Future directions in assessment: Influences of standards and implications for language learning

Troy L. Cox  
*Brigham Young University,*  
troyc@byu.edu

Margaret E. Malone

Paula Winke  
*Michigan State University*

Follow this and additional works at:  
https://scholarsarchive.byu.edu/facpub

Original Publication Citation  
https://doi.org/10.1111/flan.12326

BYU ScholarsArchive Citation  
https://scholarsarchive.byu.edu/facpub/5891

This Peer-Reviewed Article is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in Faculty Publications by an authorized administrator of BYU ScholarsArchive. For more information, please contact ellen_amatangelo@byu.edu.
Future directions in assessment: Influences of standards and implications for language learning

Troy L. Cox1 | Margaret E. Malone2 | Paula Winke3

Challenges
This paper explores how standards-based instruction and assessment have evolved and suggests areas in which future work may be helpful. How can advances in the field of applied linguistics inform our thinking? How will new approaches to assessment drive alternative conceptions of language teaching and pathways to language learning?

As Foreign Language Annals concludes its 50th anniversary, it is fitting to review the past and peer into the future of standards-based education and assessment. Standards are a common yardstick used by educators and researchers as a powerful framework for conceptualizing teaching and measuring learner success. The impact of standards on language assessment, teaching, curricula, course design, and educational policy is indisputable, but can they even be more impactful, more beneficial? In this article, we reflect upon the role of language learning standards on world language practices and assessments and discuss standards’ design, implementation, and appropriation issues that will challenge the field over the next few decades. Although predicting the future is risky, forward thinking is critical when examining an issue as large and complex as the teaching and learning of world languages.

KEYWORDS
standards-based assessment, standards-based teaching

1 Brigham Young University
2 ACTFL
3 Michigan State University

Troy L. Cox (PhD, Brigham Young University) is Associate Director, Center for Language Studies, Brigham Young University, Provo, UT.

Margaret E. Malone (PhD, Georgetown University) is Director, Center for Assessment, Research and Development, ACTFL, and Research Professor and Director, AELRC, Georgetown University, Washington, DC.

Paula Winke (PhD, Georgetown University) is Associate Professor, MA TESOL and Second Language Studies programs, Michigan State University, East Lansing.
INTRODUCTION

In the United States and around the world, education and assessment are now, unequivocally, standards-based (Glisan, 2012; Llosa, 2011). In the 1980s and 1990s, “pressure for accountability from the government, taxpayers, and policymakers initiated a systematic student outcomes assessment movement” (Ricardo-Osorio, 2008, p. 590), and the influence of that movement has now reached teaching and learning from the university level to the preschool level and across a range of academic disciplines, including the discipline of foreign and second language pedagogy. Language learning standards have streamlined and improved students’ development by providing educators with “a common yardstick” (Bärenfänger & Tschirner, 2008)—that is, a shared pedagogical framework with theory-driven and evidence-based benchmarks and goals. The standards are educational road maps that were designed to help educators provide strong and effective learning opportunities for all learners, regardless of their geographic or socioeconomic context. But more important, as Bärenfänger and Tschirner pointed out, national and international language learning standards were designed as a pedagogical tool for educators so that they could reflect on their own programs and seek to improve their curricular practices and the outcomes stemming from them. The notion was that educators could use the standards to “perhaps even revolutionize language education,” as described by Glisan (2012, p. 515).

As Foreign Language Annals concludes its 50th anniversary, it is fitting to review the past and peer into the future of standards-based education and assessment. We do this because the standards have not quite revolutionized language education, but they have improved much and certainly are, as Glisan (2012) noted, impacting language instruction, curricula, course design, and educational policy. In this article we reflect upon the somewhat nebulous issues in standards-based world language practices and assessments that will challenge the field over the next few decades. This is necessary because while language learning standards are accepted and not too controversial (compared to national standards in other subject areas, such as social studies or history; see Reagan & Osborn, 2002, pp. 7–8), they (and their use) are not entirely without controversy. Predicting the future is always dangerous yet forward thinking is critical, particularly, we think, when examining an issue as large and complex as the teaching and learning of world languages.

LOOKING BACK

2.1 Assessments came first

As Echevarria, Short, and Powers explained (2006, p. 195, referencing Tucker & Codding, 1998), “[s]tandards and the assessments that are aligned with them have become the rallying principles for improved academic performance in schools,” and the improvement is normally measured through assessments. Thus assessment is an integral part of the standards movement. As a strong player in the standards movement, ACTFL began developing and validating reliable, criterion-referenced assessments of speaking well before the standards movement began in the United States in the late 1980s and 1990s (Herzog, 2003; Menken, Hudson, & Leung, 2014). The speaking assessments, and now also the assessments of listening, reading, and writing, were at first designed with college and university students and even older learners in mind. They were based on earlier versions of the ACTFL (2012) Proficiency Guidelines, which in turn were based on the U.S. Civil Service Commission’s 1952 Interagency Language Roundtable (ILR) proficiency scale descriptors (see Herzog 2003), which themselves were created in response to real-world needs: specifically, the U.S. government’s pragmatic need to appropriately assign language learners to foreign language–dependent jobs that required various levels of linguistic skill (such as jobs in the U.S. Foreign Service). Thus the ACTFL
Proficiency Guidelines were a precursor to the foreign and second language standards, and in actuality they spun off the ILR scale, which was made to fairly and accurately assess the task-based language ability and intercultural communication competency of (predominantly) Foreign Service officers.

Briefly described, the ACTFL guidelines outline a hierarchy of global tasks that describe what a language user can consistently perform at one level (the floor) and what he or she cannot consistently perform at the next adjacent level (ceiling) across five major levels (Novice, Intermediate, Advanced, Superior, and Distinguished). The assessments that are based on these guidelines are thus based on real-world communication and are independent of how and in what context the language was used or learned: Students across educational institutions, learners with extensive time abroad, and heritage language learners can all be rated on the same scale. Language users are awarded a rating based on their performance of the floor and ceiling criteria at each major level; sublevel ratings (Low, Mid, and High from the Novice level through Advanced) are used to reflect the strength of the performance and the extent to which the test taker approaches the next level. In recent years, ACTFL and its testing arm, Language Testing International (LTI), have annually conducted more than 250,000 language tests for students, prospective teachers, government employees, and business professionals, although the speaking scale, which emphasizes interpersonal communication, continues to be the most widespread.

Because the ACTFL guidelines emerged from the ILR tradition, ACTFL’s focus has historically been on describing and measuring language proficiency in a way that allows test results to subsequently be used for a variety of purposes, such as to establish language exit goals, measure learner outcomes, and support high-quality teaching. To retain that rigor, LTI is the only provider of official ACTFL ratings (Malone, 2012), although most certainly, institutions and teachers use unofficial ACTFL-like descriptors (normally written with lower-case labels such as novice, intermediate, and advanced) to define students’ proficiency. This unofficial and widespread use, coupled with the official worldwide presence, demonstrates the extent to which the descriptors guide the operationalization, discussion, and investigation of the constructs underlying second and foreign language proficiency and confirms the extent to which the descriptors and their associated assessments have become the lingua franca in the United States and in contexts abroad, where the guidelines exist alongside other national and international scale descriptors, such as the influential Common European Framework of Reference (CEFR) for Languages (see Bärenfänger & Tschirner, 2008, for an overview), which we describe in more detail later in this article.

2.2 The standards followed

While the Proficiency Guidelines (ACTFL, 2012) and ACTFL proficiency assessments have influenced the world language teaching and learning community in the United States for many years, the effort to translate this vision into classroom practice began in 1996, with the publication of what were then known as the National Standards for Foreign Language Learning in the 21st Century (Glisan, 2012). Each subsequent revision of the Standards has outlined the necessary language competencies that support students’ ongoing learning of foreign and second languages and prepare students to succeed in careers that involve proficient use of languages other than English. Framed as an interconnected set of five goal areas (Communication, Cultures, Connections, Comparisons, and Communities; National Standards Collaborative Board, 2015, p. 27) and colloquially known as the 5 Cs, the Standards support the development of rigorous and well-articulated learning goals and facilitate the selection of learning experiences and materials that help students progress toward higher levels of proficiency. Moreover, the Standards have been adapted to align with other sets of national educational standards, including the Common Core State Standards for K–12 English language arts and mathematics (see http://www.corestandards.org/) and the College and Career Readiness
Standards, which focus on English language arts, literacy, and mathematics (see http://www.ccrscenter.org/). Such alignments demonstrate the interdependence of language learning and other content areas in the United States and the critical role that languages play in overall student success.

To better understand the extent to which the Standards have impacted the teaching and learning of world languages in the United States, ACTFL conducted an extensive survey of world language professionals in 2011. More than 89% of the more than 2,000 participants, most of whom reported teaching at the secondary, middle, and elementary school levels, indicated a familiarity with the standards (ACTFL, 2011, n.p.). In addition, 64% of respondents agreed or strongly agreed that their own department had made an effort to incorporate the Standards, and 72% indicated that they had made an effort to incorporate the Standards for teaching and learning (ACTFL, 2011, n.p.). In 2015, ACTFL partnered with language associations around the United States to release the 4th edition of the World-Readiness Standards for Language Learning (the Standards; National Standards Collaborative Board, 2015), which was accompanied by 14 language-specific supplements.

2.3 | Linking standards- and proficiency-based assessments and instruction

In 1997, ACTFL was awarded a large-scale grant from the U.S. Department of Education International Research and Studies Program to create a classroom-based assessment framework founded more squarely on the Standards (Adair-Hauck & Troyan, 2013) and thus support the work of classroom teachers who were adopting the Standards and needed a clear and valid connection between classroom foreign language teaching and assessment (see Zapata, 2016). ACTFL published the first integrated performance assessment (IPA) manual in 2003; Adair-Hauck, Glisan, Koda, Swender, and Sandrock (2006) reported on the use and adoption of this IPA in Foreign Language Annals, and Adair-Hauck, Glisan, and Troyan (2013) more recently produced an expanded guideline on how to design IPA tasks for classroom-based assessment. Unlike the standardized prepackaged ACTFL suite of speaking, listening, reading, and writing assessments offered through LTI, the IPA is a framework that was designed for teachers to use and adapt free of cost to evaluate students’ integrated abilities to read, listen, speak, and write to perform various assessment tasks. Using the IPA framework, teachers adopt or create a series of three tasks that incorporate the three modes of communication (interpretive, interpersonal, and presentational) and at least one other standard, e.g., culture or connections to other disciplines, which students work through in class over time, e.g., over three or more class sessions. As alluded to, the three tasks include (1) an interpretive communication phase (students listen to and/or read, then answer interpretive questions about a focal authentic text), (2) an interpersonal communication phase (students get feedback from Phase 1 and then pair up with a partner to discuss questions or act out and audio- or videorecord a task related to the text from Phase 1), and (3) a presentational communication phase (students present or share their research, ideas, or opinions on the topic from Phases 1 and 2). IPAs are scored by teachers (see Adair-Hauck et al., 2013) and feedback is tailored to help individual students set new learning targets and help teachers provide balanced, standards-, and proficiency-oriented instruction (see Kissau & Adams, 2016; Martel & Bailey, 2016).

2.4 | Efforts in Europe and beyond

From 2010 to 2012, ACTFL worked with partners in Europe to explore ways to establish a crosswalk between the ACTFL Proficiency Guidelines (2012) and the CEFR for Languages (http://www.coe.int/en/web/common-european-framework-reference-languages), the European standards for foreign language education we have mentioned (see Tschirner, 2012). In sum, the CEFR is a European language scale of foreign language proficiency and development that includes six levels, ranging from
A1 (the lowest) to C2 (the highest), and has independent descriptors and guides for measurement for more than 40 languages (Bärenfänger & Tschirner, 2008). At the time of the CEFR’s creation in the late 1990s and early 2000s, there were no official correspondences between the ACTFL and CEFR scales, both of which are used for learning, teaching, and assessing foreign language (Tschirner, 2012). The CEFR framework and ACTFL guidelines represent different approaches to language learning. The CEFR represents an effort by language educators and testing specialists in Europe to develop a common framework to help relate language courses and assessments to each other (Saville, 2012). The underlying principles of the CEFR include viewing language learning as a lifelong experience and helping learners to receive recognition of their qualifications or previous coursework. By reviewing both, researchers and educators will notice similarities and differences, which parallel the similarities and differences in the purposes, scope, and assessment needs within foreign and second language learning in the United States and Europe.

Undoubtedly, each body of standards benefits from the other’s existence (and from their respective efforts in expansion, refinement, and forward-moving propulsion): Any healthy and friendly competitions for use and recognition on a global scale challenge each to be the best that they can be. Both, and even more bodies of language learning standards, are certainly needed because standards must to a certain extent be customized for and calibrated to the local (i.e., national- or regional-level) language learning environment. From a practical perspective, standards may be used to regulate learning, and people and educators most certainly may want any language learning regulation to be at a more national or regional level.

3 | MOVING FORWARD: FUTURE DIRECTIONS AND RESEARCH NEEDS

Undoubtedly, additional research into the ACTFL guidelines (ACTFL, 2012), the Standards (National Standards Collaborative Board, 2015), and the way in which they work together is needed, particularly because these are “living” documents that will continue to be adapted to reflect applied linguists’ growing understanding of language acquisition and the teaching practices that most effectively support learners in their quest for the next proficiency level. As ACTFL moves on to its next 50 years of service to language teachers and learners, we highlight a few key issues and questions.

3.1 | Are all 5 Cs equally emphasized?

In his review of the Standards, Troyan (2012) noted that while the Communication goal in language education is frequently addressed and assessed, the other Cs (Cultures, Connections, Comparisons, and Communities) are addressed unevenly and to a lesser degree. Other researchers, and even ACTFL itself through its 2011 Standards Impact Survey (ACTFL, 2011), have also summarized that even experienced foreign language teachers use assessments that are more grammar-based and less standards-based. Even when the assessments are standards-based, the assessments tend to focus on the Communication standard (Kaplan, 2016; Kissau & Adams, 2016). Kaplan and Kissau and Adams, in their two respective studies, found that the Communities and Connections goal areas were the least-frequently addressed, which is somewhat surprising given that in many ways technology has made world language learning more relevant, authentic, and compelling; students can connect with learners across the globe, view authentic materials with ease, investigate areas of personal and professional interest, and easily share their opinions of current events in the target culture(s) and around the globe. But as Golonka, Bowles, Frank, Richardson, and Freynik (2014) cautioned, even though technology is
ubiquitous in foreign and second language education and helps vastly with (computer-mediated) communication and access to authentic materials, it is not a golden pathway to improved learning: Lacking in the computer-assisted language learning (CALL) field is clear and direct research on how technology improves “foreign language processes or outcomes” (p. 92). As Golonka et al. (2014) noted, even when technology is used to enhance learning, pedagogical goals still must come first. New CALL research is needed on how technology improves modern language learners’ access to and understanding of all of the 5 Cs, but especially Cultures, Connections, Comparisons, and Communities.

The apparently unequal emphasis on the 5 Cs highlights additional key areas for future work. For example, if teachers are not equally familiar with all 5 Cs, as researchers have reported (ACTFL, 2011; Kaplan, 2016; Kissau & Adams, 2016), do the teachers concomitantly not know how to approach instruction and assessment in some of these domains, or perhaps do they value the goal areas differently and thus deliberately choose to focus on some areas and essentially ignore others? Or is it that the 5 Cs are actually not equally important? Is the stress on equality across the 5 Cs stemming from a utopian Standards design aesthetic and not from an observation-tested model of language learning? If equality across the 5 Cs is not observed in language curricula, does that mean that the curricula must change (as mostly assumed by prior researchers/theorists), or that the proposed model of how languages are to be learned (through an equal weighting of the 5 Cs) must change? In addition, reiterating suggestions by Glisan (2012) and White (2016), research is needed on the way in which service learning, study abroad, and computer-mediated communication, among other approaches, support the five goal areas and the extent to which teachers across levels and languages help learners take advantage of these emerging learning opportunities. Finally, more studies like Kaplan’s (2016) that investigate assessment instruments and how they align with the Standards, particularly computer-scored assessments, are needed as communication norms and technological trends evolve. These lines of inquiry will help future educators better understand the ways in which the 5 Cs inform language instruction and assessment.

3.2 Are the communicative modes and language skills equally emphasized in assessments?

For decades, applied linguists within the fields of language learning and instructed second language acquisition viewed the four language skills (listening, speaking, reading, and writing) as mainly independent from each other. As a result, not all skills have been researched equally, and listening assessment in particular has been the most neglected, even in the context of task-based language teaching, whole-language movements, and integrated assessments.

Currently, for example, it is rare that preservice teachers’ ability to “interpret oral, printed, and video texts by demonstrating both literal and figurative or symbolic comprehension” (ACTFL and CAEP, 2015, p. 3) is assessed. While interpreting oral language is part of the Oral Proficiency Interview (OPI) and interpreting printed text is assessed through programmatic reading and writing exams, interpreting video texts is not assessed in most teacher preparation programs. This may be due in part to the ongoing theoretical debate about what the construct of listening entails (Wagner, 2013). While the ACTFL/CAEP Standards clearly list the use of video media as a critical component, the ACTFL Proficiency Guidelines (2012) and the assessments based on them are still mostly construed with audio-only listening in mind, likely because of the historical grounding of listening assessments on military and foreign service measurement needs, where audio-only listening is to this day an extremely important construct and highly needed skill.

However, the extent to which teacher candidates, and all learners for that matter, interpret video-based news broadcasts, target language materials that are accessed via technology and/or social media,
and entertainment media (movies, documentary, television, and on-demand video) is unclear. Since students’ classroom language learning experience relies on teacher candidates’ and in-service teachers’ ability to provide access to authentic listening-viewing texts—and evidence suggests that teachers do value and provide such experiences (see Colville-Hall & O’Connor, 2006; Lozano, Padilla, Sung, & Silva, 2004)—research into the most effective ways of helping learners to understand such texts and the impact of these instructional approaches on student motivation and learning is needed (see Acheson & Nelson, 2015; White, 2016).

The role of listening/viewing should also be investigated in the context of integrated learning experiences and assessments, e.g., experiences and assessments in which students read articles, listen to and view authentic texts, report back both orally and in writing on what they learned, discuss their reading and writing in groups, and work toward a production goal that involves multiple literacies and cultural understandings (Van den Branden, 2005; ACTFL IPAs, College Board Advanced Placement exams, and the Educational Testing Service’s TOEFL iBT). Moving forward, as teachers increasingly emphasize the honing of integrated and interwoven skills and the way in which language and culture interact (Kramsch, 2014), research will, in the words of Kramsch, need “to focus less on predetermined, stable, predictable facts of a linguistic, functional, or cultural nature, and more on such fluid discourse processes as comparison, contrast, analysis, interpretation, inferencing, and de- and recontextualization” (p. 308).

3.3 | What is the future role of self-assessment?

The ACTFL Proficiency Guidelines (ACTFL, 2012) have served as a U.S. framework for much of our thinking about proficiency-based instruction and assessment in the United States. It is thus not difficult to imagine their power as a point of reference for self-assessment in the United States as well. Self-assessment requires learners to reflect on and evaluate their work, knowledge, and skills. What is more, it allows learners to identify their strengths and weaknesses and to set short- and long-term goals. While institutions have used self-assessments for a variety of reasons, including providing a resource for students to improve their learning (Ziegler & Moeller, 2012), making placement decisions (Strong-Krause, 2000), and measuring language (Brown, Dewey, & Cox, 2014), the revised NCSSFL-ACTFL Can-Do Statements (ACTFL, 2015) provide an accessible tool that translates the proficiency orientation as outlined in the guidelines into specific, user-friendly terms, tasks, and situations, thus facilitating the self-assessment process. However, research has shown that self-assessments like the Can-Do statements must be used judiciously: Cox (2017) found that correlations between self-assessment and actual ability tend to be moderate at best, with learners tending to overestimate what they can do. Furthermore, some Can-Do statements work better than others (Tigchelaar, Bowles, Winke, & Gass, 2017) and may not discriminate well between the sublevels (Brown et al., 2014; Cox, 2017). Thus educators must select statements for assessment purposes carefully and understand that self-assessments are not intended to replace classroom and large-scale assessments.

These issues highlight key areas for future work. For example, to what extent do learners use Can-Do statements to assess their own learning and actually make and act on specific language learning goals? What successes do those learners experience? To what extent does the use of self-assessment promote learner agency and self-efficacy or help learners understand the domains in which future action will result in the greatest learning gains? As a result of taking responsibility for their own learning, do learners increasingly seek learning opportunities beyond the classroom? To what extent are they better prepared for, and motivated to continue into, advanced language learning experiences?
3.4 | How does standards-based education impact the learner’s overall experience?

Menken et al. (2014, p. 590) explained that standards-based education almost always results in (1) better articulation across grade levels and (2) a higher focus on assessment, because when standards are written they are generally applied to the curriculum, which affects instruction, which in turn affects assessment, and results, concretely, in standards-based tests. However, test designers still need a robust plan to make a solid validity argument: that is, that the scores from the standardized tests are interpreted correctly and applied appropriately. Applied linguists and educators must continually question, as researchers have (e.g., Bachman, 1988; Clark & Clifford, 1988; Loomis, 2015; Malone, 2003; Thompson, Cox, & Knapp, 2016; Weir, 2005) the following: Are the ways in which ACTFL standardized test scores are used appropriate? Are score interpretations meaningful? Do the ACTFL Proficiency Guidelines (2012) and Standards (National Standards Collaborative Board, 2015) push the field toward maximally improved instruction and learning?

3.5 | How do standards inform and support language teacher education?

A body of current research has investigated the extent to which the standards and the proficiency movement have influenced teacher education (Colville-Hall & O’Connor, 2006; Wilbur, 2007). First, many states have mandated that preservice teachers obtain an official ACTFL computerized OPI or OPI test score of Advanced Low or Intermediate High (ACTFL and CAEP, 2015) as required in their target language for K–12 certification and/or that they pass the portfolio/task-based Education Teaching Performance Assessment (edTPA; see Behney, 2016, for an overview). However, research has also shown that future teachers find it difficult to pass the language proficiency portion of these teacher certification tests, in part because their language classes at the university level tend to not focus on the skills that are needed, including the ability to produce sustained and lengthy discourse in both speech and writing (Russell & Davidson Devall, 2016). For example, Glisan, Swender, and Surface (2013, p. 276) found that 45% of 1,957 preservice teachers from 2006 to 2012 did not meet the minimum proficiency level for teacher certification. In light of the dramatic and ongoing shortage of language teachers across the nation (Russell & Davidson Devall, 2016), researchers have proposed changes to teaching curricula that address the proficiency needs of preservice teachers (Behney, 2016; e.g., Kissau, 2014; Troyan & Kaplan, 2015).

In addition, many states require that teacher candidates demonstrate their ability to plan, deliver, and assess standards-based instruction either through the edTPA portfolio or through the ACTFL-CAEP accreditation process. To better understand the impact of the standards on teacher certification, studies have investigated the types of content knowledge that are most important for teachers and questioned whether those bodies of knowledge are represented well on the teacher licensure tests (Kissau & Algozzine, 2017). Additional studies (Russell & Davidson Devall, 2016) have questioned the extent to which such standards-based assessments effectively document the knowledge and skills of teacher candidates who are native and heritage speakers. However, because most of the data have come from case studies or qualitative focal studies with fewer than 30 preservice teachers involved, much work is left to be done. More robust, cross-state, and nationwide longitudinal studies are needed, along with research syntheses or meta-analyses so that the field of foreign language education can better understand the impact of assessments that prioritize teacher proficiency and standards-based instruction on the entire field of foreign language education in the United States.
3.6 Are the guidelines and Standards appropriate for diverse learner populations?

Educators and program monitors have very effectively used standards-based assessments to measure classroom-based foreign and second language learning for majority first language English learners. However, the extent to which such assessments are appropriate for bilinguals or emergent bilinguals (Flores & Schissel, 2014) is theoretically problematic. First, the test scores from bilingual or emergent bilingual students are mostly not differentiated from those of their monolingual peers (Menken et al., 2014). This renders an interpretation of those scores more difficult because when students are bilingual or emergent bilinguals, as most heritage and English language learners in the United States are, their language achievement is often not directly accounted for by the level or type of instruction received. Second, individual differences such as the level of language development due to experiences in the home or local community, language dominance, prior educational history, age, native-language distance (and native-language cultural distance) from the variety of language taught, and variety of the (heritage) language spoken in home are factors that impact both the amount and rate of language achieved. Finally, for all young language learners, other factors, such as socioeconomic status and the parents’ language skills and backgrounds as well as their general education (Mori & Calder, 2017), may account for differences in educational attainment above and beyond differences in the language programs. Growth in language development for these learners can vary tremendously on the individual level. Research is needed to understand how general language learning standards, which were drafted based on observations of mostly monolingual, below-average, average, and above-average performance levels in foreign language learning, align to varieties of populations where there is no average because there is no one homogenous starting point.

4 CONCLUSIONS

In considering how the ACTFL Proficiency Guidelines (2012), the proficiency movement, and the Standards (National Standards Collaborative Board, 2015) have influenced language learning in the past and contemplating how they will do so in the future, one point becomes salient: This philosophical and practical framework provides much-needed focus for and accountability in language teaching and learning at a national level. Language learners’ assessment outcomes can help identify programs’ strengths and weaknesses and illuminate the extent to which learners are prepared to meet their personal and professional goals. Developed from U.S. government initiatives in the early 1950s, the concept of the common yardstick regardless of one’s learning circumstances continues to provide a powerful approach for conceptualizing teaching and measuring learners’ success. What is more, this shared framework, which has international influence and is undoubtedly also influenced by international trends, allows open and frank discussion among its users. Shared definitions allow researchers, program administrators, teachers, students, and other stakeholders to state and to hold each other accountable for learning outcomes and reasonable expectations on how language will be acquired. When the framework is commonly accepted and widely shared, feedback, both elicited and freely volunteered, will inevitably result in revisions and improvements not only to learners’ experiences and eventual proficiency outcomes but also to the framework itself. Moving forward, the field can benefit from studies of the deliberate and thoughtful implementation of the framework not only to enhance language teaching and learning but also to investigate content area learning. One thing is certain: As the field has experienced many changes over the last 50 years, our thinking about the desired outcomes of language teaching and learning and our approaches to designing curriculum, instruction, and assessment experiences will continue to evolve.
REFERENCES


---

**How to cite this article:** Cox TL, Malone ME, Winke P. Future directions in assessment: Influences of standards and implications for language learning. Foreign Language Annals. 2018;51:104–115. [https://doi.org/10.1111/flan.12326](https://doi.org/10.1111/flan.12326)