, integrate, modify, rearrange, substitute, plan, create, design, invent, compose, formulate, prepare, generalize, rewrite, what if?
	Evaluation	• Compare and discriminate between ideas • Assess value of theories, presentations • Make choices based on reasoned argument • Verify value of evidence • Recognize subjectivity *Question cues:* assess, decide, rank, grade, test, measure, recommend, convince, select, judge, explain, discriminate, support, conclude, compare, summarize

Distance-based assessment systems are faced with the challenge of preparing teachers to teach in a 21st-century educational and economic environment that emphasizes Bloom's "higher-order thinking skills." For teachers to help students develop such skills, teachers themselves must develop these "critical-thinking" faculties. And for teachers to develop these skills, distance education systems must instruct, model, and assess teacher-learners in higher-order ways and do so while assessing both the product of learning and the thinking process of learners (Royal & Guskey, 2015). Higher-level thinking for teachers is particularly crucial in a knowledge-based economy and in an era of rampant social media disinformation.

As Figure 17.8 demonstrates, assessing higher-order thinking skills demands a variety of distinct assessment methods and tasks.

Figure 17.8
Specific Strategies for Assessing Higher-Order Thinking (Brookhart, 2010, pp. 144–147)

To assess how well learners can . . .	Provide this kind of material . . .	And ask learners to . . .
Focus on a question/ identify the main idea	• Text, speech, problem, policy, or experiment and results	• Identify the main issue, main idea, problem, and explain their reasoning
Analyze arguments	• Text, speech, or experimental design	• Identify what evidence the author gives that supports/contradicts the argument • Identify assumptions that must be true to make the argument valid • Explain the logical structure of the argument (including irrelevant and contradictory structures)
Compare and contrast	• Two texts, events, scenarios, theories, experiments, or works of art	• Identify elements in each • Organize elements based on their similarities and differences
Evaluate materials and methods for their intended purposes	• Text, speech, problem, policy, or experiment and results	• Identify the author/designer's purpose • Identify elements in the work • Judge the value and validity of these elements in accomplishing the intended purpose • Explain their reasoning and support it with evidence
Put unlike concepts together in new ways	• Complex task/problem	• Generate multiple solutions • Produce something new
Make or evaluate a ***deductive*** conclusion	• Statement or premise	• Draw a logical conclusion based on reasoning and evidence • Select a logical conclusion from a set of choices
Make or evaluate an ***inductive*** conclusion	• Statement, scenario, information in form of graph/chart, or set of examples	• Formulate a hypothesis • Test hypothesis and revise • Formulate a definition or concept based on examples and nonexamples
Identify/define a problem	• Scenario or problem description	• Identify the problem that has to be solved • Identify the question that has to be answered
Reason with data	• Text, graph, chart, data table, or problem that requires more information or a solution	• Solve the problem and explain reasoning using data
Think creatively	• Complex problem/task requiring brainstorming innovative ideas or reorganizing existing ideas or a problem with no currently known solution	• Produce an original text, product, concept, or idea • Organize materials in new ways • Reframe a question/problem in new ways

17.4.8 Use Inferences from Multiple and Balanced Sources of Evidence versus One Sole Source of Evidence

The many teachers with whom distance programs will interact are extremely diverse. They are diverse in terms of language ability, experience, time in the classroom, gender, and educational and professional opportunities (Voltz et al., 2010). They are diverse in their approaches and attitudes toward learning. They are diverse in their likes and dislikes, in their personal strengths and weaknesses, and in their levels of commitment to teaching and learning. Therefore, just as no distance education system can impose a one-size-fits-all instructional approach, no distance education system can impose a one-size-fits-all assessment approach. It is important to make any assessment system as diverse as possible in order to be as fair and sensitive[3] as possible to a variety of learners. "Fairness" does not mean that every teacher-learner receives the same test—although for some purposes, standardized and normative tests may be necessary. Rather, it means that every learner has an equal opportunity to be assessed in the manner that best displays what he or she knows and can do (Voltz et al., 2010; Royal & Guskey, 2015).

To do this, distance programs can use a wider range of context-based, complex tasks that can be used with multiple approaches and solutions, instead of using only assessment items that are short, knowledge-focused, single-answer, and decontextualized (Black & Wiliam, 2018; Heritage, 2010; Hosp, 2010; Moon et al., 2005; Royal & Guskey, 2015; Timperley et al., 2007). They can also differentiate grading—employing self-, peer-, and instructor assessment of the products, processes, and progress of learning, both separately and cumulatively.

One model of using multiple and balanced sources of evidence comes from the National Board of Certification[4] in the United States. Teachers applying for this certification are assessed on 10 measures, including an examination of content knowledge, a comprehensive portfolio of teacher practice and student work, and interview-based methods developed by the National Board for Professional Teaching Standards (National Board for Professional Teaching Standards, 2022).

17.4.9 Where Possible and Appropriate, Take Advantage of Technology for Assessment

The past decade has seen breakthroughs in technology-based assessment that measures complex thinking; lowers the cost differential of assessment, because assessment takes less time to score and store; enables quick turnaround of assessment data to the instructor and learners; helps instructors to assess learner performance at a much more granular, detailed level; and allows for more reliable scoring and valid data interpretation (Burns, in press).

Assessment must be part of every mode of distance education delivery. While certain types of distance education have more opportunities to assess learners than other forms (e.g., online learning versus interactive audio instruction), all types of technology combined with assessment theory can identify new and better ways to assess what matters; conduct formative assessment; and involve multiple stakeholders in the formulation, design, administration, and analysis of assessment data (Morris et al., 2021).

17.4.10 Provide Language Supports to Teachers Who May Need Them

Before concluding this chapter, it is important to note that in many countries in which this guide will

[3] "Sensitive" here is used in an assessment sense—designing instruments in such a way that they accurately measure what they are supposed to measure.

[4] National Board Certification is an advanced teaching credential in the United States that complements, but does not replace, a U.S. state's teacher license. It is valid for 10 years. National Board Certification is achieved upon successful completion of a voluntary assessment program designed to recognize effective and accomplished teachers who meet high standards based on what teachers should know and be able to do. See http://www.nbpts.org/ for more information.

be read, teachers undoubtedly speak a number of languages other than the official or national language. The importance of providing assessment in a learner's first language has been widely acknowledged as a best practice in assessment. Numerous distance education systems have made reasonable accommodations for non-national language speakers by making test taking more flexible to allow examinees to have time to think and respond to questions in the national language, or by providing dictionaries, plug-ins that provide modifications such as translations and closed captioning, and visual and audio information to learners in both their native language and the language of instruction (e.g., Kannada, English, Bambara, French, Sundanese, Bahasa Indonesia, etc.) (Myung et al., 2020). Better still, of course, is conducting the assessment entirely in the learner's first language (Reid & Kleinhenz, 2015).

17.5 Conclusion

Realigning assessment within distance learning programs toward recognized best practices involves a number of approaches that will be new in many systems. These include defining and analyzing instructional quality into discrete measurable indicators that monitor teachers' progress and learning (formative assessment) and evaluating their final performance on the most critical components of teaching (summative assessment).

Quality distance education programs embrace and enact these practices. They recognize that assessment, even when summative, always has a formative component; that is, instructors should always use assessment results to further refine instruction within a distance environment. They use a multitude of measures—performance-based assessment, and traditional and alternative assessments—to assess how teachers are learning, whether they are learning, and what they are learning. They recognize that assessment—of learning, of instruction, or learners' progress, process, and products of learning—is the foundation on which a quality distance program rests.

Citation: Burns, M. (2023). Assessing Distance Learners. In *Distance Education for Teacher Training: Modes, Models and Methods.* (2nd Edition). Washington, DC: Education Development Center.

References

Archer, E. (2017). The assessment purpose triangle: Balancing the purposes of educational assessment. *Frontiers in Education*. https://doi.org/10.3389/feduc.2017.00041

Audisio, A., Taylor-Perryman, R., Tasker, T., & Steinberg, M. (2022). *Does teacher professional development improve student learning? Evidence from leading educators' teacher fellowship model. (EdWorkingPaper: 22-597).* Annenberg Institute at Brown University: https://doi.org/10.26300/ah2f-z471

Baker, R., Mitrović, A., & Mathews, M. (2010). User modeling, adaptation and personalization: Detecting gaming the system in constraint-based tutors. *Lecture Notes in Computer Science, 6075*, 267–278. doi:DOI: 10.1007/978-3-642-13470-8_25

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Prentice Hall.

Batsell Jr. , W., Perry, J., Hanley, E., & Hosseter, A. (2017). Ecological validity of the testing effect: The use of daily quizzes in introductory psychology. *Teaching of Psychology, 44*(1), 18–23. https://doi.org/10.1177/0098628316677492

Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in Education, 25*(3), 1–28. https://doi.org/10.1080/0969594X.2018.1441807

Bloom, B. (1956). *Taxonomy of educational objectives, Handbook: The cognitive domain.* David McKay.

Bloom, B. (1971). Mastery learning. In J. Block (Ed.), *Mastery learning: Theory and practice* (pp. 47–63). Holt, Rinehart and Winston.

Brame, C. (2013). *Writing good multiple choice test questions.* Center for Teaching Vanderbilt University. https://cft.vanderbilt.edu/guides-sub-pages/writing-good-multiple-choice-test-questions

Brame, C., & Biel, R. (2015). Test-enhanced learning: The potential for testing to promote greater learning in undergraduate science courses. *Cell Biology Education—Life Sciences Education, 14*, 1–12. https://doi.org/10.1187/cbe.14-11-0208

Brookhart, S. (2010). *How to assess higher-order thinking skills in your classroom.* Association of Supervision and Curriculum Development.

Buckley, J., Colosimo, L., Kantar, R., McCall, M., & Snow, E. (2021). Game-based Assessment for Education. *OECD Digital Education Outlook 2021: Pushing the Frontiers with Artificial Intelligence, Blockchain and Robots*. https://doi.org/10.1787/589b283f-en

Burns, M. (2018, August 13). *Five guidelines for developing good online assessments*. eLearning Industry. https://elearningindustry.com/developing-good-online-assessments-guidelines

Burns, M. (2021). *Background paper prepared for the 2023 Global Education Monitoring report, Technology and education: Technology in education.* United Nations Educational, Scientific and Cultural Organization. https://unesdoc.unesco.org/ark:/48223/pf0000378951

Burns, M. (in press). *Barriers and supports for technology integration: Views from teachers. Background chapter prepared for the Global Education Monitoring report: Technology and education.* United Nations Educational, Scientific and Cultural Organization.

Chirumamilla, A., & Sindre, G. (2019). Mitigation of cheating in online exams: Strengths and limitations of biometric authentication in online learning environments. In A. V. Senthil Kumar (Ed.), *Biometric authentication in online learning environments* (pp. 47–68). IGI Global. https://doi.org/10.4018/978-1-5225-7724-9.ch003

Commonwealth of Learning & Asian Development Bank (Eds.). (2008). *Quality assurance in open and distance learning: A toolkit.*

Contact North | Contact Nord. (2020, May 12). *Ten Radical ways assessment is changing*. TeachOnline.ca. https://teachonline.ca/ai-resources/ten-radical-ways-assessment-changing

Downing, S. (2006). Selected-response item formats in test development. Handbook of test development, 287, 301. In T. Haladyna, & S. Downing (Eds.), *Handbook of test development.* Educational Testing Service.

Druckman, D., & Bjork, R. (1994). *Learning, remembering, believing: Enhancing human performance.* National Academy Press.

Espinosa, T., Miller, K., Araujo, I., & Mazur, E. (2019). Reducing the gender gap in students' physics self-efficacy in a team-and project-based introductory physics class. *Physical Review Physics Education Research, 15*(1). doi:10.1103/PhysRevPhysEducRes.15.010132

Gee, J. P., & Shaffer, D. W. (2010). *Looking where the light is bad: Video games and the future of assessment. (Epistemic games group working paper No. 2010-02).* University of Wisconsin-Madison. https://www.semanticscholar.org/paper/Looking-where-the-light-is-bad:-Video-games-and-the-Gee-Shaffer/d0d7c77a88415838206f317c899e1ad7975e89ed

Goetz, T. (2011, July). The feedback loop: How technology has turned an age-old concept into an exciting new strategy for changing human behavior. *Wired*, pp. 126–162.

Guskey, T. (2010, October). Lessons of mastery learning. *Educational Leadership, 68*(2), 53–57.

Haagsman, M., Scager, K., Boonstra, J., & Koster, M. (2020). Pop-up questions within educational videos: Effects on students' learning. *Journal of Science Education and Technology, 29*(6), 713–724. https://doi.org/10.1007/s10956-020-09847-3

Heissel, J., Adam, E., Doleac, J., Figlio, D., & Meer, J. (2021, Spring). Testing, stress, and performance: How students respond physiologically to high-stakes testing. *Education Finance and Policy, 2*, 183–208. https://doi.org/10.1162/edfp_a_00306

Heritage, M. (2010). *Formative Assessment and Next-Generation Assessment Systems: Are We Losing an Opportunity?* Council of Chief State School Officers. https://files.eric.ed.gov/fulltext/ED543063.pdf

Hosp, J. (2010, December). *Linking assessment and instruction: Teacher preparation and professional development. TQ connection issue paper.* National Comprehensive Center for Teacher Quality. https://files.eric.ed.gov/fulltext/ED520827.pdf

Hultberg, P., Calonge, D., & Lee, A. (2018). Promoting long-lasting learning through instructional design. *Journal of the Scholarship of Teaching and Learning, 18*(3), 26–43. https://doi.org/10.14434/josotl.v18i3.23179

Johnson, E., Andrews-Larson, C., Keene, K., Melhuish, K., Keller, R., & Fortune, N. (2020). Inquiry and gender inequity in the undergraduate mathematics classroom. *Journal for Research in Mathematics Education, 51*(4), 504–516. https://doi.org/10.5951/jresematheduc-2020-0043

Kulhavy, R., & Stock, W. (1989). Feedback in written instruction: The place of response certitude. *Educational Psychology Review, 1*(4), 279–308. https://doi.org/10.1007/BF01320096

Lederman, D. (2020, July). *Best way to stop cheating in online courses? 'Teach better.'* Inside Higher Education. https://www.insidehighered.com/digital-learning/article/2020/07/22/technology-best-way-stop-online-cheating-no-experts-say-better

Letseka, M., & Pitsoe, V. (2013). Reflections on assessment in open distance learning (ODL): The case of the University of South Africa (UNISA). *Open Praxis, 5*(3), 197–206. https://www.learntechlib.org/p/130673/

Maeda, M. (2019). Exam cheating among Cambodian students: When, how, and why it happens. *Compare: A Journal of Comparative and International Education*, 1–19. https://doi.org/10.1080/03057925.2019.1613344

Momsen, J., Long, T., Wyse, S., & Ebert-May, D. (2017, Winter). Just the facts? Introductory undergraduate biology courses focus on low-level cognitive skills. *Cell Biology Education, 9*(4), 435-440. https://doi.org/10.1187/cbe.10-01-0001

Moon, B., Leach, J., & Stevens, M. (2005). *Designing open and distance learning for teacher education in Sub-Saharan Africa: A toolkit for educators and planners.* World Bank. http://oro.open.ac.uk/8401/1/teacher_education_toolkit_may13.pdf

Morris, E., Farrell, A., & Venetis, E. (2021, January). *A roadmap for for measuring distance learning: A Review of evidence and emerging practices.* United States Agency for International Development. https://www.edu-links.org/sites/default/files/media/file/Measuring%20Impact%20and%20Outcomes_Final_01.20.2021-508%20%281%29.pdf

Myung, J., Gallagher, A., Cottingham, B., Gong, A., Kimner, H., Witte, J., . . . Hough, H. (2020). *Supporting learning in the COVID-19 context: Research to guide distance and blended instruction.* Policy Analysis for California Education (PACE). https://files.eric.ed.gov/fulltext/ED609208.pdf

National Board for Professional Teaching Standards. (2022, August). *National board certification*. https://www.nbpts.org/certification/

Niess, M. (2011). Investigating TPACK: Knowledge growth in teaching with technology. *Journal of Educational Computing Research, 44*(3), 299–317. https://doi.org/10.2190/EC.44.3.c

Pandey, C., & Kapitanoff, S. (2011). The influence of anxiety and quality of interaction on collaborative test performance. *Active Learning in Higher Education, 12*(3), 163–174. https://doi.org/10.1177/1469787411415077

Reid, K., & Kleinhenz, E. (2015). *Supporting teacher development: Literature review.* Australian Council for Educational Research.

Reville, S., Coggins, C., Candon, J., McDermott, K., Churchill, A., & Long, B. (2005). *Reaching capacity: A blueprint for the state role in improving low performing schools and districts.* Rennie Center for Education Research & Policy.

Roediger III, H., & Butler, A. (2011). The critical role of retrieval practice in long-term retention. *Trends in cognitive sciences, 15*(1), 20–27. https://doi.org/10.1016/j.tics.2010.09.003

Rosenshine, B. (1970). Evaluation of classroom instruction. *Review of Educational Research, 40*(2), 279–300. https://doi.org/10.3102/00346543040002279

Royal, K., & Guskey, T. (2015). A case for differentiated grades. *Medical Science Educator, 25*(3), 323–325. doi:DOI: 10.1007/s40670-015-0127-5

Schacter, D., & Szpunar, K. (2015). Enhancing attention and memory during video-recorded lectures. *Scholarship of Teaching and Learning in Psychology, 1*(1), 60–71. https://doi.org/10.1037/stl0000011

See, B. (2022, August 15). Where have we gone wrong in our battle against teacher shortages? *Open Access Government*. https://www.openaccessgovernment.org/where-have-we-gone-wrong-in-our-battle-against-teacher-shortages/140272/

Spencer, S., Logel, C., & Davies, P. (2016). Stereotype threat. *Annual Review of Psychology, 67*, 415–437. https://doi.org/10.1146/annurev-psych-073115-103235

Steele, C. (1999, August). Thin ice: Stereotype threat and Black college students. *The Atlantic*. https://www.theatlantic.com/magazine/archive/1999/08/thin-ice-stereotype-threat-and-black-college-students/304663/

Tempelaar, D., Rienties, B., & Giesbers, B. (2015). Stability and sensitivity of learning analytics based prediction models. *Proceedings of 7th International conference on Computer Supported Education* (pp. 156-166). CSEDU.

The New Teacher Project. (2011, February). *Rating a teacher observation tool: Five ways to ensure classroom observations are focused and rigorous.* The New Teacher Project. https://tntp.org/assets/documents/TNTP_RatingATeacherObservationTool_Feb2011.pdf

Timmis, S., Broadfoot, P., Sutherland, R., & Oldfield, A. (2015). Rethinking assessment in a digital age: Opportunities, challenges and risks. *British Educational Research Journal Early View.* doi:10.1002/berj.3215

Timperley, H., Wilson, A., Barrar, H., & Fung, I. (2007). *Teacher professional learning and development: best evidence synthesis iteration.* New Zealand Ministry of Education. http://www.educationcounts.govt.nz/publications/series/2515/15341

Tropman, E. (2014). In defense of reading quizzes. *International Journal of Teaching and Learning in Higher Education, 26*(1), 140–146. https://files.eric.ed.gov/fulltext/EJ1043037.pdf

Voltz, D., Sims, M., & Nelson, B. (2010). *Connecting teachers, students and standards: Strategies for success in diverse and inclusive classrooms.* Association for Supervision and Curriculum Development.

Warschauer, M. (2009). Learning to write in the laptop classroom. *Writing and Pedagogy, 1*(1), 101–112. doi:doi: 10.1558/wap.v1i1.101.

Wiley Publishing. (2020, May). *Academic integrity in the age of online learning.* John Wiley & Sons (Wiley). https://www.wiley.com/en-us/network/education/instructors/teaching-strategies/academic-integrity-in-the-age-of-online-learning-3

Wyatt-Smith, C., Lingard, B., & Heck, E. (2019). *Digital learning assessments and big data: Implications for teacher professionalism.* UNESCO education research and foresight-Working papers. https://unesdoc.unesco.org/ark:/48223/pf0000370940


EDC
Education
Development
Center