

1. **Virtual Meeting Information**

1.1. The meeting is accessible to persons with disabilities. For questions about accessibility or to request an accommodation, please contact Richelle Krotts, Richelle.Krotts@tspc.oregon.gov **at least 48 hours before the meeting** [ORS 192.630((5)(b))].

2. **Educator Preparation & Pathway Committee**

Dr. Michelle York, Chair
Dr. Carrie Kondor
Samantha Herron
Dr. Alfonso Garcia Arriola
Vacant

3. **Welcome**

4. **New Program Proposals, Major Modifications, and Waivers**

4.1. ***Major Modification Requests***

4.1.1. George Fox University (GFU): Major modification request to the Undergraduate Teacher Education (UGTE) program for preservice candidates to revise coursework and program structure;

4.1.2. George Fox University (GFU): Major modification request to the Undergraduate Accelerated Teaching Degree (ATD) program for preservice candidates to restructure the program as a streamlined bachelor's degree licensure pathway

5. **Committee Business Items**

5.1. ***ORELA/NES Content Knowledge Assessment Revisions***

5.2. ***Scheduling EP&PC Meetings***

6. **Rule Proposals and Revisions: EP&PC Initial Review**

6.1. ***Program Standards Updates to Align with Current Literacy Standards***

7. **Information Items**

The remaining items are included for information only. If EP&PC members wish to discuss any of these matters further, the item(s) can be pulled for discussion at this point in the meeting.

7.1. ***Early Literacy Implementation Update***

7.2. ***Future Meetings***

8. **Adjournment**



National Benchmark Conference Report

Conducted Online October 1, 2025
Mathematics (326)

Prepared 10/24/25 by:
Evaluation Systems group of Pearson
300 Venture Way
Hadley, MA 01035

This document constitutes a working draft that includes some materials that are test secure and/or confidential for other reasons. As such, it should not be circulated to unauthorized persons. The information set forth in this document is subject to change, and any changes will fully supersede the information contained in this document.

Copyright © 2025 Pearson Education, Inc. or its affiliate(s). All Rights Reserved.
Evaluation Systems group of Pearson, 300 Venture Way, Hadley, MA 01035



Table of Contents

OVERVIEW	3
PURPOSE OF THE NATIONAL BENCHMARK CONFERENCE.....	3
POTENTIAL USES FOR THE NATIONAL EVALUATION SERIES™	4
NATIONAL BENCHMARK PROCESS.....	4
ORIENTATION	4
SIMULATED TEST-TAKING ACTIVITY	5
JUST ACCEPTABLY QUALIFIED CANDIDATE.....	5
ROUND ONE: ITEM-BASED JUDGMENTS.....	6
ROUND ONE: ANALYSIS	7
ROUND TWO: ITEM-BASED JUDGMENTS	7
ROUND TWO: ANALYSIS	7
CONFERENCE EVALUATION.....	8
NATIONAL BENCHMARK RESULTS	8
Appendix A.....	9
Appendix B	11

OVERVIEW

This report was prepared for the Educator Preparation and Pathways Committee and Oregon Teacher Standards and Practices Commission for determination related to adoption of redeveloped National Evaluation Series™ tests. This report contains the results from National Benchmark Conferences conducted for the following National Evaluation Series™ tests for teacher certification:

- Mathematics (326)

The following three groups of educators were recruited to participate on the panels:

1. Experienced public-school educators who have worked directly with teacher candidates (student teachers)
2. Early-career public school educators (with 5 or fewer years of experience)
3. Educator Preparation Program faculty members

All panelists were approved by the Director of Assessment Programs or designee prior to being invited to serve on the National Benchmark Panel. For a breakdown of the panel composition, refer to Appendix A.

PURPOSE OF THE NATIONAL BENCHMARK CONFERENCE

The purpose of the National Benchmark Conference is for expert educators and educator preparation faculty from multiple states to participate in a structured process to make individual judgments regarding recommended passing scores for teacher licensure tests. The passing score represents the minimum level of performance that a candidate must achieve to demonstrate the acceptable level of knowledge and skills required to perform the activities in which the certification allows the candidate to engage. This report presents information about the benchmarking process, a summary of the panel members, and the panel-recommended passing scores.

POTENTIAL USES FOR THE NATIONAL EVALUATION SERIES™

Each implementation of an assessment has a specific purpose. While the passing standards resulting from the National Benchmark Conference are intended for use in making certification decisions, the Series tests can be used by both states and institutions of higher education.

Adopting agencies and organizations may choose to use one or more tests for purposes other than state licensure, such as a component of completing a teacher preparation program, a component for acceptance into a preparation program, or gaining clearance to student teach. The passing standard recommended by the benchmark panel is for reference by agencies and organizations considering the role of a test for inclusion in an educator licensure program and determining a passing standard. Pearson works closely with states that opt to conduct their own standard setting rather than adopting the standard recommended by the National Benchmark panel.

NATIONAL BENCHMARK PROCESS

An iterative procedure was used in which panelists provided item-level ratings using procedures commonly referred to as a modified-Angoff method. After an orientation to the item rating methods, the panelists reviewed the test and were guided through two rounds of independent benchmark ratings. Each panel's aggregated ratings provided information about the expected test-level performance of just acceptably qualified candidates for initial licensure.

ORIENTATION

Panelists were provided an orientation that included the background to the National Evaluation Series™ test development, an overview of the test development process, including specific explanation of the benchmark process, the concept of the “just acceptably qualified candidate,” and the judgments about test items that the panelists would be asked to make. Panelists were encouraged to ask questions about the meeting's purpose before moving to the next step.

SIMULATED TEST-TAKING ACTIVITY

Panelists engaged in a simulated test-taking activity to familiarize themselves with the knowledge and skills measured by the test items. Each panelist was provided with a copy of the first operational test form and the test framework and was asked to read and answer the questions on the test without referring to the answer key. Panelists were then provided with the answer key and asked to score their own answers. Panelists were encouraged to write in their copy of the test booklet and to write notes to assist them with their review.

JUST ACCEPTABLY QUALIFIED CANDIDATE

During the orientation, panelists were introduced to the concept of the “just acceptably qualified candidate.” The facilitator guided the panel through extensive discussion of this concept, and panelists submitted a confirmation form acknowledging they understood it fully prior to the Round 1 Ratings. The definition of this term is presented below.

Just Acceptably Qualified Candidate

Just at the level of knowledge and skills important for effective job performance as a beginning teacher in a classroom in public schools

ROUND ONE: ITEM-BASED JUDGMENTS

The Pearson facilitator provided training in this step of the national benchmark process during which panelists made item-based judgments using a modified-Angoff procedure for the multiple-choice questions (MCQs).

To prepare for making their judgments, panelists were informed of the factors potentially influencing candidate performance, such as complexity of the items, content tested, item context, and language. In addition, the panelists completed a training exercise with sample test items with a range of item difficulty.

Panel members were asked to consider the group of “just acceptably qualified candidates” having the baseline pedagogical content knowledge and knowledge of professional responsibilities as in the respective test frameworks. Applying their professional judgments in response to the question below, each panel member provided an independent rating for each multiple-choice item indicating what percent of this group would provide a correct response for each item.

Imagine a hypothetical group of individuals who are just at the level of knowledge and skills required to perform the job of an initially certified educator in this field in U.S. public schools.

What percent of this group would answer the item correctly?

0% - 10%	=	1	51% - 60%	=	6
11% - 20%	=	2	61% - 70%	=	7
21% - 30%	=	3	71% - 80%	=	8
31% - 40%	=	4	81% - 90%	=	9
41% - 50%	=	5	91% - 100%	=	10

ROUND ONE: ANALYSIS

After the panelists completed their item ratings, Pearson analyzed their rating forms to produce Round One Item Rating Summaries to be given to each panelist. The reports were prepared for each test and contained the panelist's rating, the median rating provided by all panelists rating the item, and the distribution of ratings across all panelists.

ROUND TWO: ITEM-BASED JUDGMENTS

For Round Two, panelists were provided with the Item Rating Summary from Round One. Pearson facilitators explained how to read and interpret the summary. Panelists were provided with an opportunity to ask for further clarification of how to read and interpret the analyses.

Each panelist was given an opportunity to review the ratings for each item on the test in their field made by the group of panelists and compare these ratings to his/her own. Panelists then had an opportunity to discuss the items. Based on their review of the analyses provided, panelists had the opportunity to provide a second rating to replace the first rating for any items.

ROUND TWO: ANALYSIS

After the panelists completed their Round 2 ratings, Pearson analyzed their rating forms. A Round 2 Rating Summary was produced and contained the following information: a) number of scorable items on the test form, b) number of raters, c) sum of item medians (i.e., panel-based standard for each component), and d) distribution of individual rater results sorted in descending order by score value. The Round 2 Rating Summary was based on the panel members' final rating on each item (i.e., either the unchanged Round 1 rating or the Round 2 rating if it is different from the Round 1 rating). The Round 2 Rating Summary was used by facilitators to confirm the completion of the benchmarking process. Panelists were not provided with a Round 2 Rating Summary, as the information they provided during Round 2 was final.

CONFERENCE EVALUATION

Following all activities, panel members completed a Conference Evaluation Form (see Appendix B) that allowed each panelist to anonymously rate various aspects of the National Benchmark Conference and record any comments they may have had, including about specific item content.

NATIONAL BENCHMARK RESULTS

Following the conference, Pearson calculated a recommended passing standard, or National Benchmark, for each test area based on second round ratings provided by the panel members. The National Benchmark was then scaled to a score of 220 on a scale of 100–300.

The recommended standard represents the number of multiple-choice questions a candidate must answer correctly to earn a scaled score of 220.

Table 1: National Evaluation Series™ Test Fields National Benchmarks

Field Test	Total # of Items	National Benchmark
Mathematics	80	58

Appendix A

National Benchmark Conference

Panelists

Mathematics National Benchmark Panel October 1, 2025	
Affiliation	
K-12	10
EPP	0
Gender	
Female	6
Male	4
Declined to State	0
Ethnicity	
American Indian/Alaska Native	0
Asian	0
Black/African American	2
Hispanic/Latino	2
Other	0
White	6
Declined to State	0

Note:

EPP: Educator Preparation Program faculty member

Appendix B

Conference Evaluation Form



National Evaluation Series National Benchmark Conference Evaluation Form

Please respond to the following questions by selecting a rating on each scale.

1. Please enter the content area for the conference you are participating in.

Enter your answer

2. Rater ID Number (optional)

Enter your answer

3. How well do you feel you understood the training that you received for making your standard setting judgments?

1	2	3	4	5
---	---	---	---	---

Not At All Well

Very Well

⋮

4. How confident are you of the standard setting judgments you made?

1	2	3	4	5
---	---	---	---	---

Not At All Confident Very Confident

5. How satisfied are you that you were able to complete the work you were asked to do in the time that was provided?

1	2	3	4	5
---	---	---	---	---

Not At All Satisfied Very Satisfied

6. How successful were the coordination and logistics of the Standard Setting Conference?

1	2	3	4	5
---	---	---	---	---

Not At All Successful Very Successful

7. Overall, how satisfied are you with the standard setting review process in which you participated?

1	2	3	4	5
---	---	---	---	---

Not At All Satisfied Very Satisfied

⋮

8. Please provide any additional comments about the standard setting activities.

Enter your answer





National Benchmark Conference Report

Conducted Online November 14, 2024

Music (513)

Prepared 1/7/25 by:

Evaluation Systems group of Pearson
300 Venture Way
Hadley, MA 01035

This document constitutes a working draft that includes some materials that are test secure and/or confidential for other reasons. As such, it should not be circulated to unauthorized persons. The information set forth in this document is subject to change, and any changes will fully supersede the information contained in this document.

Copyright © 2025 Pearson Education, Inc. or its affiliate(s). All Rights Reserved.
Evaluation Systems group of Pearson, 300 Venture Way, Hadley, MA 01035



Table of Contents

OVERVIEW	3
PURPOSE OF THE NATIONAL BENCHMARK CONFERENCE.....	3
POTENTIAL USES FOR THE NATIONAL EVALUATION SERIES™	4
NATIONAL BENCHMARK PROCESS.....	4
ORIENTATION	4
SIMULATED TEST-TAKING ACTIVITY	5
JUST ACCEPTABLE CANDIDATE	5
ROUND ONE: ITEM-BASED JUDGMENTS.....	6
ROUND ONE: ANALYSIS	7
ROUND TWO: ITEM-BASED JUDGMENTS	7
ROUND TWO: ANALYSIS	7
CONFERENCE EVALUATION.....	8
NATIONAL BENCHMARK RESULTS	8
Appendix A.....	9
Appendix B	11

OVERVIEW

This report was prepared for the Educator Preparation and Pathways Committee and Oregon Teacher Standards and Practices Commission for determination related to adoption of redeveloped National Evaluation Series™ tests. This report contains the results from National Benchmark Conferences conducted for the following National Evaluation Series™ test(s) for teacher certification:

- Music (513)

The following three groups of educators were recruited to participate on the panels:

1. Experienced public-school educators who have worked directly with teacher candidates (student teachers)
2. Early-career public school educators (with 5 or fewer years of experience)
3. Educator Preparation Program faculty members

All panelists were approved by the Director of Assessment Programs or designee prior to being invited to serve on the National Benchmark Panel. For a breakdown of the panel composition, refer to Appendix A.

PURPOSE OF THE NATIONAL BENCHMARK CONFERENCE

The purpose of the National Benchmark Conference is for expert educators and educator preparation faculty from multiple states to participate in a structured process to make individual judgments regarding recommended passing scores for teacher licensure tests. The passing score represents the minimum level of performance that a candidate must achieve to demonstrate the acceptable level of knowledge and skills required to perform the activities in which the certification allows the candidate to engage. This report presents information about the benchmarking process, a summary of the panel members, and the panel-recommended passing scores.

POTENTIAL USES FOR THE NATIONAL EVALUATION SERIES™

Each implementation of an assessment has a specific purpose. While the passing standards resulting from the National Benchmark Conference are intended for use in making certification decisions, the Series tests can be used by both states and institutions of higher education.

Adopting agencies and organizations may choose to use one or more tests for purposes other than state licensure, such as a component of completing a teacher preparation program, a component for acceptance into a preparation program, or gaining clearance to student teach. The passing standard recommended by the benchmark panel is for reference by agencies and organizations considering the role of a test for inclusion in an educator licensure program and determining a passing standard. Pearson works closely with states that opt to conduct their own standard setting rather than adopting the standard recommended by the National Benchmark panel.

NATIONAL BENCHMARK PROCESS

An iterative procedure was used in which panelists provided item-level ratings using procedures commonly referred to as a modified-Angoff method. After an orientation to the item rating methods, the panelists reviewed the test and were guided through two rounds of independent benchmark ratings. Each panel's aggregated ratings provided information about the expected test-level performance of just acceptable candidates for initial licensure.

ORIENTATION

Panelists were provided an orientation that included the background to the National Evaluation Series™ test development, an overview of the test development process, including specific explanation of the benchmark process, the concept of the “just acceptable candidate,” and the judgments about test items that the panelists would be asked to make. Panelists were encouraged to ask questions about the meeting's purpose before moving to the next step.

SIMULATED TEST-TAKING ACTIVITY

Panelists engaged in a simulated test-taking activity to familiarize themselves with the knowledge and skills measured by the test items. Each panelist was provided with a copy of the first operational test form and the test framework and was asked to read and answer the questions on the test without referring to the answer key. Panelists were then provided with the answer key and asked to score their own answers. Panelists were encouraged to write in their copy of the test booklet and to write notes to assist them with their review.

JUST ACCEPTABLE CANDIDATE

During the orientation, panelists were introduced to the concept of the “just acceptable candidate.” The facilitator guided the panel through extensive discussion of this concept, and panelists submitted a confirmation form acknowledging they understood it fully prior to the Round 1 Ratings. The definition of this term is presented below.

Just Acceptable Candidate

Just at the level of knowledge and skills important for effective job performance as a beginning teacher in a classroom in public schools

ROUND ONE: ITEM-BASED JUDGMENTS

The Pearson facilitator provided training in this step of the national benchmark process during which panelists made item-based judgments using a modified-Angoff procedure for the multiple-choice questions (MCQs).

To prepare for making their judgments, panelists were informed of the factors potentially influencing candidate performance, such as complexity of the items, content tested, item context, and language. In addition, the panelists completed a training exercise with sample test items with a range of item difficulty.

Panel members were asked to consider the group of “just acceptable candidates” having the baseline pedagogical content knowledge and knowledge of professional responsibilities as in the respective test frameworks. Applying their professional judgments in response to the question below, each panel member provided an independent rating for each multiple-choice item indicating what percent of this group would provide a correct response for each item.

Imagine a hypothetical group of individuals who are just at the level of knowledge and skills required to perform the job of an initially certified educator in this field in U.S. public schools.

What percent of this group would answer the item correctly?

0% - 10%	=	1	51% - 60%	=	6
11% - 20%	=	2	61% - 70%	=	7
21% - 30%	=	3	71% - 80%	=	8
31% - 40%	=	4	81% - 90%	=	9
41% - 50%	=	5	91% - 100%	=	10

ROUND ONE: ANALYSIS

After the panelists completed their item ratings, Pearson analyzed their rating forms to produce Round One Item Rating Summaries to be given to each panelist. The reports were prepared for each test and contained the panelist's rating, the median rating provided by all panelists rating the item, and the distribution of ratings across all panelists.

ROUND TWO: ITEM-BASED JUDGMENTS

For Round Two, panelists were provided with the Item Rating Summary from Round One. Pearson facilitators explained how to read and interpret the summary. Panelists were provided with an opportunity to ask for further clarification of how to read and interpret the analyses.

Each panelist was given an opportunity to review the ratings for each item on the test in their field made by the group of panelists and compare these ratings to his/her own. Panelists then had an opportunity to discuss the items. Based on their review of the analyses provided, panelists had the opportunity to provide a second rating to replace the first rating for any items.

ROUND TWO: ANALYSIS

After the panelists completed their Round 2 ratings, Pearson analyzed their rating forms. A Round 2 Rating Summary was produced and contained the following information: a) number of scorable items on the test form, b) number of raters, c) sum of item medians (i.e., panel-based standard for each component), and d) distribution of individual rater results sorted in descending order by score value. The Round 2 Rating Summary was based on the panel members' final rating on each item (i.e., either the unchanged Round 1 rating or the Round 2 rating if it is different from the Round 1 rating). The Round 2 Rating Summary was used by facilitators to confirm the completion of the benchmarking process. Panelists were not provided with a Round 2 Rating Summary, as the information they provided during Round 2 was final.

CONFERENCE EVALUATION

Following all activities, panel members completed a Conference Evaluation Form (see Appendix B) that allowed each panelist to anonymously rate various aspects of the National Benchmark Conference and record any comments they may have had, including about specific item content.

NATIONAL BENCHMARK RESULTS

Following the conference, Pearson calculated a recommended passing standard, or National Benchmark, for each test area based on second round ratings provided by the panel members. The National Benchmark was then scaled to a score of 220 on a scale of 100–300.

The recommended standard represents the number of multiple-choice questions a candidate must answer correctly to earn a scaled score of 220.

Table 1: National Evaluation Series™ Test Fields National Benchmarks

Field Test	Total # of Items	National Benchmark
Music	80	47

Appendix A

National Benchmark Conference

Panelists

Music	
National Benchmark Panel	
November 14, 2024	
Affiliation	
K-12	9
EPP	0
Gender	
Female	3
Male	2
Declined to State	4
Ethnicity	
American Indian/Alaska Native	0
Asian	0
Black/African American	1
Hispanic/Latino	1
Other	0
White	6
Declined to State	1

Note:

EPP: Educator Preparation Program faculty member

Appendix B

Conference Evaluation Form



National Evaluation Series National Benchmark Conference Evaluation Form

Please respond to the following questions by selecting a rating on each scale.

1. Please enter the content area for the conference you are participating in.

Enter your answer

2. Rater ID Number (optional)

Enter your answer

3. How well do you feel you understood the training that you received for making your standard setting judgments?

1	2	3	4	5
---	---	---	---	---

Not At All Well

Very Well

⋮

4. How confident are you of the standard setting judgments you made?

1	2	3	4	5
---	---	---	---	---

Not At All Confident Very Confident

5. How satisfied are you that you were able to complete the work you were asked to do in the time that was provided?

1	2	3	4	5
---	---	---	---	---

Not At All Satisfied Very Satisfied

6. How successful were the coordination and logistics of the Standard Setting Conference?

1	2	3	4	5
---	---	---	---	---

Not At All Successful Very Successful

7. Overall, how satisfied are you with the standard setting review process in which you participated?

1	2	3	4	5
---	---	---	---	---

Not At All Satisfied Very Satisfied

⋮

8. Please provide any additional comments about the standard setting activities.

Enter your answer





National Benchmark Conference Report

Conducted Online April 23, 2025, and
April 28, 2025
Physics (323)
Physical Education (516)

Prepared 5/23/25 by:
Evaluation Systems group of Pearson
300 Venture Way
Hadley, MA 01035

This document constitutes a working draft that includes some materials that are test secure and/or confidential for other reasons. As such, it should not be circulated to unauthorized persons. The information set forth in this document is subject to change, and any changes will fully supersede the information contained in this document.

Copyright © 2025 Pearson Education, Inc. or its affiliate(s). All Rights Reserved.
Evaluation Systems group of Pearson, 300 Venture Way, Hadley, MA 01035



Table of Contents

OVERVIEW	3
PURPOSE OF THE NATIONAL BENCHMARK CONFERENCE.....	3
POTENTIAL USES FOR THE NATIONAL EVALUATION SERIES™	4
NATIONAL BENCHMARK PROCESS.....	4
ORIENTATION	4
SIMULATED TEST-TAKING ACTIVITY	5
JUST ACCEPTABLE CANDIDATE	5
ROUND ONE: ITEM-BASED JUDGMENTS.....	6
ROUND ONE: ANALYSIS	7
ROUND TWO: ITEM-BASED JUDGMENTS	7
ROUND TWO: ANALYSIS	7
CONFERENCE EVALUATION.....	8
NATIONAL BENCHMARK RESULTS	8
Appendix A.....	9
Appendix B	12

OVERVIEW

This report was prepared for the Educator Preparation and Pathways Committee and Oregon Teacher Standards and Practices Commission for determination related to adoption of redeveloped National Evaluation Series™ tests. This report contains the results from National Benchmark Conferences conducted for the following National Evaluation Series™ tests for teacher certification:

- Physics (323)
- Physical Education (516)

The following three groups of educators were recruited to participate on the panels:

1. Experienced public-school educators who have worked directly with teacher candidates (student teachers)
2. Early-career public school educators (with 5 or fewer years of experience)
3. Educator Preparation Program faculty members

All panelists were approved by the Director of Assessment Programs or designee prior to being invited to serve on the National Benchmark Panel. For a breakdown of the panel composition, refer to Appendix A.

PURPOSE OF THE NATIONAL BENCHMARK CONFERENCE

The purpose of the National Benchmark Conference is for expert educators and educator preparation faculty from multiple states to participate in a structured process to make individual judgments regarding recommended passing scores for teacher licensure tests. The passing score represents the minimum level of performance that a candidate must achieve to demonstrate the acceptable level of knowledge and skills required to perform the activities in which the certification allows the candidate to engage. This report presents information about the benchmarking process, a summary of the panel members, and the panel-recommended passing scores.

POTENTIAL USES FOR THE NATIONAL EVALUATION SERIES™

Each implementation of an assessment has a specific purpose. While the passing standards resulting from the National Benchmark Conference are intended for use in making certification decisions, the Series tests can be used by both states and institutions of higher education.

Adopting agencies and organizations may choose to use one or more tests for purposes other than state licensure, such as a component of completing a teacher preparation program, a component for acceptance into a preparation program, or gaining clearance to student teach. The passing standard recommended by the benchmark panel is for reference by agencies and organizations considering the role of a test for inclusion in an educator licensure program and determining a passing standard. Pearson works closely with states that opt to conduct their own standard setting rather than adopting the standard recommended by the National Benchmark panel.

NATIONAL BENCHMARK PROCESS

An iterative procedure was used in which panelists provided item-level ratings using procedures commonly referred to as a modified-Angoff method. After an orientation to the item rating methods, the panelists reviewed the test and were guided through two rounds of independent benchmark ratings. Each panel's aggregated ratings provided information about the expected test-level performance of just acceptable candidates for initial licensure.

ORIENTATION

Panelists were provided an orientation that included the background to the National Evaluation Series™ test development, an overview of the test development process, including specific explanation of the benchmark process, the concept of the “just acceptable candidate,” and the judgments about test items that the panelists would be asked to make. Panelists were encouraged to ask questions about the meeting's purpose before moving to the next step.

SIMULATED TEST-TAKING ACTIVITY

Panelists engaged in a simulated test-taking activity to familiarize themselves with the knowledge and skills measured by the test items. Each panelist was provided with a copy of the first operational test form and the test framework and was asked to read and answer the questions on the test without referring to the answer key. Panelists were then provided with the answer key and asked to score their own answers. Panelists were encouraged to write in their copy of the test booklet and to write notes to assist them with their review.

JUST ACCEPTABLE CANDIDATE

During the orientation, panelists were introduced to the concept of the “just acceptable candidate.” The facilitator guided the panel through extensive discussion of this concept, and panelists submitted a confirmation form acknowledging they understood it fully prior to the Round 1 Ratings. The definition of this term is presented below.

Just Acceptable Candidate

Just at the level of knowledge and skills important for effective job performance as a beginning teacher in a classroom in public schools

ROUND ONE: ITEM-BASED JUDGMENTS

The Pearson facilitator provided training in this step of the national benchmark process during which panelists made item-based judgments using a modified-Angoff procedure for the multiple-choice questions (MCQs).

To prepare for making their judgments, panelists were informed of the factors potentially influencing candidate performance, such as complexity of the items, content tested, item context, and language. In addition, the panelists completed a training exercise with sample test items with a range of item difficulty.

Panel members were asked to consider the group of “just acceptable candidates” having the baseline pedagogical content knowledge and knowledge of professional responsibilities as in the respective test frameworks. Applying their professional judgments in response to the question below, each panel member provided an independent rating for each multiple-choice item indicating what percent of this group would provide a correct response for each item.

Imagine a hypothetical group of individuals who are just at the level of knowledge and skills required to perform the job of an initially certified educator in this field in U.S. public schools.

What percent of this group would answer the item correctly?

0% - 10%	=	1	51% - 60%	=	6
11% - 20%	=	2	61% - 70%	=	7
21% - 30%	=	3	71% - 80%	=	8
31% - 40%	=	4	81% - 90%	=	9
41% - 50%	=	5	91% - 100%	=	10

ROUND ONE: ANALYSIS

After the panelists completed their item ratings, Pearson analyzed their rating forms to produce Round One Item Rating Summaries to be given to each panelist. The reports were prepared for each test and contained the panelist's rating, the median rating provided by all panelists rating the item, and the distribution of ratings across all panelists.

ROUND TWO: ITEM-BASED JUDGMENTS

For Round Two, panelists were provided with the Item Rating Summary from Round One. Pearson facilitators explained how to read and interpret the summary. Panelists were provided with an opportunity to ask for further clarification of how to read and interpret the analyses.

Each panelist was given an opportunity to review the ratings for each item on the test in their field made by the group of panelists and compare these ratings to his/her own. Panelists then had an opportunity to discuss the items. Based on their review of the analyses provided, panelists had the opportunity to provide a second rating to replace the first rating for any items.

ROUND TWO: ANALYSIS

After the panelists completed their Round 2 ratings, Pearson analyzed their rating forms. A Round 2 Rating Summary was produced and contained the following information: a) number of scorable items on the test form, b) number of raters, c) sum of item medians (i.e., panel-based standard for each component), and d) distribution of individual rater results sorted in descending order by score value. The Round 2 Rating Summary was based on the panel members' final rating on each item (i.e., either the unchanged Round 1 rating or the Round 2 rating if it is different from the Round 1 rating). The Round 2 Rating Summary was used by facilitators to confirm the completion of the benchmarking process. Panelists were not provided with a Round 2 Rating Summary, as the information they provided during Round 2 was final.

CONFERENCE EVALUATION

Following all activities, panel members completed a Conference Evaluation Form (see Appendix B) that allowed each panelist to anonymously rate various aspects of the National Benchmark Conference and record any comments they may have had, including about specific item content.

NATIONAL BENCHMARK RESULTS

Following the conference, Pearson calculated a recommended passing standard, or National Benchmark, for each test area based on second round ratings provided by the panel members. The National Benchmark was then scaled to a score of 220 on a scale of 100–300.

The recommended standard represents the number of multiple-choice questions a candidate must answer correctly to earn a scaled score of 220.

Table 1: National Evaluation Series™ Test Fields National Benchmarks

Field Test	Total # of Items	National Benchmark
Physics	80	57
Physical Education	80	57

Appendix A

National Benchmark Conference

Panelists

Physics	
National Benchmark Panel	
April 23, 2025	
Affiliation	
K-12	12
EPP	1
Gender	
Female	8
Male	5
Declined to State	0
Ethnicity	
American Indian/Alaska Native	0
Asian	2
Black/African American	3
Hispanic/Latino	0
Other	1
White	7
Declined to State	0

Note:

EPP: Educator Preparation Program faculty member

Physical Education National Benchmark Panel April 28, 2025	
Affiliation	
K-12	8
EPP	4
Gender	
Female	8
Male	4
Declined to State	0
Ethnicity	
American Indian/Alaska Native	0
Asian	0
Black/African American	1
Hispanic/Latino	0
Other	0
White	11
Declined to State	0

Note:

EPP: Educator Preparation Program faculty member

Appendix B

Conference Evaluation Form



National Evaluation Series National Benchmark Conference Evaluation Form

Please respond to the following questions by selecting a rating on each scale.

1. Please enter the content area for the conference you are participating in.

Enter your answer

2. Rater ID Number (optional)

Enter your answer

3. How well do you feel you understood the training that you received for making your standard setting judgments?

1	2	3	4	5
---	---	---	---	---

Not At All Well

Very Well

⋮

4. How confident are you of the standard setting judgments you made?

1	2	3	4	5
---	---	---	---	---

Not At All Confident Very Confident

5. How satisfied are you that you were able to complete the work you were asked to do in the time that was provided?

1	2	3	4	5
---	---	---	---	---

Not At All Satisfied Very Satisfied

6. How successful were the coordination and logistics of the Standard Setting Conference?

1	2	3	4	5
---	---	---	---	---

Not At All Successful Very Successful

7. Overall, how satisfied are you with the standard setting review process in which you participated?

1	2	3	4	5
---	---	---	---	---

Not At All Satisfied Very Satisfied

⋮

8. Please provide any additional comments about the standard setting activities.

Enter your answer





National Benchmark Conference Report

Conducted Online September 30, 2025
Social Science (325)

Prepared 10/24/25 by:
Evaluation Systems group of Pearson
300 Venture Way
Hadley, MA 01035

This document constitutes a working draft that includes some materials that are test secure and/or confidential for other reasons. As such, it should not be circulated to unauthorized persons. The information set forth in this document is subject to change, and any changes will fully supersede the information contained in this document.

Copyright © 2025 Pearson Education, Inc. or its affiliate(s). All Rights Reserved.
Evaluation Systems group of Pearson, 300 Venture Way, Hadley, MA 01035



Table of Contents

OVERVIEW	3
PURPOSE OF THE NATIONAL BENCHMARK CONFERENCE.....	3
POTENTIAL USES FOR THE NATIONAL EVALUATION SERIES™	4
NATIONAL BENCHMARK PROCESS.....	4
ORIENTATION	4
SIMULATED TEST-TAKING ACTIVITY	5
JUST ACCEPTABLY QUALIFIED CANDIDATE.....	5
ROUND ONE: ITEM-BASED JUDGMENTS.....	6
ROUND ONE: ANALYSIS	7
ROUND TWO: ITEM-BASED JUDGMENTS	7
ROUND TWO: ANALYSIS	7
CONFERENCE EVALUATION.....	8
NATIONAL BENCHMARK RESULTS	8
Appendix A.....	9
Appendix B	11

OVERVIEW

This report was prepared for the Educator Preparation and Pathways Committee and Oregon Teacher Standards and Practices Commission for determination related to adoption of redeveloped National Evaluation Series™ tests. This report contains the results from National Benchmark Conferences conducted for the following National Evaluation Series™ tests for teacher certification:

- Social Science (325)

The following three groups of educators were recruited to participate on the panels:

1. Experienced public-school educators who have worked directly with teacher candidates (student teachers)
2. Early-career public school educators (with 5 or fewer years of experience)
3. Educator Preparation Program faculty members

All panelists were approved by the Director of Assessment Programs or designee prior to being invited to serve on the National Benchmark Panel. For a breakdown of the panel composition, refer to Appendix A.

PURPOSE OF THE NATIONAL BENCHMARK CONFERENCE

The purpose of the National Benchmark Conference is for expert educators and educator preparation faculty from multiple states to participate in a structured process to make individual judgments regarding recommended passing scores for teacher licensure tests. The passing score represents the minimum level of performance that a candidate must achieve to demonstrate the acceptable level of knowledge and skills required to perform the activities in which the certification allows the candidate to engage. This report presents information about the benchmarking process, a summary of the panel members, and the panel-recommended passing scores.

POTENTIAL USES FOR THE NATIONAL EVALUATION SERIES™

Each implementation of an assessment has a specific purpose. While the passing standards resulting from the National Benchmark Conference are intended for use in making certification decisions, the Series tests can be used by both states and institutions of higher education.

Adopting agencies and organizations may choose to use one or more tests for purposes other than state licensure, such as a component of completing a teacher preparation program, a component for acceptance into a preparation program, or gaining clearance to student teach. The passing standard recommended by the benchmark panel is for reference by agencies and organizations considering the role of a test for inclusion in an educator licensure program and determining a passing standard. Pearson works closely with states that opt to conduct their own standard setting rather than adopting the standard recommended by the National Benchmark panel.

NATIONAL BENCHMARK PROCESS

An iterative procedure was used in which panelists provided item-level ratings using procedures commonly referred to as a modified-Angoff method. After an orientation to the item rating methods, the panelists reviewed the test and were guided through two rounds of independent benchmark ratings. Each panel's aggregated ratings provided information about the expected test-level performance of just acceptably qualified candidates for initial licensure.

ORIENTATION

Panelists were provided an orientation that included the background to the National Evaluation Series™ test development, an overview of the test development process, including specific explanation of the benchmark process, the concept of the “just acceptably qualified candidate,” and the judgments about test items that the panelists would be asked to make. Panelists were encouraged to ask questions about the meeting's purpose before moving to the next step.

SIMULATED TEST-TAKING ACTIVITY

Panelists engaged in a simulated test-taking activity to familiarize themselves with the knowledge and skills measured by the test items. Each panelist was provided with a copy of the first operational test form and the test framework and was asked to read and answer the questions on the test without referring to the answer key. Panelists were then provided with the answer key and asked to score their own answers. Panelists were encouraged to write in their copy of the test booklet and to write notes to assist them with their review.

JUST ACCEPTABLY QUALIFIED CANDIDATE

During the orientation, panelists were introduced to the concept of the “just acceptably qualified candidate.” The facilitator guided the panel through extensive discussion of this concept, and panelists submitted a confirmation form acknowledging they understood it fully prior to the Round 1 Ratings. The definition of this term is presented below.

Just Acceptably Qualified Candidate

Just at the level of knowledge and skills important for effective job performance as a beginning teacher in a classroom in public schools

ROUND ONE: ITEM-BASED JUDGMENTS

The Pearson facilitator provided training in this step of the national benchmark process during which panelists made item-based judgments using a modified-Angoff procedure for the multiple-choice questions (MCQs).

To prepare for making their judgments, panelists were informed of the factors potentially influencing candidate performance, such as complexity of the items, content tested, item context, and language. In addition, the panelists completed a training exercise with sample test items with a range of item difficulty.

Panel members were asked to consider the group of “just acceptably qualified candidates” having the baseline pedagogical content knowledge and knowledge of professional responsibilities as in the respective test frameworks. Applying their professional judgments in response to the question below, each panel member provided an independent rating for each multiple-choice item indicating what percent of this group would provide a correct response for each item.

Imagine a hypothetical group of individuals who are just at the level of knowledge and skills required to perform the job of an initially certified educator in this field in U.S. public schools.

What percent of this group would answer the item correctly?

0% - 10%	=	1	51% - 60%	=	6
11% - 20%	=	2	61% - 70%	=	7
21% - 30%	=	3	71% - 80%	=	8
31% - 40%	=	4	81% - 90%	=	9
41% - 50%	=	5	91% - 100%	=	10

ROUND ONE: ANALYSIS

After the panelists completed their item ratings, Pearson analyzed their rating forms to produce Round One Item Rating Summaries to be given to each panelist. The reports were prepared for each test and contained the panelist's rating, the median rating provided by all panelists rating the item, and the distribution of ratings across all panelists.

ROUND TWO: ITEM-BASED JUDGMENTS

For Round Two, panelists were provided with the Item Rating Summary from Round One. Pearson facilitators explained how to read and interpret the summary. Panelists were provided with an opportunity to ask for further clarification of how to read and interpret the analyses.

Each panelist was given an opportunity to review the ratings for each item on the test in their field made by the group of panelists and compare these ratings to his/her own. Panelists then had an opportunity to discuss the items. Based on their review of the analyses provided, panelists had the opportunity to provide a second rating to replace the first rating for any items.

ROUND TWO: ANALYSIS

After the panelists completed their Round 2 ratings, Pearson analyzed their rating forms. A Round 2 Rating Summary was produced and contained the following information: a) number of scorable items on the test form, b) number of raters, c) sum of item medians (i.e., panel-based standard for each component), and d) distribution of individual rater results sorted in descending order by score value. The Round 2 Rating Summary was based on the panel members' final rating on each item (i.e., either the unchanged Round 1 rating or the Round 2 rating if it is different from the Round 1 rating). The Round 2 Rating Summary was used by facilitators to confirm the completion of the benchmarking process. Panelists were not provided with a Round 2 Rating Summary, as the information they provided during Round 2 was final.

CONFERENCE EVALUATION

Following all activities, panel members completed a Conference Evaluation Form (see Appendix B) that allowed each panelist to anonymously rate various aspects of the National Benchmark Conference and record any comments they may have had, including about specific item content.

NATIONAL BENCHMARK RESULTS

Following the conference, Pearson calculated a recommended passing standard, or National Benchmark, for each test area based on second round ratings provided by the panel members. The National Benchmark was then scaled to a score of 220 on a scale of 100-300.

The recommended standard represents the number of multiple-choice questions a candidate must answer correctly to earn a scaled score of 220.

Table 1: National Evaluation Series™ Test Fields National Benchmarks

Field Test	Total # of Items	National Benchmark
Social Science	80	50

Appendix A

National Benchmark Conference

Panelists

Social Science National Benchmark Panel September 30, 2025	
Affiliation	
K-12	8
EPP	0
Gender	
Female	5
Male	3
Declined to State	0
Ethnicity	
American Indian/Alaska Native	0
Asian	0
Black/African American	3
Hispanic/Latino	0
Other	1
White	4
Declined to State	0

Note:

EPP: Educator Preparation Program faculty member

Appendix B

Conference Evaluation Form



National Evaluation Series National Benchmark Conference Evaluation Form

Please respond to the following questions by selecting a rating on each scale.

1. Please enter the content area for the conference you are participating in.

Enter your answer

2. Rater ID Number (optional)

Enter your answer

3. How well do you feel you understood the training that you received for making your standard setting judgments?

1	2	3	4	5
---	---	---	---	---

Not At All Well

Very Well

⋮

4. How confident are you of the standard setting judgments you made?

1	2	3	4	5
---	---	---	---	---

Not At All Confident Very Confident

5. How satisfied are you that you were able to complete the work you were asked to do in the time that was provided?

1	2	3	4	5
---	---	---	---	---

Not At All Satisfied Very Satisfied

6. How successful were the coordination and logistics of the Standard Setting Conference?

1	2	3	4	5
---	---	---	---	---

Not At All Successful Very Successful

7. Overall, how satisfied are you with the standard setting review process in which you participated?

1	2	3	4	5
---	---	---	---	---

Not At All Satisfied Very Satisfied

⋮

8. Please provide any additional comments about the standard setting activities.

Enter your answer



Alignment of National Evaluation Series Social Science Framework with 2024 Oregon Social Science Standards

This alignment study identifies the state standards that are addressed in whole or in part by each competency of the assessment framework. An indication of alignment does not necessarily imply complete congruence of the content of a Series test competency with the relevant standard. The information in this document is subject to change if revisions are made to the assessment framework. Any changes will fully supersede the information contained in this document.

<u>Series 325: Social Science</u>	<u>2024 Oregon Social Science Standards</u>
I. World History and Historical Thinking and Skills	
0001 Understand major developments in world history from early civilizations to 1600 CE.	6/7.H.CH.3 6/7.H.CC.4-.5 6/7.H.CP.14-.15 6/7.C.PI.1-.2, .4 6/7.C.IR.6-.7 6/7.C.DP.9 6/7.E.IC.9 6/7.E.10.ST-.11.ST

<u>Series 325: Social Science</u>	<u>2024 Oregon Social Science Standards</u>
<p>0002 Understand major developments in world history from 1600 to the present.</p>	<p>HS.US.CC.7, .9 HS.US.CE.10 HS.US.CP.16 HS.WR.CH.1-.2 HS.WR.CC.3-.6 HS.WR.CE.7-.9 HS.WR.CP.10-.19 6/7.C.PI.4 6/7.E.IC.9 6/7.E.10.ST 8.E.MI.7</p>
<p>0003 Understand historical thinking and skills.</p>	<p>Essential Disciplinary Practices II, III, IV, V, VI, VII, VIII, IX</p>

II. U.S. History**0004** Understand major developments in early U.S. history to 1877.

6/7.H.CC.6

6/7.H.CE.7

6/7.H.CP.13

8.H.CH.1-.2

8.H.CC.4

8.H.CE.5-.8

HS.US.CH.1, .3-.4

HS.US.CP.18

HS.WR.CH.2

8.C.PI.2-.5

8.C.CE.12

8.G.HE.7

6/7.E.10.ST

8.E.MI.6

8.E.ST.9

HS.E.MI.7

HS.E.ST.15

<u>Series 325: Social Science</u>	<u>2024 Oregon Social Science Standards</u>
<p>0005 Understand major developments in U.S. history from 1877 to the present.</p>	<p>HS.US.CH.2-.4 HS.US.CC.5-.8 HS.US.CE.10-.11 HS.US.CP.12-.19 HS.WR.CH.2 8.C.CE.13-.14, .18 HS.C.DP.12 8.E.MI.6 HS.E.MI.7 HS.E.ST.15</p>

<u>Series 325: Social Science</u>	<u>2024 Oregon Social Science Standards</u>
III. American Government	
<p>0006 Understand civic participation and skills, the roles and systems of government, and the rights and responsibilities of U.S. citizenship.</p>	<p>6/7.C.PI.3 6/7.C.IR.6, .8 6/7.C.CE.10 8.C.PI.1, .6 8.C.IR.9 8.C.CE.12-.14 HS.C.PI.1-2 HS.C.IR.9-10 HS.C.DP.17 HS.C.CE.19</p>
<p>0007 Understand basic principles of the U.S. Constitution and the structure and functions of different branches and levels of government in the United States.</p>	<p>6/7.C.PI.5 8.C.PI.7-.8 8.C.DP.10-.11 HS.C.PI.3, .5-.7 HS.C.IR.8 HS.C.IR.11 HS.C.DP.13-.16</p>

<u>Series 325: Social Science</u>	<u>2024 Oregon Social Science Standards</u>
IV. World Geography	
<p>0008 Understand spatial thinking, geographic tools and skills, and physical and cultural features of places and regions.</p>	<p>6/7.G.GR.1-.2 6/7.G.MM.3 6/7.G.HI.4 8.G.GR.1-.2 8.G.MM.3 8.G.HI.5-.6 8.G.HE.7 HS.G.GR.1 HS.G.MM.2 HS.G.HI.6-.7</p>
<p>0009 Understand human systems.</p>	<p>6/7.G.MM.3 6/7.G.HI.4-.5 6/7.G.HE.6-.7 8.G.GR.1-.2 8.G.MM.3-.4 8.G.HI.5-.6 8.G.HE.7 HS.G.MM.2-.5 HS.G.HI.7-8 HS.G.HE.9-.11</p>

<u>Series 325: Social Science</u>	<u>2024 Oregon Social Science Standards</u>
V. Economics and Financial Literacy	
0010 Understand the core ideas in economics.	6/7.E.MI.4-.7 6/7.E.IC.8 6/7.E.12.ST 8.E.MI.5 8.E.IC.8 HS.E.MI.1-.4, .6 HS.E.IC.8-.10 HS.E.ST.11-.14
0011 Understand fundamental concepts of financial literacy.	6/7.E.ES.1-.3 8.E.ES.1-.4 HS.E.MI.5