ACTFL-Aligned Cut Scores
Standard Setting Methodology and Results

ESPAÑOL

LAS LINKS ESPAÑOL

Grades 6–12

Version 1.0

TABLE OF CONTENTS

SUMMARY

On October 24–28, 2022, Data Recognition Corporation (DRC) sponsored a standard setting study for LAS Links® Español. The purpose of the study was to develop cut scores (passing scores) for LAS Links Español that align to the American Council on the Teaching of Foreign Languages (ACTFL®) Proficiency Guidelines. Specifically, the study sought to develop cut scores that aligned to the proficiency levels Novice Low (Principiante Bajo) through Advanced High (Avanzado Alto).

The standard setting study engaged an international committee of six language educators in the Yes/No Angoff procedure (Impara & Plake, 1997) to recommend ACTFL-aligned cut scores for the test.

During the online, five-day workshop, educators (a) discussed the expectations for students in each ACTFL proficiency level and (b) recommended cut scores. Their domain-specific cut score recommendations were combined into a single set of Overall cut scores, which may be applied to two grade spans: grade span 9–12 and grade span 6–8. The Overall scale score ranges associated with each ACTFL proficiency level are shown in Table 1.

Table 1. ACTFL-aligned scale ranges for the LAS Links Español Overall composite score

Test Score	Scale Ranges							
	Novice Low & Mid	Novice High	Inter. Low	Inter. Mid	Inter. High	Adv. Low	Adv. Mid	Adv. High
Overall	0–417	418–450	451–475	476–526	527–568	569–612	613–670	671–999

This document describes the standard setting process and results. Guidance for how to properly interpret the results of ACTFL-aligned cut scores for LAS Links Español is found at the conclusion of the document.

LAS Links Español is a research-based assessment of Spanish-language proficiency. Two parallel forms of the test have been published: Form Español A and Form Español B. These forms are available in five different grade spans: K–1, 2–3, 4–5, 6–8, and 9–12.

BACKGROUND

In the United States, Spanish is the most studied world language: nearly 70% of K–12 student enrollments in world language programs are for those in Spanish (American Councils for International Education, 2017). The successful acquisition of an additional language calls for careful monitoring as students progress in their learning. One way schools monitor students' Spanish-language proficiency is by administering LAS Links Español, the leading test of Spanish-language proficiency currently available.

Although LAS Links and LAS Links Español have their own proficiency standards with associated cut scores and proficiency level definitions, many educators and administrators seek to describe students' language proficiency using the ACTFL Proficiency Guidelines. Therefore, DRC implemented the standard setting described in this white paper to provide LAS Links Español users with cut scores that are linked to the ACTFL Proficiency Guidelines.

ABOUT LAS LINKS ESPAÑOL

LAS Links Español is a research-based assessment of Spanish-language proficiency. Two parallel forms of the test have been published: Form Español A and Form Español B. These forms are available in five different grade spans: K–1, 2–3, 4–5, 6–8, and 9–12. The test items for each grade span are designed to measure developmentally appropriate language skills with special emphasis on academic language, the language needed to access instruction in the K–12 classroom. To this end, LAS Links Español measures students' language proficiency in four key domains: listening (escuchando), speaking (hablando), reading (lectura), and writing (escrita). Students' scores from these four domains are averaged to create a single composite score called Overall.

DRC, along with former publisher CTB/McGraw-Hill, has published language assessments for K–12 students for decades, starting with the Language Assessment Scales (LAS®) and then LAS Links® (CTB/McGraw-Hill, 2013). The Second Edition of LAS Links was published in 2013 in both English and Spanish. LAS Links Español has the same structure and measurement philosophy as the English-language version, LAS Links. However, LAS Links Español is not a simple translation of LAS Links. Although the two testing programs share many commonalities, one may view LAS Links Español as a transadaptation of LAS Links: the test items and the language skills measured by LAS Links Español specifically reflect the types of Spanish-language skills that students acquire in both social and academic settings.

ABOUT THE ACTFL PROFICIENCY GUIDELINES

The ACTFL Proficiency Guidelines have been informing language learning in the United States since 1986 (ACTFL, 2012). The Guidelines describe an individual's language abilities in terms of speaking, writing, listening, and reading, and identify five major levels of proficiency: Novice, Intermediate, Advanced, Superior, and Distinguished. Within the levels Novice, Intermediate, and Advanced, there are sublevels of Low, Mid, and High. The ACTFL Proficiency Guidelines span the full continuum of Spanish language proficiency, and present a valuable instrument for the evaluation of functional language ability.

At the standard setting, DRC sought to establish cut scores that: (a) reflected the expectations summarized in the ACTFL Proficiency Guidelines, (b) linked students' scores on the tests to these expectations, and (c) were appropriate for each domain and grade span.

A total of eight cut scores were established to define nine performance levels: Novice Low, Novice Mid, Novice High, Intermediate Low, Intermediate Mid, Intermediate High, Advanced Low, Advanced Mid, and Advanced High.

THE STANDARD SETTING

YES/NO ANGOFF METHOD

The Yes/No Angoff method was selected for the LAS Links Español ACTFL standard setting. The Yes/No Angoff method is a modification of the modified Angoff (1971) procedure, one of the most frequently implemented methods to establish performance standards on educational assessments. In the Yes/No method, panelists are directed to make a dichotomous ("yes" or "no") judgment about whether the hypothetical threshold examinees would be able to answer each question correctly. The Yes/No Angoff method is well suited to assessments like LAS Links Español.

The materials used at the standard setting workshop were based on empirical test data from LAS Links Español, as well as English- and Spanish-language versions of the ACTFL Proficiency Guidelines.

ACTFL PROFICIENCY THRESHOLD STUDENTS

The ACTFL Proficiency Guidelines have three major levels, Novice, Intermediate, and Advanced, that are subdivided into Low, Mid, and High (e.g., Novice Low, Novice Mid, and Novice High). Each major level describes a specific range of abilities in terms of speaking, writing, listening, and reading.

At the standard setting workshop, participants thought deeply about the skills and abilities of the eight threshold students (Novice Mid, Novice High, Intermediate Low, Intermediate Mid, Intermediate High, Advanced Low, Advanced Mid, and Advanced High). A threshold student is a student who is just entering a proficiency level.

MATERIALS

Participants were provided support materials to successfully engage in the Yes/No Angoff process. Materials included: student test forms of 17–30 items, item maps summarizing information about the items in a test form (item number and answer key), rubrics and associated scoring guides, the ACTFL Proficiency Guidelines, and a collaborative space to draft threshold descriptors.

As participants studied these items, they considered the Spanish-language skills that students needed to answer an item correctly. The test also contains constructed-response items worth multiple points. As participants studied these items, they used the items and associated scoring guides (rubrics) to consider the Spanish-language skills that students needed to earn the first score point, and then they considered the additional knowledge and skills needed to earn two points, and so on.

STAFF

Staff members from DRC served as facilitators and in support roles throughout the standard setting. These staff members did not contribute to the cut score recommendations during the workshop. They were responsible for facilitating the workshop, training participants, entering participant results into a database, performing data analyses, and tracking secure materials. Experts in language testing from DRC were present throughout the workshop and worked with participants to provide specialist support.

WORKSHOP PARTICIPANTS

The committee comprised a purposeful mix of educators with a variety of backgrounds. Special care was taken to promote diversity among participants in terms of background and location.

Before the workshop, participants were asked to describe their background and demographics. Based on these self-reported characteristics, all six of the participants were women; five were white and one preferred not to indicate her race; and two were of Hispanic, Latino/a, or Spanish origin. All six were proficient in both Spanish and English.

Of the six participants, four were currently employed as educational consultants, one as a Spanish-language teacher, and one as a school-level administrator. All six participants held a bachelor's degree or higher, and five held a master's or professional degree. Four of the participants had worked in education for more than 20 years, and all six worked in education for more than 10 years.

WORKSHOP PROCESS

The standard setting workshop took place over a five-day period on Zoom. All participants began the workshop with an opening session and training led by DRC. Participants understood that they would consider the Spanish-language skills expected of students in each proficiency level, and they would engage in the Yes/No Angoff method to make cut score judgments.

The participants read the ACTFL Proficiency Guidelines to consider the Spanish-language skills that students were expected to demonstrate at the threshold of each proficiency level and engaged in discussions about the Spanish-language skills they expected to be demonstrated by each of the eight threshold students. The eight threshold students were just in each ACTFL proficiency level, from just Novice Mid to just Advanced High. Participants recorded their expectations for students at the thresholds of each ACTFL proficiency level on electronic whiteboards. Participants were encouraged to review these descriptions frequently throughout the workshop and to consider the threshold students when they made their Yes/No ratings. By the end of this discussion, participants had thoroughly considered the ACTFL Proficiency Guidelines and threshold students, and they reached an understanding of the types of skills that the threshold student for each proficiency level should have.

Participants then examined the tests in terms of what each item measured. Participants were instructed to take notes on the item maps about the Spanish-language skills required to answer the items correctly.

On the morning of day 2, DRC gave additional training for Yes/No Angoff ratings. Participants were reminded how to make Yes/No Angoff ratings, and how these ratings could be transformed into cut score recommendations. This was followed by a mid-process evaluation and all six participants indicating that they were ready to proceed with Round 1.

Individually, participants made their Round 1 Yes/No Angoff ratings. They referred to their test books, item maps, threshold student descriptions, and the ACTFL Proficiency Guidelines. Following Round 1, DRC

calculated the Yes/No Angoff cut score recommendations and presented the results to participants. Specifically, participants were shown their calculated cut score recommendations, as well as the overall median cut score recommendation for the group. Participants were also shown a histogram of the range of the group's Round 1 cut score recommendations.

The Round 1 discussion focused on items where participants disagreed in their Yes/No ratings. Participants referred to their test books, item maps, threshold student descriptions, and the ACTFL Proficiency Guidelines throughout the discussions. To further inform participants' work, DRC shared a crosswalk drafted by DRC Test Development that illustrated in broad strokes where the ACTFL proficiency levels and LAS Links Español grade-span proficiency level descriptors may intersect. This illustration provided no instructional information and was only meant to further participants conversations.

Throughout the standard setting process, participants completed multiple evaluations to assess their understanding of the workshop goals, Yes/No Angoff method, readiness to proceed, and their confidence in their recommendations. Generally, participants were satisfied with their recommendations and with the workshop as a whole.

Following this discussion, participants made their Round 2 Yes/No Angoff ratings. Participants were reminded that Yes/No Angoff rating is an individual activity. Participants were also reminded that they would be free to retain their Yes/No Angoff ratings for any/all items from Round 1 or to change one or more of them; however, in either case, participants would need to have content-based rationales for their decisions.

DRC calculated the Yes/No Angoff cut score recommendations for Round 2, and participants were presented with their calculated cut score recommendation. The presentation included the overall median cut score recommendation for the group and histogram representation of the range of their cut score recommendations. After participants completed the process for reading, they repeated the process for listening, writing, and speaking.

After making their cut score recommendations, participants were presented with the cut score recommendations for all four domains. Participants were informed that they could recommend adjustments to the cut scores, if needed, had their conceptualizations of the threshold students changed over the course of the workshop. However, the committee indicated it was generally satisfied with its recommendations.

Participants indicated their perceived validity of the workshop and their recommendations as part of the post-workshop evaluation. Hambleton (2001) noted that evaluations are important evidence for establishing the validity of performance levels. Throughout the standard setting process, participants completed multiple evaluations to assess their understanding of the workshop goals, Yes/No Angoff method, readiness to proceed, and their confidence in their recommendations. Generally, participants were satisfied with their recommendations and with the workshop as a whole. Particularly, participants understood the connection between their conceptions of threshold students and their corresponding cut score recommendations. Participants generally agreed that the final recommendations reflected the work of the standard setting committee.

RESULTS

PARTICIPANTS' RECOMMENDATIONS AFTER ROUND 2

Table 2 shows participants' Overall recommendation from Round 2 of the Yes/No Angoff procedure, as expressed on the test scale. The Overall cut score is the average of the cut scores for the four domains. The Overall cut scores are rounded down, when applicable, benefitting students very near the cut score.

During the workshop, participants made Yes/No Angoff recommendations on the raw-score metric for each domain. To calculate the group's Overall recommendation, each participant's raw-score cut score recommendation was transformed onto the test scale, and then the median of those cut scores was taken.

Table 2. Recommendations from Round 2 of the standard setting, scale-score metric

Test Score	Scale-Score Cut-Score Recommendations							
	Novice Mid	Novice High	Inter. Low	Inter. Mid	Inter. High	Adv. Low	Adv. Mid	Adv. High
Overall	362	418	446	465	527	569	613	672

After the standard setting, DRC reviewed participants' cut score recommendations, and noted that (a) the standard setting had been conducted according to industry best practices, and (b) that the cut scores were recommended in a defensible manner. They also noted that DRC has the latitude to adjust cut score recommendations within a band defined by conditional standard error of measurement (CSEM) and that such adjustments for technical reasons are very common.

The adjustments made by DRC were all within a range of ±1 CSEM of the cut scores recommended by participating educators. Adjustments made within a range of ±2 CSEM are typically seen as being consistent with standard setting participants' content-based expectations, given the expected variability surrounding the recommendation process.

Table 3 shows the adjusted cut scores, as developed after the standard setting by DRC, expressed on the scale-score metric.

Table 3. Adjusted cut scores, scale-score metric

Test Score	Scale-Score Cut-Scores							
	Novice Mid	Novice High	Inter. Low	Inter. Mid	Inter. High	Adv. Low	Adv. Mid	Adv. High
Overall	362	418	451	476	527	569	613	671

In response to participants' feedback, DRC collapsed the lowest two levels into a single level, Novice Low & Mid. This decision is described in the next section. Table 4 shows the scale ranges, as based on the adjusted cut scores from Table 3, developed after the standard setting by DRC. These scale ranges are expressed on the Overall composite metric.

Table 4. ACTFL-aligned scale ranges for LAS Links Español for the Overall composite score

Test Score	Scale Ranges							
	Novice Low & Mid	Novice High	Inter. Low	Inter. Mid	Inter. High	Adv. Low	Adv. Mid	Adv. High
Overall	0–417	418–450	451–475	476–526	527–568	569–612	613–670	671–999

SUGGESTED USES OF THE ACTFL-ALIGNED CUT SCORES

LAS LINKS ESPAÑOL AND ACTFL

LAS Links Español provides important information for determining Spanish language proficiency and subsequently for planning instructional programs and making decisions about language dominance, identification, classification, eligibility, and placement. The test results can also be used to identify students who would benefit from taking assessments in the native language, to identify difficulties in the native language, and to monitor and track progress in attaining Spanish language proficiency.

The ACTFL Proficiency Guidelines have been informing language learning in the United States since 1986 (ACTFL, 2012). The Guidelines describe an individual's language abilities in terms of speaking, writing, listening, and reading. The ACTFL Proficiency Guidelines span the proficiency continuum of little to no ability to a fluent individual, and they present a valuable instrument for the evaluation of functional language ability.

At the standard setting, DRC sought to establish cut scores that: (a) reflected the expectations summarized in the ACTFL Proficiency Guidelines, (b) linked students' scores on the tests to these expectations, and (c) were appropriate for each domain and grade span.

CONSIDERATIONS FOR NOVICE LOW

After the standard setting, DRC determined that LAS Links Español does not provide an interpretable distinction between students in the Novice Low and Novice Mid levels of ACTFL. As part of the standard setting, participants noted that the skills expected of students in these ACTFL levels are only very slightly different. Moreover, DRC noted that LAS Links Español measurement error was higher among items at this point on the test scale. For these reasons, an ACTFL-aligned Novice Low cut score should not be applied to LAS Links Español results. Instead, one can interpret test performances below the Novice High cut score (i.e., below 418) as being consistent with the Novice Low & Novice Mid levels.

APPLICATION OF CUT SCORES TO GRADE SPAN 6–8

To evaluate the possibility of applying the ACTFL-aligned scale ranges to grades 6–8, the committee examined the hardest items from grade span 6–8, and they found that the structure and overall difficulty of these items was comparable with grade span 9–12, albeit with somewhat less complex stimuli used in the grades 6–8 assessment. Study participants verified the appropriateness of applying most of the ACTFL-aligned scale score ranges from the grades 9–12 recommendations to grade span 6–8 in their responses to a post-workshop survey where this question was formally posed again. Participants and DRC staff members agreed, however, that Advanced High should not be reported for grade span 6–8. That is, students who take the grade span 6–8 test can use the same cut scores as their peers in high school, but the highest proficiency level that they can achieve is Advanced Mid. For example, if a student scores 671 in grade 8, they should be considered "Advanced Mid or higher." LAS Links Español tests below grade span 6–8. LAS Links users who need ACTFL levels for highly skilled students in grades 4 or 5 should consider administration of the tests for grade span 6–8.

References

American Council on the Teaching of Foreign Languages. (2012). ACTFL Proficiency Guidelines 2012. https://www.actfl.org/educator-resources/actfl-proficiency-guidelines

American Councils for International Education. (2017). The national K–12 foreign language enrollment survey report. https://www.americancouncils.org/sites/default/files/FLE-report-June17.pdf

Angoff, W. H. (1971). Scales, norms and equivalent scores. In R.L. Thorndike, (Ed.), *Educational Measurement, 2nd Ed*. Washington, DC: American Council on Education, 508-600.

CTB/McGraw-Hill. (2013). LAS Links second edition technical manual: Español, form B. Maple Grove, MN: Data Recognition Corporation.

Hambleton, R. K. (2001). Setting performance standards on educational assessments and criteria for evaluating the process. In G. J. Cizek (Ed.), *Setting performance standards: Concepts, methods, and perspectives* (pp. 89–116). Mahwah, NJ: Lawrence Erlbaum.

Impara, J. C., & Plake, B. S. (1997). Standard setting: An alternative approach. *Journal of Educational Measurement*, 34, 353–366. https://www.jstor.org/stable/1435114

DATA RECOGNITION
DRC
CORPORATION