



MELAB 2016

Report

Contact Information

All correspondence and mailings should be addressed to:

CaMLA

Argus 1 Building
535 West William St., Suite 310
Ann Arbor, Michigan
48103-4978 USA

T +1 866.696.3522

T +1 734.615.9629

F +1 734.763.0369

info@cambridgemichigan.org

CambridgeMichigan.org



© 2017 Cambridge Michigan Language Assessments®



TABLE OF CONTENTS

1. Description of the Test.....	1
1.1 General Description	1
1.2 Test Format	1
2. Scoring and Reporting of Results	2
2.1 Explanation of Scoring	2
2.2 Procedures for Reporting Scores	2
3. Interpreting MELAB Results	2
4. Test-Taking Population	3
5. Test Statistics	4
5.1 Distribution of Results for the Test Population	4
5.2 Reliability Figures for Listening and GCVR	5
5.3. Rater Agreement Figures for Writing.....	5
5.4. Speaking Rater Performance.....	5
6. References	5

LIST OF TABLES

Table 1: Format and Content of the MELAB.....	1
Table 2: Final MELAB CEFR Score Bands	2
Table 3: List of First Language Backgrounds	3
Table 4: Distribution (in %) of MELAB Test Takers by Age.....	3
Table 5: Distribution (in %) of MELAB Test Takers by Gender.....	3
Table 6: Distribution (in %) of MELAB Test Takers by Purpose for Taking the Test	3
Table 7: Descriptive Statistics for MELAB Scaled Scores.....	4
Table 8: MELAB Scaled Scores Corresponding to Specified Percentiles.....	4
Table 9: Distribution (in %) of MELAB Speaking Test Scores	4
Table 10: Distribution (in %) of MELAB Test Takers by CEFR Level	4

1. DESCRIPTION OF THE TEST

1.1 GENERAL DESCRIPTION

The Michigan English Language Assessment Battery (MELAB) is a standardized, English as a foreign language examination, designed by CaMLA, for adult nonnative speakers of English who will need to use English for academic or professional purposes. The four component skills of listening, reading, writing, and speaking are evaluated through a combination of tasks.

The MELAB is aimed at the B1–C1 levels of the Common European Framework of Reference (CEFR; Council of Europe, 2001), and the score report is valid for two years. MELAB scores are used by students applying to universities in the United States, Canada, and United Kingdom where the language of instruction is English. They are also used for certification of English proficiency for various organizations and licensing professionals.

CaMLA is committed to the excellence of its tests, which are developed in accordance with the highest standards in educational measurement. All parts of the examination are written following specified guidelines, and items are pretested to ensure that they function properly. CaMLA works closely with test centers to ensure that its tests are administered in a way that is fair and accessible to test takers and that the MELAB is open to all people who wish to take the exam, regardless of the school they attend.

1.2 TEST FORMAT

The MELAB measures four skill areas: listening, reading, writing, and speaking. The writing, listening, and reading sections are taken in one sitting. The speaking test is optional and is taken separately. Table 1 describes the format and content of the MELAB. Test preparation resources are available on the CaMLA website.

Table 1: Format and Content of the MELAB

Section	Time	Description	Number of Items
Writing	30 minutes	Test takers write an essay based on one of two topic choices.	1 task
Listening	35–40 minutes	Part 1 (multiple choice) A short recorded question or statement is accompanied by three printed responses. Test takers choose the statement that conveys a reasonable answer or response.	18
		Part 2 (multiple choice) A short recorded conversation is accompanied by three printed statements. Test takers choose the statement that conveys the same meaning as what was heard, or that is true based upon the conversation.	22
		Part 3 (multiple choice) Four recorded interviews, such as those that might be heard on the radio, are each followed by recorded comprehension questions. The questions and answer choices are printed in the test booklet. Test takers choose the correct answer from the choices.	20
Grammar Cloze Vocabulary Reading (GCVR)	80 minutes	Grammar (multiple choice) An incomplete sentence is followed by a choice of four words or phrases to complete it. Only one choice is grammatically correct.	32
		Cloze (multiple choice) Two passages with deletions are followed by choices of words and phrases to complete the text. Test takers must choose the word or phrase that best fills the blank in terms of grammar and meaning.	24
		Vocabulary (multiple choice) An incomplete sentence is followed by a choice of four words or phrases to complete it. Test takers must choose the option that best completes the sentence in terms of meaning.	31
		Reading (multiple choice) Four reading passages are followed by comprehension questions. Test takers choose the correct answer from the printed answer choices.	23
Speaking	15 minutes	Test takers engage in a conversation with an examiner.	

2. SCORING AND REPORTING OF RESULTS

2.1 EXPLANATION OF SCORING

The MELAB listening and grammar, cloze, vocabulary, and reading (GCVR) sections are scored by computer at CaMLA. Each correct answer contributes proportionally within each section, and there are no points deducted for wrong answers. A scaled score is calculated using an advanced mathematical model based on Item Response Theory. This method ensures that scores are comparable across different administrations.

The speaking and writing sections are graded according to scales established by CaMLA (see our website for the rating scales). The speaking section is conducted and assessed by a CaMLA certified speaking examiner, and the writing section is assessed by at least two CaMLA certified writing examiners.

2.2 PROCEDURES FOR REPORTING SCORES

MELAB scores are reported on an official score report form. Official score reports of the MELAB are sent directly from CaMLA to an admissions office. The score report provides the following information:

- A score for each of the required sections (writing, listening, and GCVR)
- A final MELAB score (0–99), which is the average of the writing, listening, and GCVR section scores.
- A speaking test score for test takers who opted for this part of the test.
- Additional comments about the test performance (where this is relevant).

The section scores reported in the score report can also be used to provide information on the test taker's CEFR level: B1–C1. Table 2 shows the MELAB scaled scores that correspond to these levels. These correspondences are based on a standard setting study conducted by CaMLA (CaMLA, 2016).

3. INTERPRETING MELAB RESULTS

The MELAB is a multilevel exam, covering a range of proficiency levels on the CEFR (Council of Europe, 2001) from B1 to C1. Selected CEFR performance descriptors illustrating what test takers should be able to do at each level are available from the CaMLA website.

When interpreting a MELAB score report, it is important to remember that the MELAB estimates a test taker's true proficiency by approximating the kinds of tasks that may be encountered in real life. Also, temporary factors unrelated to a test taker's proficiency, such as fatigue, anxiety, or illness, may affect exam results. Additionally, check the date the test was taken. While the MELAB score report is valid for two years, language ability changes over time. This ability can improve with active use and further study of language, or it may diminish if the report holder does not continue to study or to use English on a regular basis.

When using test scores for decision-making, both section scores and the final score should be considered. Two test takers who have the same final scores but quite different section scores may differ in their language skills. Such differences may affect their ability to use English effectively in different contexts. It is also important to remember that test performance is only one aspect to be considered. Communicative language ability consists of both knowledge of language and knowledge of the world. Therefore, one would need to consider how factors other than language affect how well someone can communicate. For example, in the context of academic studies in English, the ability to function effectively involves not only knowledge of English, but also other knowledge and skills such as content knowledge and study skills. Since language ability is just one of many factors that affect success, or lack of success, in an academic setting, MELAB scores should not be used to predict academic success or failure.

Table 2: Final MELAB CEFR Score Bands

CEFR Level	Listening	GCVR	Writing	Speaking*
C1	92 – 100	91 – 100	90 – 97	4- – 4
B2	81 – 91	78 – 90	77 – 87	3- – 3+
B1	62 – 80	57 – 77	67 – 75	2- – 2+
Below B1	0 – 61	0 – 56	0 – 65	1 – 1+

Note that the speaking scale includes “plus” and “minus” bands between 1 and 4. The full scale is 1, 1+, 2-, 2, 2+, 3-, 3, 3+, 4-, 4

4. TEST-TAKING POPULATION

This section presents an overview of the test takers who took the MELAB in 2016, providing demographic information for the testing population. Every test taker completed a registration form, which asked for their gender, date of birth, first language, and purpose for taking the exam. Cases where information was not given or was not correctly given were treated as missing data.

Table 3 lists the first language backgrounds of the test takers. The test takers represented 45 different first language backgrounds, but it should be noted that the test population primarily consisted of test takers whose first language was Arabic, Chinese (Cantonese/Mandarin), German, Farsi/Persian, or Spanish.

Table 3: List of First Language Backgrounds

Albanian	Hindi	Shona
Amharic	Hungarian	Sinhalese
Arabic	Ibo (Igbo)	Slovak
Armenian	Indonesian	Slovenian/Slovene
Benga	Japanese	Spanish
Bengali	Korean	Tagalog/Filipino
Bulgarian	Kurdish	Tamil
Chinese	Lithuanian	Telugu
(Cantonese/ Mandarin)	Macedonian	Thai
English	Malayalam	Tibetan
Ewe	Mongolian	Turkish
Farsi/Persian	Pashtu	Ukrainian
French	Portuguese	Urdu
German	Punjabi	Vietnamese
Gujarati	Romanian	Yoruba
	Russian	

Tables 4, 5, and 6 present the distribution of test takers by age, gender, and purpose for taking the test, respectively. Each of these tables presents the test takers who took the MELAB with the speaking test separately from those who did not take the speaking test. Overall, 42.18% of the MELAB test takers took the optional speaking test.

Table 4 shows that a substantial number of MELAB test takers were between the ages of 17 and 22 years old (44.36%), which implies that they were college-aged test takers who were likely still in school, applying for college, or in the very early stages of their careers. However, the table also shows that many of the MELAB test takers who took the speaking test were over 29 years old (28.97%), meaning they were adults beyond college-age. Table 5 shows that the majority of MELAB test takers were male (52.05%). However, for the MELAB test takers who took the speaking

test, it shows that the majority of test takers were female (53.80%). Table 6 shows that the majority of test takers reported taking the test for education program admissions (62.82%).

Table 4: Distribution (in %) of MELAB Test Takers by Age

Age	Without Speaking	With Speaking	Overall
13–16	0.00	0.00	0.00
17–19	0.44	0.61	0.51
20–22	24.17	21.88	23.21
23–25	28.82	10.64	21.15
26–29	14.41	7.90	11.67
30–39	13.08	13.68	13.33
≥ 40	12.64	18.24	15.00
Missing Data	5.10	26.14	13.97

Table 5: Distribution (in %) of MELAB Test Takers by Gender

Gender	Without Speaking	With Speaking	Overall
Male	57.87	44.07	52.05
Female	40.80	53.80	46.28
Missing Data	1.33	2.13	1.67

Table 6: Distribution (in %) of MELAB Test Takers by Purpose for Taking the Test

Reason	Without Speaking	With Speaking	Overall
Education Program Admissions	65.63	58.97	62.82
Language Course Requirement	23.06	8.51	16.92
Scholarship	0.44	0.30	0.38
Obtain Employment	0.44	3.65	1.79
Improve Employment	1.33	4.26	2.56
Personal Interest	2.22	1.22	1.79
Other	2.22	18.84	9.23
Missing Data	4.66	4.26	4.49

Table 7: Descriptive Statistics for MELAB Scaled Scores

Section	Mean	Standard Deviation	Median	Minimum	Maximum
Writing	76.83	7.53	77	53	97
Listening	77.06	11.26	78	0	98
GCVR	74.06	13.62	75	0	99
Final Score	76.83	7.53	77	53	97

5. TEST STATISTICS

5.1 DISTRIBUTION OF RESULTS FOR THE TEST POPULATION

Table 7 presents descriptive statistics for the writing, listening, and GCVR sections, as well as for the final MELAB score. It shows that while test takers scored very similarly across the three test sections, the test takers tended to score slightly lower on the GCVR section.

Table 8 shows the MELAB scores that correspond to specified percentiles. Using this table, it is possible to see how a test taker has performed in relation to the test population. For example, if a test taker has received a final MELAB score of 85, they have performed in the 90th percentile of the population, which means that they have scored higher than 90% of the test takers included in this analysis.

Table 8: MELAB Scaled Scores Corresponding to Specified Percentiles

Percentile	Writing	Listening	GCVR	Final Score
99th	93	97	98	93
95th	90	94	94	90
90th	85	91	91	85
75th	83	86	85	83
50th	77	78	75	77
25th	73	69	65	73
10th	67	62	55	67
5th	65	57	49	65
1st	57	50	44	57

The percentage of test takers at each score point for the optional speaking test is presented in Table 9. It shows that 44.07% of the test takers obtained a score of 4- or 4, suggesting that a large proportion of the test takers who opted for the speaking test were highly proficient English speakers.

Table 9: Distribution (in %) of MELAB Speaking Test Scores

Score	% of Test Takers
1	0.00
1+	0.00
2-	0.61
2	1.82
2+	8.21
3-	6.38
3	15.20
3+	23.71
4-	22.19
4	21.88

Table 10 lists the percentage of test taker scores by CEFR level for each MELAB section. It shows that the majority of test takers were at the B1 level for the listening and GCVR sections, at the B2 level for the writing section, and at the B2 and C1 levels for the speaking section.

Table 10: Distribution (in %) of MELAB Test Takers by CEFR Level

Section	A2	B1	B2	C1
Listening	9.87	49.23	32.44	8.46
GCVR	11.15	46.79	30.77	11.28
Writing	7.56	39.10	47.31	6.03
Speaking	0.00	10.64	45.29	44.07

5.2 RELIABILITY FIGURES FOR LISTENING AND GCVR

Test scores are a numerical measure of a test taker's ability. *Reliability* refers to the consistency of the measurement. In theory, a test taker's test score should be the same each time the test is taken or across different forms of the same test. In practice, even when the test conditions are carefully controlled, an individual's performance on a set of test items will vary from one administration to another due to variation in the items across different forms of the same test or due to variability in individual performance. Among the reasons for this are temporary factors unrelated to a test taker's proficiency, such as fatigue, anxiety, or illness. As a result, test scores always contain a small amount of measurement error. The aim, however, is to keep this error to a minimum. For high-stakes exams such as the MELAB, a reliability figure of 0.80 and above is expected and acceptable. Apart from monitoring reliability, the estimated variability in test-taker performance is also monitored through the standard error of measurement (SEM) estimate.

Reliability and SEM estimates are obtained for each administration of the MELAB. The reliability estimates are calculated in BILOG using the Bayes MAP (maximum a posteriori) method. The SEM estimates are calculated using the reliability estimates and the scaled scores. In this report, the reliability and SEM estimates are summarized as averages across the different 2016 MELAB administrations. For the listening section, the average reliability estimate was 0.83, and the average SEM estimate was 4.33. For the GCVR section, the average reliability estimate was 0.92, and the average SEM estimate was 3.72. These values demonstrate that the reliability figures for both exam sections are not only above the minimally acceptable value of 0.80 but are consistently near 0.90. Additionally, the SEM estimates as a proportion of the 100-point scale are very small. These values suggest excellent consistency of measurement for the MELAB listening and GCVR sections.

5.3. RATER AGREEMENT FIGURES FOR WRITING

The examiners for the writing section are native speakers of English who are trained and certified according to standards set by CaMLA. Each writing performance is rated separately by two accredited raters. If the raters reach exact agreement, then the agreed score is the score awarded. If the raters award adjacent scores, the score awarded is an average of the two rater scores. If the raters do not reach exact or adjacent agreement on the score to be awarded, the writing performance is evaluated separately by a third rater.

CaMLA monitors rater agreement for quality control

purposes. The percentage of exact and adjacent agreement between the first and second raters is monitored, along with the Pearson correlation coefficient. In this report, these two rater agreement figures are summarized across the different 2016 MELAB administrations. The overall percentage of exact and adjacent agreement was 92.95%, and the overall Pearson correlation was 0.847. Both of these values are high, which suggests excellent agreement among raters.

5.4. SPEAKING RATER PERFORMANCE

The examiners for the speaking test are native or highly proficient nonnative speakers of English who are trained and certified according to standards set by CaMLA. Recordings of speaking tests are sent to CaMLA for monitoring and review.

6. REFERENCES

CaMLA (2016). *Linking the Common European Framework of Reference and the Michigan English Language Assessment Battery*. CaMLA Technical Report, CaMLA. Retrieved from <http://cambridgemichigan.org/wp-content/uploads/2017/04/MELAB-CEFR-Linking-Technical-Report.pdf>

Council of Europe (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.