FR-10-77a

HumRRO

Human Resources Research Organization

Missouri Assessment Program End-of-Course (EOC) Assessment Forms Validation Study: Appendices A through D

Leslie R. Taylor
Hilary L. Campbell
Richard C. Deatz
Rebecca N. Dvorak
Lisa E. Koger
Arthur A. Thacker

Prepared for: Missouri Department of Elementary and Secondary Education
205 Jefferson Street
P.O. Box 480
Jefferson City, Missouri 65102

Prepared under: Contract No: C308004001-003

March 8, 2011

66 Canal Center Plaza, Suite 400 • Alexandria, VA 22314 • 703.549.3611 • fax. 703.549.9025
www.humrro.org

FR-10-77a

HumRRO

Human Resources Research Organization

Missouri Assessment Program End-of-Course (EOC) Assessment Forms Validation Study: Appendices A through D

Leslie R. Taylor
Hilary L. Campbell
Richard C. Deatz
Rebecca N. Dvorak
Lisa E. Koger
Arthur A. Thacker

Prepared for: Missouri Department of Elementary and Secondary Education
205 Jefferson Street
P.O. Box 480
Jefferson City, Missouri 65102

Prepared under: Contract No: C308004001-003

March 8, 2011

66 Canal Center Plaza, Suite 400 • Alexandria, VA 22314 • 703.549.3611 • fax. 703.549.9025
www.humrro.org

MISSOURI ASSESSMENT PROGRAM END-OF-COURSE (EOC) ASSESSMENT FORMS ALIGNMENT FORMS VALIDATION STUDY: APPENDICES A THROUGH E

TABLE OF CONTENTS

Appendix A EOC English II: Detailed Statistical Results A-1

Webb Alignment Indicators A-1
Categorical Concurrence A-1
Depth-of-Knowledge Consistency.....A-2
Range-of-Knowledge Correspondence.....A-11
Balance-of-Knowledge Representation A-13
Consensus DOK Ratings on English II CLEs.....A-15
English II Item Alignment to CLEs.....A-18
English II Item Quality Ratings A-21
Panelist Comments on English II Items A-24
English II CLEs Matched to Items A-26

Appendix B EOC Algebra I: Detailed Statistical Results B-1

Webb Alignment Indicators B-1
Categorical Concurrence B-1
Depth-of-Knowledge Consistency.....B-2
Range-of-Knowledge Correspondence.....B-10
Balance-of-Knowledge Representation B-13
Consensus DOK Ratings on Algebra I CLEs B-15
Algebra I Item Alignment to CLEs B-16
Algebra I Item Quality Ratings B-19
Panelist Comments on Algebra I Items B-22
Algebra I CLEs Matched to Items B-24

Appendix C EOC Biology: Detailed Statistical Results..... C-1

Webb Alignment Indicators C-1
Categorical Concurrence C-1
Depth-of-Knowledge Consistency..... C-3
Range-of-Knowledge Correspondence..... C-10
Balance-of-Knowledge Representation C-12
Consensus DOK Ratings on CLEs C-14
Biology Item Alignment to CLEsC-..... C-17
Biology Item Quality Ratings C-20
Panelist Comments on Biology Items C-23
Biology CLEs Matched to Items C-27

Appendix D Sample Alignment Review Materials C-1

English C-1
Reading DOK Levels C-1
Writing DOK Levels C-3
Algebra I C-5

DOK Levels *C-5*
Biology C-7
Biology DOK Levels *C-7*

List of Tables

Table A-1. Categorical Concurrence for English II 2009 Test Form: Mean Number of Items per Big Idea with Multiple-Choice and Performance Event Items A-1
Table A-2. Categorical Concurrence for English II 2009, 2010, and 2011 Test Forms: Mean Number of Items per Big Idea with Multiple-Choice Items Only.....A-2
Table A-3. DOK Consistency for English II 2009 Test Form: Mean Percentage of Multiple-Choice and Performance Event Items Below, At, or Above Corresponding CLEs A-3
Table A-4. DOK Consistency for English II 2009 Test Form: Mean Percentage of Multiple-Choice Items Below, At, or Above Corresponding CLEs.....A-4
Table A-5. DOK Consistency for English II 2010 Test Form: Mean Percentage of Multiple-Choice Items Below, At, or Above Corresponding CLEs.....A-5
Table A-6. DOK Consistency for English II 2011 Test Form: Mean Percentage of Multiple-Choice Items Below, At, or Above Corresponding CLEs.....A-6
Table A-7. DOK Consistency for English II 2009 Test Form: Mean Percentage of CLEs Assessed by Multiple-Choice Items and Performance Events Below, At, or Above Expected DOK Level. A-7
Table A-8. DOK Consistency for English II 2009 Test Form: Mean Percentage of CLEs Assessed by Multiple-Choice Items Below, At, or Above Expected DOK Level. .A-8
Table A-9. DOK Consistency for English II 2010 Test Form: Mean Percentage of CLEs Assessed by Multiple-Choice Items Below, At, or Above Expected DOK Level. .A-9
Table A-10. DOK Consistency for English II 2011 Test Form: Mean Percentage of CLEs Assessed by Multiple-Choice Items Below, At, or Above Expected DOK Level. .A-10
Table A-7. Range-of-Knowledge for English II 2009 Test Form: Mean CLEs per Big Idea Linked with Multiple-Choice and Performance Event Items.....A-11
Table A-8. Range-of-Knowledge for English II 2009, 2010, and 2011 Test Forms: Mean CLEs per Big Idea Linked with Multiple-Choice Items Only.....A-12
Table A-9. Balance-of-Knowledge Representation for English II 2009 Test Form: Balance Index per Big Idea A-13
Table A-10. Balance-of-Knowledge Representation for English II 2009, 2010, and 2011 Test Forms: Balance Index per Big Idea A-14
Table A-11. English II: Group Consensus Ratings on DOK Level per CLE.....A-15
Table A-12. Mean Overall Alignment Rating per Item for Each English II Test Form.A-18
Table A-13. Mean Overall Quality Rating per Item for Each English II Test Form.....A-21
Table A-14. Reviewer Comments on 2009 Test Form Items for English II.....A-24
Table A-15. Reviewer Comments on 2010 Test Form Items for English II.....A-25
Table A-16. Reviewer Comments on 2011 Test Form Items for English II.....A-25

Table A-17. English II Items Matched to CLEs by Test Form Year A-26
Table B-1. Categorical Concurrence for Algebra I 2009 Test Form: Mean Number of Multiple-Choice and Performance Event Items per Strand B-1
Table B-2. Categorical Concurrence for Algebra I 2009, 2010, and 2011 Test Forms: Mean Number Multiple-Choice Items Only per Strand B-2
Table B-3. DOK Consistency for Algebra I 2009 Test Form: Mean Percentage of Multiple-Choice Items and Performance Event Items Below, At, or Above Corresponding CLEs B-3
Table B-4. DOK Consistency for Algebra I 2009 Test Form: Mean Percentage of Multiple-Choice Items Below, At, or Above Corresponding CLEs.....B-4
Table B-5. DOK Consistency for Algebra I 2010 Test Form: Mean Percentage of Multiple-Choice Items Below, At, or Above Corresponding CLEs.....B-5
Table B-6. DOK Consistency for Algebra I 2011 Test Form: Mean Percentage of Multiple-Choice Items Below, At, or Above Corresponding CLEs.....B-6
Table B-7. DOK Consistency for Algebra I 2009 Test Form: Mean Percentage of CLEs Assessed by Multiple-Choice Items and Performance Events Below, At, or Above Expected DOK Level B-7
Table B-8. DOK Consistency for Algebra I 2009 Test Form: Mean Percentage of CLEs Assessed by Multiple-Choice Items Below, At, or Above Expected DOK Level B-8
Table B-9. DOK Consistency for Algebra I 2010 Test Form: Mean Percentage of CLEs Assessed by Multiple-Choice Items Below, At, or Above Expected DOK Level B-9
Table B-10. DOK Consistency for Algebra I 2011 Test Form: Mean Percentage of CLEs Assessed by Multiple-Choice Items Below, At, or Above Expected DOK Level ..B-10
Table B-7. Range-of-Knowledge for Algebra I 2009 Test Form: Mean CLEs per Strand Linked with Multiple Choice and Performance Items B-11
Table B-8. Range-of-Knowledge for Algebra I 2009, 2010, and 2011 Test Forms: Mean CLEs per Strand Linked with Multiple Choice Items Only.....B-12
Table B-9. Balance-of-Knowledge Representation for Algebra I 2009 Test Form: Balance Index per Strand B-13
Table B-10. Balance-of-Knowledge Representation for Algebra I 2009, 2010, and 2011 Test Forms: Balance Index per Strand B-14
Table B-11. Algebra I: Group Consensus Ratings on DOK Level per CLE B-15
Table B-12. Mean Overall Alignment Rating per Item for Each Algebra I Test Form .B-16
Table B-13. Mean Overall Quality Rating per Item for Each Algebra I Test Form B-19
Table B-14. Reviewer Comments on 2009 Test Form Items for Algebra I B-22
Table B-15. Reviewer Comments on 2010 Test Form Items for Algebra I B-22
Table B-16. Reviewer Comments on 2011 Test Form Items for Algebra I B-23
Table B-17. Items Matched to Algebra I CLEs by Test Form Year.....B-24
Table C-1. Categorical Concurrence for Biology 2009 Test Form: Mean Number Multiple-Choice and Performance Event Items per Strand C-1
Table C-2. Categorical Concurrence for Biology 2009, 2010, and 2011 Test Forms: Mean Number Multiple-Choice Only Items per Strand C-2
Table C-3. DOK Consistency for Biology 2009 Test Form: Mean Percentage of Multiple-Choice Items and Performance Event Below, At, or Above Corresponding CLEs C-3

Table C-4. DOK Consistency for Biology 2009 Test Form: Mean Percentage of Multiple-Choice Items Below, At, or Above Corresponding CLEs C-4
Table C-5. DOK Consistency for Biology 2010 Test Form: Mean Percentage of Multiple-Choice Items Below, At, or Above Corresponding CLEs C-5
Table C-6. DOK Consistency for Biology 2011 Test Form: Mean Percentage of Multiple-Choice Items Below, At, or Above Corresponding CLEs C-6
Table C-7. DOK Consistency for Biology 2009 Test Form: Mean Percentage of CLEs Assessed by Multiple-Choice Items and Performance Events Below, At, or Above Expected DOK Level. C-7
Table C-8. DOK Consistency for Biology 2009 Test Form: Mean Percentage of CLEs Assessed By Multiple-Choice Items Below, At, or Above Expected DOK Level. C-8
Table C-9. DOK Consistency for Biology 2010 Test Form: Mean Percentage of CLEs Assessed By Multiple-Choice Items Below, At, or Above Expected DOK Level. C-9
Table C-10. DOK Consistency for Biology 2011 Test Form: Mean Percentage of CLEs Assessed By Multiple-Choice Items Below, At, or Above Expected DOK Level. C-10
Table C-7. Range-of-Knowledge for Biology 2009 Test Form: Mean CLEs per Strand Linked with Multiple-Choice and Performance Event Items..... C-11
Table C-8. Range-of-Knowledge for Biology 2009, 2010, and 2011 Test Form: Mean CLEs per Strand Linked with Multiple-Choice Items..... C-11
Table C-9. Balance-of-Knowledge Representation for Biology 2009 Test Form: Balance Index per Strand C-12
Table C-10. Balance-of-Knowledge Representation for Biology 2009, 2010, and 2011 Test Forms: Balance Index per Strand C-13
Table C-11. Biology: Group Consensus Ratings on DOK Level per CLE C-14
Table C-12. Mean Overall Alignment Rating per Item for Each Biology Test Form... C-17
Table C-13. Mean Overall Quality Rating per Item for Each Biology Test Form C-20
Table C-14. Reviewer Comments on 2009 Test Form Items for Biology..... C-23
Table C-15. Reviewer Comments on 2010 Test Form Items for Biology..... C-24
Table C-16. Reviewer Comments on 2010 Test Form Items for Biology..... C-25
Table C-17. Items Matched to Biology CLEs by Test Form Year C-27

Appendix A
EOC English II: Detailed Statistical Results

In Appendix A, we present the full alignment results for English II. These results include (a) the four Webb measures, (b) consensus DOK ratings by CLE, (c) item alignment and quality ratings, (d) summary reviewer comments, and (e) items matched to course-level expectations (CLEs). Note that we performed the analyses for English II at the level of the Big Idea per strand.

For each analysis, we display the results first for the 2009 test form conducted on *all* operational items (multiple-choice and performance events). We then present results of analyses on the three test forms (2009 included) with only the multiple-choice items[1].

Webb Alignment Indicators

The following tables include complete statistical results on the four Webb alignment indicators: Categorical Concurrence, Depth-of-Knowledge (DOK) Consistency, Range of Knowledge, and Balance of Knowledge.

Categorical Concurrence

Tables A-1 and A-2 include categorical concurrence results: the mean number of items matched to Big Idea by panelists, the standard deviation (S.D.) among panelists' ratings, and the final alignment conclusion (Yes or No). The criterion for acceptable Categorical Concurrence is a minimum of six items per Big Idea.

Table A-1. Categorical Concurrence for English II 2009 Test Form: Mean Number of Items per Big Idea with Multiple-Choice and Performance Event Items

Big Idea	2009 Test Form		
	Mean Items per Big Idea	S.D.	At Least Six Items per Big Idea
Reading - Processes	12.40	2.19	Yes
Reading - Fiction	9.60	1.52	Yes
Reading - Nonfiction	7.80	1.30	Yes
Writing - Process	0	0	No
Writing - Text Development	5.00	0.71	No
Writing - Forms/Types	1.00	0.00	No
Big Ideas Matched to Six or More Items			3 of 6

Note: The *total* number of items matched to the Writing strand does meet the minimum requirement of six items.

[1] As a reminder to the reader, reviewers only rated performance events for the 2009 test forms.

Table A-2. Categorical Concurrence for English II 2009, 2010, and 2011 Test Forms: Mean Number of Items per Big Idea with Multiple-Choice Items Only

Big Idea	2009 Test Form			2010 Test Form			2011 Test Form		
	Mean Items per Big Idea	S.D.	At Least Six Items per Big Idea	Mean Items per Big Idea	S.D.	At Least Six Items per Big Idea	Mean Items per Big Idea	S.D.	At Least Six Items per Big Idea
Reading - Processes	12.60	2.19	Yes	13.40	2.07	Yes	11.60	1.34	Yes
Reading - Fiction	9.60	1.52	Yes	10.40	1.52	Yes	5.60	1.95	[a] Yes
Reading - Nonfiction	7.80	1.30	Yes	6.20	2.17	Yes	12.60	3.21	Yes
Writing - Process	0	0	No	0	0	No	0	0	No
Writing - Text Development	4.80	0.45	No	5.00	0.00	No	5.00	0.00	No
Writing - Forms/Types	0	0	No	0	0	No	0	0	No
Big Ideas Matched to Six or More Items			3 of 6			3 of 6			3 of 6

[a] Mean number of items is just below minimum decision criterion. Range = 4 to 9 items.

Depth-of-Knowledge Consistency

Tables A-3 through A-10 present results of a comparative analysis between assessment items and CLEs on depth-of-knowledge (DOK). Tables A-3 through A-6 focus on the test item DOK relative to the corresponding CLEs. Specifically, these tables include the mean percentage of items per Strand rated below, at the same level, or above the DOK of the corresponding CLE. Webbs' criterion for acceptable DOK consistency is that item DOK must be At or Above the DOK level of the matched standard for at least 50% of items. Across the CLEs per Strand, we note (Yes or No) whether 50% of total items assessed CLEs as the appropriate cognitive level. Note that the Webb method compares item DOK values to the consensus DOK values determined by reviewers, which may differ from the State published DOK levels per CLE in some cases.

Table A-3. DOK Consistency for English II 2009 Test Form: Mean Percentage of Multiple-Choice and Performance Event Items Below, At, or Above Corresponding CLEs

Big Idea	2009 Test Form									
	[a] Mean Items per Big Idea	[b] Percent Items Below		[c] Percent Items Same		[d] Percent Items Above		[e] Percent Items At/Above DOK of CLE		[f] 50% or More Items At/Above DOK of CLE
		M	S.D.	M	S.D.	M	S.D.	M	S.D.	
Reading - Processes	12.40	39%	0.17	57%	0.13	4%	0.05	61%	0.17	Yes
Reading - Fiction	9.60	31%	0.21	60%	0.18	9%	0.10	69%	0.21	Yes
Reading - Nonfiction	7.80	63%	0.20	37%	0.20	0	0	37%	0	No
Writing - Process	0	0	0	0	0	0	0	0	0	No
Writing - Text Development	5.00	0	0	97%	0.07	3%	0.07	100%	0	Yes
Writing - Forms/Types	1.00	100%	0	0	0	0	0	0	0	No
Big Ideas with CLEs Assessed Appropriately										3 of 6

[a] Mean Items per Big Idea = Mean number of items matched to strand (Categorical Concurrence)
[b] Items Below = Percentage of items below DOK level of CLEs per strand.
[c] Items Same = Percentage of items with same DOK level as CLEs per strand.
[d] Items Above = Percentage of items above DOK level of CLEs per strand.
[e] Items At and Above = Percentage of items, when added, with DOK at the Same level and Above level of CLE.
[f] 50% or More Items = At least half of items assessing strand matched DOK level of corresponding CLEs.

Table A-4. DOK Consistency for English II 2009 Test Form: Mean Percentage of Multiple-Choice Items Below, At, or Above Corresponding CLEs

Big Idea	2009 Test Form									
	[a] Mean Items per Big Idea	[b] Percent Items Below		[c] Percent Items At		[d] Percent Items Above		[e] Percent Items At/Above DOK of CLE		[f] 50% or More Items At/Above DOK of CLE
		M	S.D.	M	S.D.	M	S.D.	M	S.D.	
Reading - Processes	12.60	39%	0.17	57%	0.13	4%	0.05	61%	0.17	Yes
Reading - Fiction	9.60	31%	0.21	60%	0.18	9%	0.10	69%	0.21	Yes
Reading - Nonfiction	7.80	63%	0.20	37%	0.20	0	0	37%	0	No
Writing - Process	0	0	0	0	0	0	0	0	0	No
Writing - Text Development	4.80	0	0	100%	0	0	0	100%	0	Yes
Writing - Forms/Types	0	0	0	0	0	0	0	0	0	No
Big Ideas with CLEs Assessed Appropriately										3 of 6

[a] Mean Items per Big Idea = Mean number of items matched to strand (Categorical Concurrence)
[b] Items Below = Percentage of items below DOK level of CLEs per strand.
[c] Items Same = Percentage of items with same DOK level as CLEs per strand.
[d] Items Above = Percentage of items above DOK level of CLEs per strand.
[e] Items At and Above = Percentage of items, when added, with DOK at the Same level and Above level of CLE.
[f] 50% or More Items = At least half of items assessing strand matched DOK level of corresponding CLEs.

Table A-5. DOK Consistency for English II 2010 Test Form: Mean Percentage of Multiple-Choice Items Below, At, or Above Corresponding CLEs

Big Idea	2010 Test Form									
	[a] Mean Items per Big Idea	[b] Percent Items Below		[c] Percent Items Same		[d] Percent Items Above		[e] Percent Items At/Above DOK of CLE		[f] 50% or More Items At/Above DOK of CLE
		M	S.D.	M	S.D.	M	S.D.	M	S.D.	
Reading - Processes	13.40	48%	0.14	52%	0.14	0%	0	52%	0.14	Yes
Reading - Fiction	10.40	53%	0.20	47%	0.20	0	0	47%	0.20	No
Reading - Nonfiction	6.20	58%	0.30	42%	0.30	0	0	42%	0.30	No
Writing - Process	0	0	0	0	0	0	0	0	0	No
Writing - Text Development	5.00	0	0	84%	0.26	16%	0.26	100%	0	Yes
Writing - Forms/Types	0	0	0	0	0	0	0	0	0	No
Big Ideas with CLEs Assessed Appropriately										2 of 6

[a] Mean Items per Big Idea = Mean number of items matched to strand (Categorical Concurrence)
[b] Items Below = Percentage of items below DOK level of CLEs per strand.
[c] Items Same = Percentage of items with same DOK level as CLEs per strand.
[d] Items Above = Percentage of items above DOK level of CLEs per strand.
[e] Items At and Above = Percentage of items, when added, with DOK at the Same level and Above level of CLE.
[f] 50% or More Items = At least half of items assessing strand matched DOK level of corresponding CLEs.

Table A-6. DOK Consistency for English II 2011 Test Form: Mean Percentage of Multiple-Choice Items Below, At, or Above Corresponding CLEs

Big Idea	2011 Test Form									
	[a] Mean Items per Big Idea	[b] Percent Items Below		[c] Percent Items Same		[d] Percent Items Above		[e] Percent Items At/Above DOK of CLE		[f] 50% or More Items At/Above DOK of CLE
		M	S.D.	M	S.D.	M	S.D.	M	S.D.	
Reading - Processes	11.60	37%	0.12	62%	0.1	1%	0.03	63%	0.12	Yes
Reading - Fiction	5.60	60%	0.20	40%	0.20	0	0	40%	0.20	No
Reading - Nonfiction	12.60	59%	0.20	41%	0.20	0	0	41%	0.20	No
Writing - Process	0	0	0	0	0	0	0	0	0	No
Writing - Text Development	5.00	0	0	96%	0.09	4%	0.09	100%	0	Yes
Writing - Forms/Types	0	0	0	0	0	0	0	0	0	No
Big Ideas with CLEs Assessed Appropriately										2 of 6

[a] Mean Items per Big Idea = Mean number of items matched to strand (Categorical Concurrence)
[b] Items Below = Percentage of items below DOK level of CLEs per strand.
[c] Items Same = Percentage of items with same DOK level as CLEs per strand.
[d] Items Above = Percentage of items above DOK level of CLEs per strand.
[e] Items At and Above = Percentage of items, when added, with DOK at the Same level and Above level of CLE.
[f] 50% or More Items = At least half of items assessing strand matched DOK level of corresponding CLEs.

Tables A-7 and A-10 summarize the same data in a different way by focusing on the percentage of CLEs assessed At or Above the DOK level expected. Tables display the mean percentage of *standards* (CLEs) per Big Idea assessed at the appropriate DOK level (item DOK and standard DOK are the same), as well as the number of standards assessed below and above the level expected. At least 50% of items must be At or Above the DOK level of the corresponding CLE in order for the assessment of that CLE to be judged minimally appropriate.

Table A-7. DOK Consistency for English II 2009 Test Form: Mean Percentage of CLEs Assessed by Multiple-Choice Items and Performance Events Below, At, or Above Expected DOK Level.

Big Idea	[a] Number of CLEs	2009 Test Form								
		[b] Percent CLEs Assessed Below DOK		[c] Percent CLEs Assessed At DOK		[d] Percent CLEs Assessed Above DOK		[e] Percent CLEs Assessed At/Above DOK Expected		[f] 50% or More Items At/Above DOK of CLE
		M	S.D.	M	S.D.	M	S.D.	M	S.D.	
Reading - Processes	3	28%	0.18	67%	0.2	5%	0.11	72%	0.18	Yes
Reading - Fiction	3	27%	0.28	55%	0.2	18%	0.17	73%	0.28	Yes
Reading - Nonfiction	3	65%	0.1	35%	0.1	0	0	35%	0	No
Writing - Process	1	0	0	0	0	0	0	0	0	No
Writing - Text Development	5	0	0	100%	0	0	0	100%	0	Yes
Writing - Forms/Types	1	100%	0	0	0	0	0	0	0	No
Big Ideas with CLEs Assessed Appropriately										3 of 6

[a] Number of CLEs = Per Course Level Expectations 2.0, the number of CLEs per Strand.
[b] CLEs Assessed Below DOK = Percentage of CLEs per Strand assessed below consensus DOK level.
[c] CLEs Assessed At DOK = Percentage of CLEs per Strand assessed at the consensus DOK level.
[d] CLEs Assessed Above DOK = Percentage of CLEs per Strand assessed above consensus DOK level.
[e] CLEs Assessed At/Above DOK = Combined percentage of CLEs per Strand assessed at or above consensus DOK level.
[f] 50% or More CLEs = At least half of CLEs per strand were assessed by items at the appropriate DOK level.

Table A-8. DOK Consistency for English II 2009 Test Form: Mean Percentage of CLEs Assessed by Multiple-Choice Items Below, At, or Above Expected DOK Level.

Big Idea	[a] Number of CLEs	2009 Test Form								
		[b] Percent CLEs Assessed Below DOK		[c] Percent CLEs Assessed At DOK		[d] Percent CLEs Assessed Above DOK		[e] Percent CLEs Assessed At/Above DOK Expected		[f] 50% or More Items At/Above DOK of CLE
		M	S.D.	M	S.D.	M	S.D.	M	S.D.	
Reading - Processes	3	28%	0.18	67%	0.20	5%	0.11	72%	0.18	Yes
Reading - Fiction	3	27%	0.28	55%	0.20	18%	0.17	73%	0.28	Yes
Reading - Nonfiction	3	65%	0.09	35%	0.09	0	0	35%	0.09	No
Writing - Process	1	0	0	0	0	0	0	0	0	No
Writing - Text Development	5	0	0	100%	0.00	0	0	100%	0.00	Yes
Writing - Forms/Types	1	0	0	0	0	0	0	0	0	No
Big Ideas with CLEs Assessed Appropriately										3 of 6

[a] Number of CLEs = Per Course Level Expectations 2.0, the number of CLEs per Strand.
[b] CLEs Assessed Below DOK = Percentage of CLEs per Strand assessed below consensus DOK level.
[c] CLEs Assessed At DOK = Percentage of CLEs per Strand assessed at the consensus DOK level.
[d] CLEs Assessed Above DOK = Percentage of CLEs per Strand assessed above consensus DOK level.
[e] CLEs Assessed At/Above DOK = Combined percentage of CLEs per Strand assessed at or above consensus DOK level.
[f] 50% or More CLEs = At least half of CLEs per strand were assessed by items at the appropriate DOK level.

Table A-9. DOK Consistency for English II 2010 Test Form: Mean Percentage of CLEs Assessed by Multiple-Choice Items Below, At, or Above Expected DOK Level.

Big Idea	[a] Number of CLEs	2010 Test Form								
		[b] Percent CLEs Assessed Below DOK		[c] Percent CLEs Assessed At DOK		[d] Percent CLEs Assessed Above DOK		[e] Percent CLEs Assessed At/Above DOK Expected		[f] 50% or More Items At/Above DOK of CLE
		M	S.D.	M	S.D.	M	S.D.	M	S.D.	
Reading - Processes	3	30%	0.27	70%	0.27	0	0	70%	0.27	Yes
Reading - Fiction	3	48%	0.29	52%	0.29	0	0	52%	0.29	Yes
Reading - Nonfiction	3	58%	0.28	42%	0.28	0	0	42%	0.28	No
Writing - Process	1	0	0	0	0	0	0	0	0	No
Writing - Text Development	5	0	0	80%	0.45	20%	0.45	100%	0.00	Yes
Writing - Forms/Types	1	0	0	0	0	0	0	0	0	No
Big Ideas with CLEs Assessed Appropriately										3 of 6

[a] Number of CLEs = Per Course Level Expectations 2.0, the number of CLEs per Strand.
[b] CLEs Assessed Below DOK = Percentage of CLEs per Strand assessed below consensus DOK level.
[c] CLEs Assessed At DOK = Percentage of CLEs per Strand assessed at the consensus DOK level.
[d] CLEs Assessed Above DOK = Percentage of CLEs per Strand assessed above consensus DOK level.
[e] CLEs Assessed At/Above DOK = Combined percentage of CLEs per Strand assessed at or above consensus DOK level.
[f] 50% or More CLEs = At least half of CLEs per strand were assessed by items at the appropriate DOK level.

Table A-10. DOK Consistency for English II 2011 Test Form: Mean Percentage of CLEs Assessed by Multiple-Choice Items Below, At, or Above Expected DOK Level.

Big Idea	[a] Number of CLEs	2011 Test Form								
		[b] Percent CLEs Assessed Below DOK		[c] Percent CLEs Assessed At DOK		[d] Percent CLEs Assessed Above DOK		[e] Percent CLEs Assessed At/Above DOK Expected		[f] 50% or More Items At/Above DOK of CLE
		M	S.D.	M	S.D.	M	S.D.	M	S.D.	
Reading - Processes	3	30%	0.27	70%	0.27	0	0	70%	0.27	Yes
Reading - Fiction	3	63%	0.34	37%	0.34	0	0	37%	0.34	No
Reading - Nonfiction	3	63%	0.25	37%	0.25	0	0	37%	0.25	No
Writing - Process	1	0	0	0	0	0	0	0	0	No
Writing - Text Development	5	0	0	100%	0.00	0	0	100%	0.00	Yes
Writing - Forms/Types	1	0	0	0	0	0	0	0	0	No
Big Ideas with CLEs Assessed Appropriately										2 of 6

[a] Number of CLEs = Per Course Level Expectations 2.0, the number of CLEs per Strand.
[b] CLEs Assessed Below DOK = Percentage of CLEs per Strand assessed below consensus DOK level.
[c] CLEs Assessed At DOK = Percentage of CLEs per Strand assessed at the consensus DOK level.
[d] CLEs Assessed Above DOK = Percentage of CLEs per Strand assessed above consensus DOK level.
[e] CLEs Assessed At/Above DOK = Combined percentage of CLEs per Strand assessed at or above consensus DOK level.
[f] 50% or More CLEs = At least half of CLEs per strand were assessed by items at the appropriate DOK level.

Range-of-Knowledge Correspondence

Tables A-7 and A-8 present the range-of-knowledge correspondence results. The tables include the mean number (and percentage) of CLEs matched to at least one item per Big Idea. For acceptable range, a minimum of 50% of CLEs within each Big Idea should be matched to at least one item.

Table A-7. Range-of-Knowledge for English II 2009 Test Form: Mean CLEs per Big Idea Linked with Multiple-Choice and Performance Event Items

Big Idea	Number of CLEs	2009 Test Form					
		Mean Items per Big Idea	Number of CLEs Assessed		% CLEs Assessed		50% or More CLEs
			M	S.D.	M	S.D.	
Reading - Processes	3	12.40	3.00	0.00	100	0.00	Yes
Reading - Fiction	3	9.60	3.00	0.00	100	0.00	Yes
Reading - Nonfiction	3	7.80	3.00	0.00	100	0.00	Yes
Writing - Process	1	0	0	0	0	0	No
Writing - Text Development	5	5.00	1.00	0.00	20	0.00	No
Writing - Forms/Types	1	1.00	1.00	0.00	100	0.00	Yes
Big Ideas with CLEs Assessed by At Least One Item							4 of 6

Table A-8. Range-of-Knowledge for English II 2009, 2010, and 2011 Test Forms: Mean CLEs per Big Idea Linked with Multiple-Choice Items Only

Big Idea	Number of CLEs	2009 Test Form							2010 Test Form						2011 Test Form					
		Mean Items per Big Idea	Number of CLEs Assessed		% CLEs Assessed		50% or More CLEs		Mean Items per Big Idea	Number of CLEs Assessed		% CLEs Assessed		50% or More CLEs	Mean Items per Big Idea	Number of CLEs Assessed		% CLEs Assessed		50% or More CLEs
			M	S.D.	M	S.D.				M	S.D.	M	S.D.			M	S.D.	M	S.D.	
Reading - Processes	3	12.6	3.00	0.00	100	0.00	Yes		13.4	2.00	0.00	67	0.00	Yes	11.6	2.00	0.00	67	0.00	Yes
Reading - Fiction	3	9.6	3.00	0.00	100	0.00	Yes		10.4	2.40	0.55	80	0.18	Yes	5.6	1.80	0.45	60	0.15	Yes
Reading - Nonfiction	3	7.8	3.00	0.00	100	0.00	Yes		6.2	2.80	0.45	93	0.15	Yes	12.6	2.80	0.45	93	0.15	Yes
Writing - Process	1	0	0	0	0	0	No		0	0	0	0	0	No	0	0	0	0	0	No
Writing - Text Development	5	4.8	1.00	0.00	20	0.00	No		5	1.00	0.00	20	0.00	No	5	1.00	0.00	20	0.00	No
Writing - Forms/Types	1	0	0	0	0	0	No		0	0	0	0	0	No	0	0	0	0	0	No
Big Ideas with CLEs Assessed by At Least One Item							3 of 6							3 of 6						3 of 6

Balance-of-Knowledge Representation

Tables A-9 and A-10 display the balance indices for each Big Idea per strand. This index is based on the mean number of items matched to each CLE. The minimum acceptable balance index is 70 out of 100. The table also includes the percentage of items linked to each Big Idea per strand.

Table A-9. Balance-of-Knowledge Representation for English II 2009 Test Form: Balance Index per Big Idea

Big Idea	CLEs per Big Idea	2009 Test Form					
		Mean CLEs Linked with Items	Mean Items per Big Idea	% Items Linked to Big Idea		Balance Index	Balance Index Target Met
				M	S.D.		
Reading - Processes	3	3.00	12.40	35	0.40	0.81	Yes
Reading - Fiction	3	3.00	9.60	28	0.28	0.83	Yes
Reading - Nonfiction	3	3.00	7.80	22	0.24	0.82	Yes
Writing - Process	1	0	0	0	0	0	N/A
Writing - Text Development	5	1.00	5.00	14	0.08	1.00	Yes
Writing - Forms/Types	1	1.00	1.00	3	0.00	1.00	Yes
Big Ideas Met Minimum Index							5 of 6

Note: N/A indicates that no balance index was calculated for Writing-Process because reviewers did not match items to this Big Idea.

Table A-10. Balance-of-Knowledge Representation for English II 2009, 2010, and 2011 Test Forms: Balance Index per Big Idea

Big Idea	CLEs per Big Idea	2009 Test Form						2010 Test Form						2011 Test Form					
		Mean CLEs Linked with Items	Mean Items per Big Idea	% Items Linked to Big Idea		Balance Index	Balance Index Target Met	Mean CLEs Linked with Items	Mean Items per Big Idea	% Items Linked to Big Idea		Balance Index	Balance Index Target Met	Mean CLEs Linked with Items	Mean Items per Big Idea	% Items Linked to Big Idea		Balance Index	Balance Index Target Met
				M	S.D.					M	S.D.					M	S.D.		
Reading - Processes	3	3.00	12.60	36	0.40	0.81	Yes	2.00	13.40	38	0.36	0.71	Yes	2.00	11.60	33	0.21	0.91	Yes
Reading - Fiction	3	3.00	9.60	28	0.28	0.83	Yes	2.40	10.40	30	0.26	0.75	Yes	1.80	5.60	16	0.30	0.91	Yes
Reading - Nonfiction	3	3.00	7.80	22	0.24	0.82	Yes	2.80	6.20	18	0.38	0.90	Yes	2.80	12.60	36	0.49	0.83	Yes
Writing - Process	1	0	0	0	0	0	N/A	0	0	0	0	0	N/A	0	0	0	0	0	N/A
Writing - Text Development	5	1.00	4.80	14	0.08	1.00	Yes	1.00	5.00	14	0.00	1.00	Yes	1.00	5.00	14	0.00	1.00	Yes
Writing - Forms/Types	1	0	0	0	0	0	N/A	0	0	0	0	0	No	0	0	0	0	0	N/A
Big Ideas Met Minimum Index							4 of 6						4 of 6						4 of 6

Note: N/A indicates that no balance index was calculated for Writing-Process because reviewers did not match items to this Big Idea.

Consensus DOK Ratings on English II CLEs

Table A-11 presents DOK ratings established through group consensus for each English II CLE. Column 1 lists the Strand letter, Big Idea number, and Concept letter, while Column 3 displays the CLE content description. If the CLE includes multiple parts, these parts are lettered consecutively, as reflected in Column 2. Column 4 indicates the DOK rating assigned to the CLE by the group. Note that a single DOK rating applies to an *entire* CLE (including each part).

Table A-11. English II: Group Consensus Ratings on DOK Level per CLE

Strand, Big Idea, Concept	CLE Component	CLE_Description	DOK
R1E	a	Develop vocabulary through text, using roots and affixes	2
	b	Develop vocabulary through text, using context clues	
	c	Develop vocabulary through text, using glossary, dictionary and thesaurus	
R1H	a	Apply post-reading skills to comprehend, interpret, analyze, and evaluate text: identify and explain the relationship between the main idea and supporting details	3
	d	Apply post-reading skills to comprehend, interpret, analyze, and evaluate text: draw conclusions	
	e	Apply post-reading skills to comprehend, interpret, analyze, and evaluate text: paraphrase	
	f	Apply post-reading skills to comprehend, interpret, analyze, and evaluate text: summarize	
R1I	a	Compare, contrast, analyze and evaluate connections: text to text (information and relationships in various fiction and non-fiction works)	3
R2A		Analyze and evaluate the text features in grade-level text	2
R2B	a	Identify and explain literary techniques, in text emphasizing understatement	3
	b	Identify and explain literary techniques, in text emphasizing parallelism	
	c	Identify and explain literary techniques, in text emphasizing allusion	
	d	Identify and explain literary techniques, in text emphasizing analogy	
	e	Identify and explain literary techniques, in text emphasizing analyze and evaluate literary techniques, sensory details, figurative language and sound devices previously introduced	
R2C	a	Use details from text(s) to demonstrate comprehension skills previously introduced	3
	b	Use details from text(s) to analyze character, plot, setting, point of view	
	c	Use details from text(s) to analyze the development of a theme across genres	
	d	Use details from text(s) to identify and analyze tone	
R3A		Explain, analyze and evaluate the author's use of text features to clarify meaning	3

Strand, Big Idea, Concept	CLE Component	CLE_Description	DOK
R3B	a	Identify and explain literary techniques, in text emphasizing understatement	3
	b	Identify and explain literary techniques, in text emphasizing parallelism	
	c	Identify and explain literary techniques, in text emphasizing allusion	
	d	Identify and explain literary techniques, in text emphasizing analogy	
	e	Identify and explain literary techniques, in text emphasizing analyze and evaluate literary techniques, sensory details, figurative language and sound devices previously introduced	
R3C	a	Use details from informational and persuasive text(s) to analyze and evaluate the organizational patterns	3
	b	Use details from informational and persuasive text(s) to identify and analyze faulty reasoning and unfounded inferences	
	c	Use details from informational and persuasive text(s) to evaluate proposed solutions	
	d	Use details from informational and persuasive text(s) to evaluate for accuracy and adequacy of evidence	
	e	Use details from informational and persuasive text(s) to evaluate effect of tone on the overall meaning of work	
	f	Use details from informational and persuasive text(s) to analyze and evaluate point of view	
	g	Use details from informational and persuasive text(s) to analyze and evaluate author's viewpoint/perspective	
	h	Use details from informational and persuasive text(s) to demonstrate comprehension skills previously introduced	
W1A		Apply a writing process to write effectively in various forms and types of writing	3
W2A	a	Compose text showing awareness of audience	3
	b	Compose text choosing a form and point of view appropriate to purpose and audience	
W2B	a	Compose text with strong, controlling idea	3
	b	Compose text with relevant specific details	
	c	Compose text with complex ideas	
	d	Compose text with freshness of thought	
W2C	a	Compose text with effective beginning, middle and end	3
	b	Compose text with a logical order	
	c	Compose text with effective paragraphing	
	d	Compose text with cohesive devices	
	e	Compose text with varied sentence structure	
	f	Compose text with clarity of expression	
	g	Compose text with active voice	
W2D	a	Compose text using precise and vivid language	3
	b	Compose text using writing techniques, such as imagery, humor, voice and figurative language	
W2E	a	In written text apply conventions of capitalization	1

Strand, Big Idea, Concept	CLE Component	CLE_Description	DOK
	b	In written text apply conventions of punctuation	
	c	In written text apply standard usage	
W3A	a	Compose a variety of texts, using narrative, descriptive, expository, and/or persuasive features	4
	b	Compose a variety of texts, in various formats, including workplace communication	
	c	Compose a variety of texts, including summary	
	d	Compose a variety of texts, including literary analysis	
	e	Compose a variety of texts, including reflective writing	

English II Item Alignment to CLEs

Table A-12 provides the mean alignment rating per item based on the Overall Alignment rating scale (from '1=not aligned to any CLE' to '4=fully aligned to CLE; exemplary'). This rating serves as a confidence measure of the extent to which an item targets selected CLEs. The English II panel included five reviewers.

Table A-12. Mean Overall Alignment Rating per Item for Each English II Test Form

Form Year	Item Number	Degree of Alignment	
		M	S.D.
2009	1	3.20	0.45
2009	2	3.20	0.45
2009	3	3.20	0.45
2009	4	3.20	0.45
2009	5	2.40	1.14
2009	6	3.00	0.00
2009	7	3.40	0.55
2009	8	3.40	0.55
2009	9	3.20	0.45
2009	10	3.20	0.45
2009	11	3.20	0.45
2009	12	3.40	0.55
2009	24	3.60	0.55
2009	25	3.20	0.45
2009	26	3.40	0.55
2009	27	3.00	0.71
2009	28	3.20	0.45
2009	29	3.20	0.45
2009	30	3.20	0.45
2009	31	3.20	0.45
2009	32	3.00	0.00
2009	33	3.20	0.45
2009	34	3.20	0.45
2009	35	3.20	0.45
2009	36	3.20	0.45
2009	37	3.20	0.45
2009	38	3.20	0.45
2009	39	3.00	0.00
2009	40	3.20	0.45
2009	41	3.20	0.45
2009	43	3.00	0.00
2009	44	3.20	0.45
2009	45	3.20	0.45

Form Year	Item Number	Degree of Alignment	
		M	S.D.
2009	46	3.20	0.45
2009	47	3.20	0.45
2010	1	2.80	0.45
2010	2	3.40	0.55
2010	3	3.20	0.45
2010	4	2.80	0.45
2010	5	3.20	0.45
2010	6	3.20	0.45
2010	7	3.40	0.55
2010	8	3.40	0.55
2010	9	3.40	0.55
2010	10	3.00	0.00
2010	11	3.40	0.55
2010	12	3.00	0.00
2010	24	3.40	0.55
2010	25	3.20	0.45
2010	26	3.00	0.00
2010	27	3.20	0.45
2010	28	3.40	0.55
2010	29	3.40	0.55
2010	30	3.40	0.55
2010	31	3.20	0.45
2010	32	3.20	0.45
2010	33	3.40	0.55
2010	34	3.40	0.55
2010	35	3.20	0.45
2010	36	3.40	0.55
2010	37	3.00	0.00
2010	38	3.20	0.45
2010	39	3.40	0.55
2010	40	3.00	0.00
2010	41	3.40	0.55
2010	43	3.40	0.55
2010	44	3.40	0.55
2010	45	3.40	0.55
2010	46	3.40	0.55
2010	47	3.40	0.55
2011	1	3.40	0.55
2011	2	3.20	0.45
2011	3	3.20	0.45
2011	4	3.40	0.55

Form Year	Item Number	Degree of Alignment	
		M	S.D.
2011	5	3.00	0.00
2011	6	3.40	0.55
2011	7	3.40	0.55
2011	8	3.40	0.55
2011	9	3.40	0.55
2011	10	3.40	0.55
2011	11	3.40	0.55
2011	12	3.40	0.55
2011	24	3.40	0.55
2011	25	3.40	0.55
2011	26	3.20	0.45
2011	27	2.80	0.84
2011	28	3.20	0.84
2011	29	3.20	0.45
2011	30	3.40	0.55
2011	31	3.40	0.55
2011	32	3.40	0.55
2011	33	3.20	0.45
2011	34	1.80	0.45
2011	35	3.20	0.45
2011	36	3.40	0.55
2011	37	3.00	0.00
2011	38	3.20	0.45
2011	39	3.40	0.55
2011	40	3.00	0.00
2011	41	3.00	0.00
2011	43	3.40	0.55
2011	44	3.40	0.55
2011	45	3.40	0.55
2011	46	3.40	0.55
2011	47	3.40	0.55

English II Item Quality Ratings

Table A-13 provides mean item quality ratings based on the Overall Item Quality rating scale (from '1= poor quality' to '4=exceptional quality'). This rating provides a global judgment on the format and clarity of items. The English II panel included five reviewers.

Table A-13. Mean Overall Quality Rating per Item for Each English II Test Form

Form Year	Item Number	Degree of Alignment	
		M	S.D.
2009	1	3.00	0.00
2009	2	3.00	0.00
2009	3	3.00	0.00
2009	4	3.00	0.00
2009	5	2.40	0.89
2009	6	3.00	0.00
2009	7	3.20	0.45
2009	8	3.20	0.45
2009	9	3.00	0.00
2009	10	3.00	0.00
2009	11	3.00	0.00
2009	12	3.20	0.45
2009	24	3.40	0.55
2009	25	3.00	0.71
2009	26	3.20	0.45
2009	27	3.00	0.00
2009	28	3.00	0.00
2009	29	3.00	0.00
2009	30	3.20	0.45
2009	31	3.00	0.00
2009	32	3.00	0.00
2009	33	3.00	0.00
2009	34	3.00	0.00
2009	35	3.00	0.00
2009	36	3.00	0.00
2009	37	3.00	0.00
2009	38	3.00	0.00
2009	39	3.00	0.00
2009	40	3.00	0.00
2009	41	3.00	0.00
2009	43	3.00	0.00
2009	44	3.00	0.00
2009	45	3.00	0.00
2009	46	3.00	0.00

Form Year	Item Number	Degree of Alignment	
		M	S.D.
2009	47	3.00	0.00
2010	1	3.00	0.00
2010	2	3.00	0.00
2010	3	3.00	0.00
2010	4	3.00	0.00
2010	5	3.00	0.00
2010	6	3.00	0.00
2010	7	3.00	0.00
2010	8	3.00	0.00
2010	9	3.00	0.00
2010	10	3.00	0.00
2010	11	3.00	0.00
2010	12	3.00	0.00
2010	24	3.00	0.00
2010	25	3.00	0.00
2010	26	3.00	0.00
2010	27	3.00	0.00
2010	28	3.00	0.00
2010	29	3.00	0.00
2010	30	2.80	0.45
2010	31	3.00	0.00
2010	32	2.80	0.45
2010	33	3.00	0.00
2010	34	3.00	0.00
2010	35	3.00	0.00
2010	36	3.00	0.00
2010	37	3.00	0.00
2010	38	3.00	0.00
2010	39	3.00	0.00
2010	40	3.00	0.00
2010	41	3.00	0.00
2010	43	3.20	0.45
2010	44	3.20	0.45
2010	45	3.00	0.00
2010	46	3.20	0.45
2010	47	3.00	0.00
2011	1	3.00	0.00
2011	2	3.00	0.00
2011	3	3.00	0.00
2011	4	3.00	0.00
2011	5	3.00	0.00

Form Year	Item Number	Degree of Alignment	
		M	S.D.
2011	6	3.00	0.00
2011	7	3.00	0.00
2011	8	3.20	0.45
2011	9	3.00	0.00
2011	10	3.00	0.00
2011	11	3.00	0.00
2011	12	3.00	0.00
2011	24	3.00	0.00
2011	25	3.00	0.00
2011	26	3.00	0.00
2011	27	2.80	0.45
2011	28	2.80	0.45
2011	29	3.00	0.00
2011	30	3.00	0.00
2011	31	3.20	0.45
2011	32	3.00	0.00
2011	33	3.00	0.00
2011	34	1.80	0.45
2011	35	3.00	0.00
2011	36	3.00	0.00
2011	37	3.00	0.00
2011	38	3.00	0.00
2011	39	3.00	0.00
2011	40	3.00	0.00
2011	41	3.00	0.00
2011	43	3.20	0.45
2011	44	3.20	0.45
2011	45	3.20	0.45
2011	46	3.20	0.45
2011	47	3.20	0.45

Panelist Comments on English II Items

Tables A-14 through A-16 present panelists' comments on the individual items for the English II test forms. To maintain test security, no individual item identifiers are included.

Table A-14. Reviewer Comments on 2009 Test Form Items for English II

Test Form	Item Number	Reviewer Comment
2009	5	Alignment is too vague
2009	5	Challenges Consideration 1 on NCEO. Item refers to genre. No MO CLE addresses author's choice of genre. It appears as if item is referring to the genre of drama as a text-feature. Perhaps item was designed to reference R2A. However, this is not the best choice for this item.
2009	5	Genre question does not directly relate to CLE
2009	25	This aligns closely to 2 CLEs: R1H and R3C
2009	27	Question does not link to relationship of main ideas to support
2009	30	Good distractors
2009	48	In order to effectively answer the prompt, one must use features of both W3A and W2C. It would be better to code this item under "W" and not specify designations beyond that.
2009	48	In order to effectively compose a text (W3A) students must also fulfill the requirements of W2C.
2009	48	Since this portion of the exam is scored wholistically, then two CLEs are necessary for this item.
2009	48	The students could use a variety of texts from the first CLE to complete the writing prompt but they would achieve a better score using the criteria of the second CLE.
2009	48	The writing prompt covers more than one aspect of the standards. I feel that the two CLE codes have to be used together to fully cover what is being required in the prompt.

Table A-15. Reviewer Comments on 2010 Test Form Items for English II

Test Form	Item Number	Reviewer Comment
2010	1	Author's purpose is not clearly spelled out in any fiction CLEs
2010	4	The CLE R2C asks for "theme across genre" although this is only one text, it's the best fit.
2010	9	There might be a better example to use
2010	12	If the intent of this item is to measure "accuracy and adequacy of evidence, it is a poor match. If it intends to measure supporting ideas it's not a perfect match, but an adequate one.
2010	26	The CLE R2C asks for "theme across genre" although this is only one text, it's the best fit.
2010	30	It's a simple definition; not very challenging item
2010	32	MLA style currently advocates use of italics, not quotation marks, for emphasis.
2010	41	The CLE R2C asks for "theme across genre" although this is only one text, it's the best fit.

Table A-16. Reviewer Comments on 2011 Test Form Items for English II

Test Form	Item Number	Reviewer Comment
2011	5	most likely is worded poorly
2011	8	This item does a good job of asking students to evaluate adequacy of evidence and present a solution.
2011	27	It's difficult to identify a clear CLE since this question calls for simple identification of an item in the essay.
2011	27	Not aligned closely with a standard. An argument can be made for either R1H or R3C. R1H identifies main idea and details, which this question loosely reflects. R3C covers basic comprehension of non-fiction works.
2011	28	The question really fit into any one category
2011	34	confusing question; appears to be seeking an analogy, but correct choice is so boring that most higher thinking students wouldn't choose it as the correct response
2011	34	The item asks students to create their own analogy; this is not reflected in the CLEs.
2011	34	This is poorly written and students almost have to create their own analogy. This would fit better as a science or writing question.
2011	34	This item asks students to create an analogy from the text. The CLEs for the state of MO do not address this expectation. Perhaps the item was designed to be coded as R3B. It is, however, a poorly constructed question to assess the CLE.
2011	34	This question appears to more of a reasoning question asking the student to create their own analogy. It does not closely align to the CLEs for English II concerned with identifying and analyzing analogies in literature.
2011	38	The CLE R2C asks for "theme across genre" although this is only one text, it's the best fit.

English II CLEs Matched to Items

Table A-17 displays the English II CLEs matched to items (by sequential item number) per test form by reviewers.

Table A-17. English II Items Matched to CLEs by Test Form Year

Form Year	CLE	Item Number
2009	R1E	7
2009	R1E	9
2009	R1E	24
2009	R1E	29
2009	R1E	36
2009	R1E	43
2009	R1H	4
2009	R1H	8
2009	R1H	11
2009	R1H	25
2009	R1H	27
2009	R1H	31
2009	R1H	33
2009	R1H	37
2009	R1H	40
2009	R1I	34
2009	R1I	35
2009	R2A	3
2009	R2A	5
2009	R2B	2
2009	R2B	6
2009	R2B	9
2009	R2B	38
2009	R2B	39
2009	R2B	41
2009	R2C	1
2009	R2C	3
2009	R2C	4
2009	R2C	8
2009	R2C	30
2009	R2C	37
2009	R2C	39
2009	R3A	6
2009	R3A	32
2009	R3B	9

Form Year	CLE	Item Number
2009	R3B	10
2009	R3B	26
2009	R3B	28
2009	R3B	30
2009	R3C	12
2009	R3C	25
2009	R3C	26
2009	R3C	27
2009	R3C	33
2009	R3C	35
2009	W2E	43
2009	W2E	44
2009	W2E	45
2009	W2E	46
2009	W2E	47
2010	R1E	1
2010	R1E	7
2010	R1E	24
2010	R1E	30
2010	R1E	36
2010	R1H	1
2010	R1H	2
2010	R1H	3
2010	R1H	4
2010	R1H	5
2010	R1H	6
2010	R1H	12
2010	R1H	29
2010	R1H	31
2010	R1H	33
2010	R1H	35
2010	R1H	37
2010	R1H	38
2010	R1H	39
2010	R1H	40
2010	R2A	32
2010	R2B	25
2010	R2B	26
2010	R2B	27
2010	R2B	28
2010	R2B	40

Form Year	CLE	Item Number
2010	R2B	41
2010	R2C	4
2010	R2C	5
2010	R2C	26
2010	R2C	28
2010	R2C	29
2010	R2C	34
2010	R2C	36
2010	R2C	37
2010	R2C	38
2010	R2C	39
2010	R2C	41
2010	R3A	8
2010	R3A	32
2010	R3B	9
2010	R3B	10
2010	R3C	8
2010	R3C	11
2010	R3C	12
2010	R3C	31
2010	R3C	34
2010	R3C	35
2010	W2E	43
2010	W2E	44
2010	W2E	45
2010	W2E	46
2010	W2E	47
2011	R1E	1
2011	R1E	7
2011	R1E	11
2011	R1E	24
2011	R1E	36
2011	R1H	9
2011	R1H	11
2011	R1H	12
2011	R1H	25
2011	R1H	27
2011	R1H	29
2011	R1H	30
2011	R1H	32
2011	R1H	33

Form Year	CLE	Item Number
2011	R1H	34
2011	R1H	35
2011	R2B	6
2011	R2B	37
2011	R2B	38
2011	R2B	40
2011	R2B	41
2011	R2C	4
2011	R2C	5
2011	R2C	31
2011	R2C	38
2011	R2C	39
2011	R3A	2
2011	R3A	5
2011	R3A	8
2011	R3A	12
2011	R3A	31
2011	R3B	3
2011	R3B	6
2011	R3B	10
2011	R3B	26
2011	R3B	28
2011	R3B	34
2011	R3C	2
2011	R3C	4
2011	R3C	5
2011	R3C	8
2011	R3C	9
2011	R3C	25
2011	R3C	27
2011	R3C	33
2011	R3C	35
2011	R3C	39
2011	W2E	43
2011	W2E	44
2011	W2E	45
2011	W2E	46
2011	W2E	47

Appendix B
EOC Algebra I: Detailed Statistical Results

In Appendix B, we present the full alignment results for Algebra I. These results include (a) the four Webb measures, (b) consensus DOK ratings by CLE, (c) item alignment and quality ratings, (d) summary reviewer comments, and (e) items matched to course-level expectations (CLEs). All results are reported at the level of the content Strand.

For each analysis, we display the results first for the 2009 test form conducted on *all* operational items (multiple-choice and performance events). We then present results of analyses on the three test forms (2009 included) with only the multiple-choice items[2].

Webb Alignment Indicators

The following tables include complete statistical results on the four Webb alignment indicators: Categorical Concurrence, Depth-of-Knowledge (DOK) Consistency, Range of Knowledge, and Balance of Knowledge.

Categorical Concurrence

Tables B-1 and B-2 include categorical concurrence results: the mean number of items matched to strand by panelists, the standard deviation (S.D.) among panelists' ratings, and the final alignment conclusion (Yes or No). The criterion for acceptable Categorical Concurrence is a minimum of six items per strand.

Table B-1. Categorical Concurrence for Algebra I 2009 Test Form: Mean Number of Multiple-Choice and Performance Event Items per Strand

Strand	2009 Test Form		
	Mean Items per Strand	S.D.	At Least Six Items per Strand
Number and Operations	9.25	0.96	Yes
Algebraic Relationships	20.25	0.50	Yes
Data and Probability	7.50	1.00	Yes
Strands Matched to Six or More Items			3 of 3

[2] As a reminder to the reader, reviewers only rated performance events for the 2009 test forms.

Table B-2. Categorical Concurrence for Algebra I 2009, 2010, and 2011 Test Forms: Mean Number Multiple-Choice Items Only per Strand

Strand	2009 Test Form			2010 Test Form			2011 Test Form		
	Mean Items per Strand	S.D.	At Least Six Items per Strand	Mean Items per Strand	S.D.	At Least Six Items per Strand	Mean Items per Strand	S.D.	At Least Six Items per Strand
Number and Operations	9.25	0.96	Yes	8.75	0.50	Yes	9.50	0.58	Yes
Algebraic Relationships	18.25	0.50	Yes	18.00	0.00	Yes	19.25	0.50	Yes
Data and Probability	7.50	1.00	Yes	8.25	0.50	Yes	6.25	0.50	Yes
Strands Matched to Six or More Items			3 of 3			3 of 3			3 of 3

Depth-of-Knowledge Consistency

Tables B-3 through B-10 present results of a comparative analysis between assessment items and CLEs on depth-of-knowledge (DOK). Tables B-3 through B-6 focus on the test item DOK relative to the corresponding CLEs. Specifically, these tables include the mean percentage of items per Strand rated below, at the same level, or above the DOK of the corresponding CLE. Webbs' criterion for acceptable DOK consistency is that item DOK must be At or Above the DOK level of the matched standard for at least 50% of items. Across the CLEs per Strand, we note (Yes or No) whether 50% of total items assessed CLEs as the appropriate cognitive level. Note that the Webb method compares item DOK values to the consensus DOK values determined by reviewers, which may differ from the State published DOK levels per CLE in some cases.

Table B-3. DOK Consistency for Algebra I 2009 Test Form: Mean Percentage of Multiple-Choice Items and Performance Event Items Below, At, or Above Corresponding CLEs

Strand	2009 Test Form									
	[a] Mean Items per Strand	[b] Percent Items Below		[c] Percent Items At		[d] Percent Items Above		[e] Percent Items At/Above DOK of CLE		[f] 50% or More Items At/Above DOK of CLE
		M	S.D.	M	S.D.	M	S.D.	M	S.D.	
Number and Operations	9.25	15%	0.13	76%	0.09	9%	0.12	85%	0.13	Yes
Algebraic Relationships	20.25	22%	0.1	73%	0.15	5%	0.06	78%	0.1	Yes
Data and Probability	7.5	42%	0.11	58%	0.11	0	0	58%	0.11	Yes
Strands with CLEs Assessed Appropriately										3 of 3

[a] Mean Items per Big Idea = Mean number of items matched to strand (Categorical Concurrence)
[b] Items Below = Percentage of items below DOK level of CLEs per strand.
[c] Items Same = Percentage of items with same DOK level as CLEs per strand.
[d] Items Above = Percentage of items above DOK level of CLEs per strand.
[e] Items At and Above = Percentage of items, when added, with DOK at the Same level and Above level of CLE.
[f] 50% or More Items = At least half of items assessing strand matched DOK level of corresponding CLEs.

Table B-4. DOK Consistency for Algebra I 2009 Test Form: Mean Percentage of Multiple-Choice Items Below, At, or Above Corresponding CLEs

Strand	2009 Test Form									
	[a] Mean Items per Strand	[b] Items Below		[c] Items Same		[d] Items Above		[e] Items At/Above DOK Level of CLE		[f] 50% or More Items
		M	S.D.	M	S.D.	M	S.D.	M	S.D.	
Number and Operations	9.25	15%	0.13	76%	0.09	9%	0.12	85%	0.13	Yes
Algebraic Relationships	18.25	25%	0.11	75%	0.11	0	0	75%	0.11	Yes
Data and Probability	7.50	42%	0.11	58%	0.11	0	0	58%	0.11	Yes
Strands with CLEs Assessed Appropriately										3 of 3

[a] Mean Items per Big Idea = Mean number of items matched to strand (Categorical Concurrence)
[b] Items Below = Percentage of items below DOK level of CLEs per strand.
[c] Items Same = Percentage of items with same DOK level as CLEs per strand.
[d] Items Above = Percentage of items above DOK level of CLEs per strand.
[e] Items At and Above = Percentage of items, when added, with DOK at the Same level and Above level of CLE.
[f] 50% or More Items = At least half of items assessing strand matched DOK level of corresponding CLEs.

Table B-5. DOK Consistency for Algebra I 2010 Test Form: Mean Percentage of Multiple-Choice Items Below, At, or Above Corresponding CLEs

Strand	2010 Test Form									
	Mean Items per Strand	Items Below		Items Same		Items Above		[e] Items At/Above DOK Level of CLE		[f] 50% or More Items
		M	S.D.	M	S.D.	M	S.D.	M	S.D.	
Number and Operations	8.75	6%	0.06	83%	0.15	12%	0.14	94%	0.06	Yes
Algebraic Relationships	18.00	12%	0.05	88%	0.05	0	0	88%	0.05	Yes
Data and Probability	8.25	61%	0.11	39%	0.11	0	0	39%	0.11	No
Strands with CLEs Assessed Appropriately										2 of 3

[a] Mean Items per Big Idea = Mean number of items matched to strand (Categorical Concurrence)
[b] Items Below = Percentage of items below DOK level of CLEs per strand.
[c] Items Same = Percentage of items with same DOK level as CLEs per strand.
[d] Items Above = Percentage of items above DOK level of CLEs per strand.
[e] Items At and Above = Percentage of items, when added, with DOK at the Same level and Above level of CLE.
[f] 50% or More Items = At least half of items assessing strand matched DOK level of corresponding CLEs.

Table B-6. DOK Consistency for Algebra I 2011 Test Form: Mean Percentage of Multiple-Choice Items Below, At, or Above Corresponding CLEs

Strand	2011 Test Form									
	Mean Items per Strand	Items Below		Items Same		Items Above		[e] Items At/Above DOK Level of CLE		[f] 50% or More Items
		M	S.D.	M	S.D.	M	S.D.	M	S.D.	
Number and Operations	9.50	29%	0.11	66%	0.10	5%	0.10	71%	0.11	Yes
Algebraic Relationships	19.25	26%	0.10	74%	0.10	0	0	74%	0.10	Yes
Data and Probability	6.25	77%	0.19	23%	0.19	0	0	23%	0.19	No
Strands with CLEs Assessed Appropriately										2 of 3

[a] Mean Items per Big Idea = Mean number of items matched to strand (Categorical Concurrence)
[b] Items Below = Percentage of items below DOK level of CLEs per strand.
[c] Items Same = Percentage of items with same DOK level as CLEs per strand.
[d] Items Above = Percentage of items above DOK level of CLEs per strand.
[e] Items At and Above = Percentage of items, when added, with DOK at the Same level and Above level of CLE.
[f] 50% or More Items = At least half of items assessing strand matched DOK level of corresponding CLEs.

Tables B-7 and B-10 summarize the same data in a different way by focusing on the percentage of CLEs assessed appropriately. These tables display the mean percentage of *standards* (CLEs) per Strand assessed at the appropriate DOK level (item DOK and standard DOK are the same), as well as the number of standards assessed below and above the level expected. At least 50% of items must be At or Above the DOK level of the corresponding CLE in order for the assessment of that CLE to be judged minimally appropriate.

Table B-7. DOK Consistency for Algebra I 2009 Test Form: Mean Percentage of CLEs Assessed by Multiple-Choice Items and Performance Events Below, At, or Above Expected DOK Level

Strand	[a] Number of CLEs	2009 Test Form								
		[b] Percent CLEs Assessed Below DOK		[c] Percent CLEs Assessed At DOK		[d] Percent CLEs Assessed Above DOK		[e] CLEs Assessed At/Above DOK Expected		[f] 50% or More CLEs
		M	S.D.	M	S.D.	M	S.D.	M	S.D.	
Number and Operations	2	13%	0.25	88%	0.25	0	0	88%	0.25	Yes
Algebraic Relationships	10	21%	0.13	75%	0.15	5%	0.06	79%	0.13	Yes
Data and Probability	5	29%	0.08	71%	0.08	0	0	71%	0.08	Yes
Strands with CLEs Assessed Appropriately										3 of 3

[a] Number of CLEs = Per Course Level Expectations 2.0, the number of CLEs per Strand.
[b] CLEs Assessed Below DOK = Percentage of CLEs per Strand assessed below consensus DOK level.
[c] CLEs Assessed At DOK = Percentage of CLEs per Strand assessed at the consensus DOK level.
[d] CLEs Assessed Above DOK = Percentage of CLEs per Strand assessed above consensus DOK level.
[e] CLEs Assessed At/Above DOK = Combined percentage of CLEs per Strand assessed at or above consensus DOK level.
[f] 50% or More CLEs = At least half of CLEs per strand were assessed by items at the appropriate DOK level.

Table B-8. DOK Consistency for Algebra I 2009 Test Form: Mean Percentage of CLEs Assessed by Multiple-Choice Items Below, At, or Above Expected DOK Level

		2009 Test Form								
Strand	[a] Number of CLEs	[b] Percent CLEs Assessed Below DOK		[c] Percent CLEs Assessed At DOK		[d] Percent CLEs Assessed Above DOK		[e] CLEs Assessed At/Above DOK Expected		[f] 50% or More CLEs
		M	S.D.	M	S.D.	M	S.D.	M	S.D.	
Number and Operations	2	13%	0.25	88%	0.25	0	0	88%	0.25	Yes
Algebraic Relationships	10	22%	0.14	78%	0.14	0	0	78%	0.14	Yes
Data and Probability	5	29%	0.08	71%	0.08	0	0	71%	0.08	Yes
Strands with CLEs Assessed Appropriately										3 of 3

[a] Number of CLEs = Per Course Level Expectations 2.0, the number of CLEs per Strand.
[b] CLEs Assessed Below DOK = Percentage of CLEs per Strand assessed below consensus DOK level.
[c] CLEs Assessed At DOK = Percentage of CLEs per Strand assessed at the consensus DOK level.
[d] CLEs Assessed Above DOK = Percentage of CLEs per Strand assessed above consensus DOK level.
[e] CLEs Assessed At/Above DOK = Combined percentage of CLEs per Strand assessed at or above consensus DOK level.
[f] 50% or More CLEs = At least half of CLEs per strand were assessed by items at the appropriate DOK level.

Table B-9. DOK Consistency for Algebra I 2010 Test Form: Mean Percentage of CLEs Assessed by Multiple-Choice Items Below, At, or Above Expected DOK Level

Strand	[a] Number of CLEs	2010 Test Form								
		[b] Percent CLEs Assessed Below DOK		[c] Percent CLEs Assessed At DOK		[d] Percent CLEs Assessed Above DOK		[e] CLEs Assessed At/Above DOK Expected		[f] 50% or More CLEs
		M	S.D.	M	S.D.	M	S.D.	M	S.D.	
Number and Operations	2	0	0	83%	0.19	0.17	0.19	100%	0.00	Yes
Algebraic Relationships	10	16%	0.05	84%	0.05	0	0	84%	0.05	Yes
Data and Probability	5	56%	0.13	44%	0.13	0	0	44%	0.13	No
Strands with CLEs Assessed Appropriately										2 of 3

[a] Number of CLEs = Per Course Level Expectations 2.0, the number of CLEs per Strand.
[b] CLEs Assessed Below DOK = Percentage of CLEs per Strand assessed below consensus DOK level.
[c] CLEs Assessed At DOK = Percentage of CLEs per Strand assessed at the consensus DOK level.
[d] CLEs Assessed Above DOK = Percentage of CLEs per Strand assessed above consensus DOK level.
[e] CLEs Assessed At/Above DOK = Combined percentage of CLEs per Strand assessed at or above consensus DOK level.
[f] 50% or More CLEs = At least half of CLEs per strand were assessed by items at the appropriate DOK level.

Table B-10. DOK Consistency for Algebra I 2011 Test Form: Mean Percentage of CLEs Assessed by Multiple-Choice Items Below, At, or Above Expected DOK Level

Strand	[a] Number of CLEs	2011 Test Form								
		[b] Percent CLEs Assessed Below DOK		[c] Percent CLEs Assessed At DOK		[d] Percent CLEs Assessed Above DOK		[e] CLEs Assessed At/Above DOK Expected		[f] 50% or More CLEs
		M	S.D.	M	S.D.	M	S.D.	M	S.D.	
Number and Operations	2	13%	0.25	75%	0.29	13%	0.25	88%	0.25	Yes
Algebraic Relationships	10	29%	0.05	71%	0.05	0	0	71%	0.05	Yes
Data and Probability	5	75%	0.19	25%	0.19	0	0	25%	0.19	No
Strands with CLEs Assessed Appropriately										2 of 3

[a] Number of CLEs = Per Course Level Expectations 2.0, the number of CLEs per Strand.
[b] CLEs Assessed Below DOK = Percentage of CLEs per Strand assessed below consensus DOK level.
[c] CLEs Assessed At DOK = Percentage of CLEs per Strand assessed at the consensus DOK level.
[d] CLEs Assessed Above DOK = Percentage of CLEs per Strand assessed above consensus DOK level.
[e] CLEs Assessed At/Above DOK = Combined percentage of CLEs per Strand assessed at or above consensus DOK level.
[f] 50% or More CLEs = At least half of CLEs per strand were assessed by items at the appropriate DOK level.

Range-of-Knowledge Correspondence

Tables B-7 and B-8 present the range-of-knowledge correspondence results. These tables include the mean number (and percentage) of CLEs matched to at least one item per Strand. For acceptable range, a minimum of 50% of CLEs within each Strand should be matched to at least one item.

Table B-7. Range-of-Knowledge for Algebra I 2009 Test Form: Mean CLEs per Strand Linked with Multiple Choice and Performance Items

Strand	Number of CLEs	2009 Test Form					
		Mean Items per Strand	Number of CLEs Assessed		% CLEs Assessed		50% or More CLEs
			M	S.D.	M	S.D.	
Number and Operations	2	9.25	2.00	0.00	100%	0.00	Yes
Algebraic Relationships	10	20.25	9.00	0.82	90%	0.08	Yes
Data and Probability	5	7.50	4.25	0.50	85%	0.10	Yes
Strands with CLEs Assessed by At Least One Item							3 of 3

Table B-8. Range-of-Knowledge for Algebra I 2009, 2010, and 2011 Test Forms: Mean CLEs per Strand Linked with Multiple Choice Items Only

Strand	Number of CLEs	2009 Test Form						2010 Test Form						2011 Test Form					
		Mean Items per Strand	Number of CLEs Assessed		% CLEs Assessed		50% or More CLEs	Mean Items per Strand	Number of CLEs Assessed		% CLEs Assessed		50% or More CLEs	Mean Items per Strand	Number of CLEs Assessed		% CLEs Assessed		50% or More CLEs
			M	S.D.	M	S.D.			M	S.D.	M	S.D.			M	S.D.	M	S.D.	
Number and Operations	2	9.25	2.00	0.00	100%	0.00	Yes	8.75	2.00	0.00	100%	0.00	Yes	9.50	2.00	0.00	100%	0.00	Yes
Algebraic Relationships	10	18.25	9.00	0.82	90%	0.08	Yes	18.00	9.50	0.58	95%	0.06	Yes	19.25	7.75	0.50	78%	0.05	Yes
Data and Probability	5	7.50	4.25	0.50	85%	0.10	Yes	8.25	4.50	0.58	90%	0.12	Yes	6.25	4.25	0.50	85%	0.10	Yes
Strands with CLEs Assessed by At Least One Item							3 of 3						3 of 3						3 of 3

Balance-of-Knowledge Representation

Tables B-9 and B-10 display the balance indices for each Algebra I strand. This index is based on the mean number of items matched to each CLE. The minimum acceptable balance index is 70 out of 100. The table also includes the percentage of items linked to each Strand per strand.

Table B-9. Balance-of-Knowledge Representation for Algebra I 2009 Test Form: Balance Index per Strand

Strand	CLEs per Strand	2009 Test Form					
		Mean CLEs Linked with Items	Mean Items per Strand	% Items Linked to Strand		Balance Index	Balance Index Target Met
				M	S.D.		
Number and Operations	2	2.00	9.25	26%	0.39	0.98	Yes
Algebraic Relationships	10	9.00	20.25	54%	0.20	0.96	Yes
Data and Probability	5	4.25	7.50	21%	0.41	0.81	Yes
Balance Index Met							3 of 3

Table B-10. Balance-of-Knowledge Representation for Algebra I 2009, 2010, and 2011 Test Forms: Balance Index per Strand

Strand	CLEs per Strand	2009 Test Form						2010 Test Form						2011 Test Form					
		Mean CLEs Linked with Items	Mean Items per Strand	% Items Linked to Strand		Balance Index	Balance Index Target Met	Mean CLEs Linked with Items	Mean Items per Strand	% Items Linked to Strand		Balance Index	Balance Index Target Met	Mean CLEs Linked with Items	Mean Items per Strand	% Items Linked to Strand		Balance Index	Balance Index Target Met
				M	S.D.					M	S.D.					M	S.D.		
Number and Operations	2	2.00	9.25	26%	0.39	0.98	Yes	2.00	8.75	25%	0.50	0.74	Yes	2.00	9.50	27%	0.37	0.82	Yes
Algebraic Relationships	10	9.00	18.25	52%	0.20	0.98	Yes	9.50	18.00	51%	0.00	0.75	Yes	7.75	19.25	55%	0.32	0.74	Yes
Data and Probability	5	4.25	7.50	21%	0.41	0.81	Yes	4.50	8.25	24%	0.50	0.81	Yes	4.25	6.25	18%	0.32	0.83	Yes
Balance Index Met							3 of 3						3 of 3						3 of 3

Consensus DOK Ratings on Algebra I CLEs

Table B-11 presents DOK ratings established through group consensus for each Algebra I CLE. Column 1 lists the Strand letter, Big Idea number, and Concept letter, while Column 2 displays the CLE content description. Column 3 indicates the DOK rating assigned to the CLE by the group. The titles corresponding with the Strand letters are as follows: N = Numbers and Operations, A = Algebraic Relationships, and D = Data and Probability.

Table B-11. Algebra I: Group Consensus Ratings on DOK Level per CLE

Strand, Big Idea, Concept	CLE Description	DOK Rating
N1A	Compare and order rational and irrational numbers, including finding their approximate locations on a number line	1
N1B	Use real numbers and various models, drawing, etc. to solve problems	2
A1B	Generalize patterns using explicitly or recursively defined functions	2
A1C	Compare and contrast various forms of representations of patterns	2
A1D	Understand and compare the properties of linear and nonlinear functions	2
A1E	Describe the effects of parameter changes on linear, exponential growth/decay and quadratic functions including intercepts	2
A2A	Use symbolic algebra to represent and solve problems that involve linear and quadratic relationships including equations and inequalities	2
A2B	Describe and use algebraic manipulations, including factoring and rules of integer exponents and apply properties of exponents (including order of operations) to simplify expressions	2
A2C	Use and solve equivalent forms of equations (linear, absolute value and quadratic)	2
A2D	Use and solve systems of linear equations or inequalities with 2 variables	2
A3A	Identify quantitative relationships and determine the type(s) of functions that might model the situation to solve the problem	3
A4A	Analyze linear and quadratic functions by investigating rates of change, intercepts and zeros	3
D1A	Formulate questions and collect data about a characteristic which include sample spaces and distributions	3
D1C	Select and use appropriate graphical representation of data and given one-variable quantitative data, display the distribution and describe its shape	2
D2A	Apply statistical measures of center to solve problems	2
D2C	Given a scatterplot, determine an equation for a line of best fit	2
D3A	Make conjectures about possible relationships between 2 characteristics of a sample on the basis of scatter plots of the data	3

Algebra I Item Alignment to CLEs

Table B-12 provides the mean alignment rating per item based on the Overall Alignment rating scale (from '1=not aligned to any CLE' to '4=fully aligned to CLE; exemplary'). This rating serves as a confidence measure of the extent to which an item targets selected CLEs. The Algebra I panel included four reviewers.

Table B-12. Mean Overall Alignment Rating per Item for Each Algebra I Test Form

Form Year	Item Number	Degree of Alignment	
		M	S.D.
2009	1	4.00	0.00
2009	2	4.00	0.00
2009	3	2.50	1.00
2009	4	4.00	0.00
2009	5	3.75	0.50
2009	10	3.00	0.00
2009	11	4.00	0.00
2009	12	3.75	0.50
2009	13	4.00	0.00
2009	14	3.75	0.50
2009	15	3.00	0.00
2009	16	3.25	0.50
2009	17	4.00	0.00
2009	18	3.00	0.00
2009	19	4.00	0.00
2009	20	3.75	0.50
2009	21	3.25	0.50
2009	26	3.25	0.96
2009	27	3.25	0.50
2009	28	2.75	0.96
2009	29	3.75	0.50
2009	30	4.00	0.00
2009	31	3.75	0.50
2009	32	3.75	0.50
2009	33	2.50	0.58
2009	34	3.25	0.96
2009	35	4.00	0.00
2009	36	4.00	0.00
2009	37	3.75	0.50
2009	38	4.00	0.00
2009	43	4.00	0.00
2009	44	3.75	0.50

Form Year	Item Number	Degree of Alignment	
		M	S.D.
2009	45	4.00	0.00
2009	46	4.00	0.00
2009	47	4.00	0.00
2010	1	4.00	0.00
2010	2	4.00	0.00
2010	3	4.00	0.00
2010	4	4.00	0.00
2010	5	3.75	0.50
2010	10	4.00	0.00
2010	11	4.00	0.00
2010	12	4.00	0.00
2010	13	4.00	0.00
2010	14	4.00	0.00
2010	15	4.00	0.00
2010	16	3.00	1.15
2010	17	3.50	0.58
2010	18	3.75	0.50
2010	19	4.00	0.00
2010	20	4.00	0.00
2010	21	4.00	0.00
2010	26	4.00	0.00
2010	27	4.00	0.00
2010	28	4.00	0.00
2010	29	4.00	0.00
2010	30	4.00	0.00
2010	31	4.00	0.00
2010	32	3.50	1.00
2010	33	4.00	0.00
2010	34	4.00	0.00
2010	35	4.00	0.00
2010	36	2.25	0.50
2010	37	3.75	0.50
2010	38	4.00	0.00
2010	43	4.00	0.00
2010	44	4.00	0.00
2010	45	3.75	0.50
2010	46	4.00	0.00
2010	47	4.00	0.00
2011	1	4.00	0.00
2011	2	4.00	0.00

Form Year	Item Number	Degree of Alignment	
		M	S.D.
2011	3	3.25	0.50
2011	4	4.00	0.00
2011	5	4.00	0.00
2011	10	3.75	0.50
2011	11	3.75	0.50
2011	12	3.75	0.50
2011	13	4.00	0.00
2011	14	4.00	0.00
2011	15	4.00	0.00
2011	16	2.00	0.00
2011	17	4.00	0.00
2011	18	3.75	0.50
2011	19	4.00	0.00
2011	20	4.00	0.00
2011	21	4.00	0.00
2011	26	4.00	0.00
2011	27	4.00	0.00
2011	28	3.75	0.50
2011	29	2.50	0.58
2011	30	4.00	0.00
2011	31	4.00	0.00
2011	32	4.00	0.00
2011	33	4.00	0.00
2011	34	3.75	0.50
2011	35	4.00	0.00
2011	36	4.00	0.00
2011	37	4.00	0.00
2011	38	4.00	0.00
2011	43	4.00	0.00
2011	44	4.00	0.00
2011	45	4.00	0.00
2011	46	4.00	0.00
2011	47	4.00	0.00

Algebra I Item Quality Ratings

Table B-13 provides mean item quality ratings based on the Overall Item Quality rating scale (from '1= poor quality' to '4=exceptional quality'). This rating provides a global judgment on the format and clarity of items. The Algebra I panel included four reviewers.

Table B-13. Mean Overall Quality Rating per Item for Each Algebra I Test Form

Form Year	Item Number	Degree of Alignment	
		M	S.D.
2009	1	4.00	0.00
2009	2	4.00	0.00
2009	3	3.75	0.50
2009	4	4.00	0.00
2009	5	4.00	0.00
2009	10	3.75	0.50
2009	11	4.00	0.00
2009	12	3.75	0.50
2009	13	4.00	0.00
2009	14	3.50	1.00
2009	15	4.00	0.00
2009	16	2.75	0.96
2009	17	4.00	0.00
2009	18	3.50	0.58
2009	19	3.75	0.50
2009	20	3.50	0.58
2009	21	3.75	0.50
2009	26	3.75	0.50
2009	27	3.75	0.50
2009	28	2.50	0.58
2009	29	4.00	0.00
2009	30	4.00	0.00
2009	31	3.75	0.50
2009	32	4.00	0.00
2009	33	2.25	1.26
2009	34	4.00	0.00
2009	35	4.00	0.00
2009	36	4.00	0.00
2009	37	4.00	0.00
2009	38	4.00	0.00
2009	43	4.00	0.00
2009	44	4.00	0.00
2009	45	4.00	0.00

Form Year	Item Number	Degree of Alignment	
		M	S.D.
2009	46	4.00	0.00
2009	47	4.00	0.00
2010	1	2.50	0.58
2010	2	4.00	0.00
2010	3	4.00	0.00
2010	4	4.00	0.00
2010	5	4.00	0.00
2010	10	4.00	0.00
2010	11	4.00	0.00
2010	12	4.00	0.00
2010	13	4.00	0.00
2010	14	4.00	0.00
2010	15	4.00	0.00
2010	16	2.00	0.00
2010	17	4.00	0.00
2010	18	3.50	1.00
2010	19	4.00	0.00
2010	20	4.00	0.00
2010	21	4.00	0.00
2010	26	4.00	0.00
2010	27	4.00	0.00
2010	28	4.00	0.00
2010	29	4.00	0.00
2010	30	4.00	0.00
2010	31	4.00	0.00
2010	32	4.00	0.00
2010	33	4.00	0.00
2010	34	3.00	1.15
2010	35	4.00	0.00
2010	36	2.00	0.00
2010	37	4.00	0.00
2010	38	3.50	1.00
2010	43	4.00	0.00
2010	44	4.00	0.00
2010	45	4.00	0.00
2010	46	4.00	0.00
2010	47	4.00	0.00
2011	1	4.00	0.00
2011	2	4.00	0.00
2011	3	4.00	0.00
2011	4	4.00	0.00

Form Year	Item Number	Degree of Alignment	
		M	S.D.
2011	5	4.00	0.00
2011	10	4.00	0.00
2011	11	4.00	0.00
2011	12	4.00	0.00
2011	13	4.00	0.00
2011	14	4.00	0.00
2011	15	4.00	0.00
2011	16	1.50	0.58
2011	17	3.50	0.58
2011	18	4.00	0.00
2011	19	4.00	0.00
2011	20	4.00	0.00
2011	21	3.75	0.50
2011	26	4.00	0.00
2011	27	4.00	0.00
2011	28	4.00	0.00
2011	29	1.50	0.58
2011	30	4.00	0.00
2011	31	4.00	0.00
2011	32	4.00	0.00
2011	33	4.00	0.00
2011	34	3.75	0.50
2011	35	4.00	0.00
2011	36	4.00	0.00
2011	37	4.00	0.00
2011	38	4.00	0.00
2011	43	4.00	0.00
2011	44	4.00	0.00
2011	45	4.00	0.00
2011	46	3.50	1.00
2011	47	4.00	0.00

Panelist Comments on Algebra I Items

Tables B-14 through B-16 present panelists' comments on the individual items for the Algebra I test forms. To maintain test security, no individual item identifiers are included.

Table B-14. Reviewer Comments on 2009 Test Form Items for Algebra I

Test Form	Item Number	Reviewer Comment
2009	3	I think this question could fall equally under both categories
2009	3	If a student missed this question it would be difficult to say which concept they did not understand.
2009	3	WHAT IS THE DIFFERENCE BETWEEN A1E AND A4A
2009	14	I do not agree that students must know that a number is irrational in order to work with numbers
2009	16	I do not agree that this is a fair question to be asked of Algebra I students. It is not an Algebra I concept.
2009	16	This is not an Alg 1 Concept.
2009	28	I do not agree that this is a fair question to be asked of Algebra I students. It is not an Algebra I concept.
2009	28	This is not an Alg 1 Concept.
2009	28	This is not an algebraic problem
2009	33	I DON'T UNDERSTAND THE WORDING OF "TERM" AND "TERM NUMBER"
2009	33	I feel like this question is a stretch to this CLE. The variance in the choices is a slight technical change and would confuse students. I'm not completely sure this is testing what the CLE is written to assess.
2009	33	Poor question. Term vs Term number
2009	33	The distractors are unfair. C and D are testing whether or not the student reads term or term number correctly.
2009	34	I am unsure regarding the CLE alignment for this question. I'm also wondering if it could be A3A

Table B-15. Reviewer Comments on 2010 Test Form Items for Algebra I

Test Form	Item Number	Reviewer Comment
2010	1	Difficult to do online
2010	1	Difficult to do online. Easier to do with booklet.
2010	1	I'm concerned that students taking the test online will have an extremely difficult time unless the graph is live and students can plot the points on the graph.
2010	1	Very difficult to do online. Students taking it online can make simply mistakes
2010	16	Not a traditionally covered in Algebra 1

Test Form	Item Number	Reviewer Comment
2010	16	Not an algebraic topic
2010	16	Not specifically an Algebra I concept.
2010	16	this concept is not in the algebra CLE's
2010	18	Not specifically an Algebra I concept.
2010	32	Unsure which standard this question is testing
2010	34	Not an algebraic topic
2010	34	Not specifically an Algebra I concept.
2010	36	Hard to place and not an algebra I topic!
2010	36	Not a traditionally covered in Algebra 1
2010	36	This is not an algebra CLE
2010	36	Trouble deciding if this should be coded in N1B or D1A - not sure of the intent of the question. Also, not specifically an Algebra I concept.
2010	38	Not specifically an Algebra I concept.

Table B-16. Reviewer Comments on 2011 Test Form Items for Algebra I

Test Form	Item Number	Reviewer Comment
2011	10	Not specifically an Algebra I concept
2011	16	Not an Algebra 1 concept!!!!!!
2011	16	Problem is too difficult for Algebra I students. Not specifically an Algebra I concept.
2011	16	This is not taught in Algebra 1 and doesn't clearly fit into any CLEs
2011	16	This problem is not an algebra 1 concept
2011	17	Difficult for students to read and comprehend
2011	21	Difficult for students to read and comprehend the answer choices
2011	29	NOT an Algebra 1 concept!!!!!!!!
2011	29	This is not an algebra 1 concept
2011	29	This is not an Algebra I concept and does not clearly fit into an Algebra I CLE
2011	29	This is not taught in Algebra 1 and doesn't clearly fit into any CLEs
2011	34	Difficult for students to read and comprehend

Algebra I CLEs Matched to Items

Table B-17 displays the Algebra I CLEs matched to items (by sequential item number) per test form by reviewers.

Table B-17. Items Matched to Algebra I CLEs by Test Form Year

Form Year	CLE	Item Number
2009	A1B	11
2009	A1B	27
2009	A1C	33
2009	A1D	5
2009	A1D	21
2009	A1D	27
2009	A1D	44
2009	A1E	3
2009	A2A	1
2009	A2A	15
2009	A2A	21
2009	A2A	29
2009	A2A	34
2009	A2A	35
2009	A2B	13
2009	A2B	37
2009	A2B	38
2009	A2C	19
2009	A2C	35
2009	A2D	46
2009	A3A	29
2009	A4A	3
2009	A4A	17
2009	A4A	31
2009	D1A	16
2009	D1A	28
2009	D1C	47
2009	D2A	4
2009	D2A	32
2009	D2A	43
2009	D2C	36
2009	D3A	12
2009	D3A	20
2009	N1A	2
2009	N1A	10

2009	N1A	14
2009	N1A	30
2009	N1A	45
2009	N1B	14
2009	N1B	16
2009	N1B	18
2009	N1B	26
2009	N1B	28
2009	N1B	34
2009	N1B	38
2010	A1B	11
2010	A1C	1
2010	A1C	33
2010	A1D	5
2010	A1D	21
2010	A1D	33
2010	A1D	44
2010	A1E	3
2010	A1E	29
2010	A1E	32
2010	A2A	15
2010	A2A	35
2010	A2A	37
2010	A2A	45
2010	A2B	13
2010	A2C	17
2010	A2C	19
2010	A2D	19
2010	A2D	46
2010	A3A	27
2010	A4A	5
2010	A4A	32
2010	D1A	36
2010	D1C	26
2010	D1C	30
2010	D1C	47
2010	D2A	4
2010	D2A	43
2010	D2C	28
2010	D3A	12
2010	D3A	20
2010	N1A	2

2010	N1A	10
2010	N1A	14
2010	N1A	31
2010	N1B	16
2010	N1B	18
2010	N1B	26
2010	N1B	28
2010	N1B	34
2010	N1B	36
2010	N1B	38
2011	A1B	11
2011	A1D	1
2011	A1D	21
2011	A1D	30
2011	A1D	32
2011	A1D	47
2011	A1E	5
2011	A1E	17
2011	A1E	30
2011	A1E	43
2011	A2A	3
2011	A2A	10
2011	A2A	15
2011	A2A	26
2011	A2A	34
2011	A2A	36
2011	A2B	13
2011	A2B	38
2011	A2D	26
2011	A2D	28
2011	A2D	45
2011	A3A	1
2011	A3A	21
2011	A3A	47
2011	A4A	19
2011	A4A	32
2011	D1A	16
2011	D1C	33
2011	D2A	4
2011	D2A	44
2011	D2C	37
2011	D3A	12
2011	D3A	20

2011	N1A	2
2011	N1A	14
2011	N1A	31
2011	N1B	10
2011	N1B	16
2011	N1B	18
2011	N1B	27
2011	N1B	29
2011	N1B	35
2011	N1B	46

Appendix C
EOC Biology: Detailed Statistical Results

In Appendix C, we present the full alignment results for Biology. These results include (a) the four Webb measures, (b) consensus DOK ratings by CLE, (c) item alignment and quality ratings, (d) summary reviewer comments, and (e) items matched to course-level expectations (CLEs). All results are reported at the level of the content Strand.

For each analysis, we display the results first for the 2009 test form conducted on *all* operational items (multiple-choice and performance events). We then present results of analyses on the three test forms (2009 included) with only the multiple-choice items[3]. Note that the tables reporting results on analyses *only* of multiple-choice items for the 2009, 2010, and 2011 test forms do not list Scientific Inquiry because this strand is intended for assessment by performance events, as specified in the test specifications. No reviewers applied this strand as their primary match to multiple-choice items, and only two reviewers applied the strand to three multiple-choice items as a secondary match.

Webb Alignment Indicators

The following tables include complete statistical results on the four Webb alignment indicators: Categorical Concurrence, Depth-of-Knowledge (DOK) Consistency, Range of Knowledge, and Balance of Knowledge.

Categorical Concurrence

Tables C-1 and C-2 include categorical concurrence results: the mean number of items matched to strand by panelists, the standard deviation (S.D.) among panelists' ratings, and the final alignment conclusion (Yes or No). The criterion for acceptable Categorical Concurrence is a minimum of six items per strand.

Table C-1. Categorical Concurrence for Biology 2009 Test Form: Mean Number Multiple-Choice and Performance Event Items per Strand

	2009		
Strand	Mean Items per Strand	S.D.	At Least Six Items per Strand
Living Organisms	21.83	0.41	Yes
Ecology	13.00	0.00	Yes
Scientific Inquiry	14.83	0.75	Yes
Strands Matched to Six or More Items			3 of 3

[3] As a reminder to the reader, reviewers only rated performance events for the 2009 test forms.

Table C-2. Categorical Concurrence for Biology 2009, 2010, and 2011 Test Forms: Mean Number Multiple-Choice Only Items per Strand

	2009			2010			2011		
Strand	Mean Items per Strand	S.D.	At Least Six Items per Strand	Mean Items per Strand	S.D.	At Least Six Items per Strand	Mean Items per Strand	S.D.	At Least Six Items per Strand
Living Organisms	21.83	0.41	Yes	21.64	0.41	Yes	22.00	0.00	Yes
Ecology	13.00	0.00	Yes	13.46	0.23	Yes	13.00	0.00	Yes
Scientific Inquiry	0	0	No	0	0	No	0	0	No
Strands Matched to Six or More Items			2 of 3			2 of 3			2 of 3

Depth-of-Knowledge Consistency

Tables C-3 through C-10 present results of a comparative analysis between assessment items and CLEs on depth-of-knowledge (DOK). Tables C-3 through C-6 focus on the test item DOK relative to the corresponding CLEs. Specifically, these tables include the mean percentage of items per Strand rated below, at the same level, or above the DOK of the corresponding CLE. Webbs' criterion for acceptable DOK consistency is that item DOK must be At or Above the DOK level of the matched standard for at least 50% of items. Across the CLEs per Strand, we note (Yes or No) whether 50% of total items assessed CLEs as the appropriate cognitive level.

Table C-3. DOK Consistency for Biology 2009 Test Form: Mean Percentage of Multiple-Choice Items and Performance Event Below, At, or Above Corresponding CLEs

Strand	2009									
	[a] Mean Items per Strand	[b] Percent Items Below		[c] Percent Items At		[d] Percent Items Above		[e] Percent Items At/Above DOK of CLE		[f] DOK Target Met
		M	SD	M	SD	M	SD	M	SD	
Living Organisms	21.83	16%	0.09	77%	0.08	7%	0.04	84%	0.09	Yes
Ecology	13.00	29%	0.06	69%	0.07	1%	0.03	71%	0.06	Yes
Scientific Inquiry	14.83	86%	0.12	14%	0.12	0	0	14%	0.12	No
Strands with CLEs Assessed Appropriately										2 of 3

[a] Mean Items per Big Idea = Mean number of items matched to strand (Categorical Concurrence)
[b] Items Below = Percentage of items below DOK level of CLEs per strand.
[c] Items Same = Percentage of items with same DOK level as CLEs per strand.
[d] Items Above = Percentage of items above DOK level of CLEs per strand.
[e] Items At and Above = Percentage of items, when added, with DOK at the Same level and Above level of CLE.
[f] 50% or More Items = At least half of items assessing strand matched DOK level of corresponding CLEs.

Table C-4. DOK Consistency for Biology 2009 Test Form: Mean Percentage of Multiple-Choice Items Below, At, or Above Corresponding CLEs

Strand	2009									
	[a]Mean Items per Strand	[b] Percent Items Below		[c] Percent Items At		[d] Percent Items Above		[e] Percent Items At/ Above DOK of CLE		[f] DOK Target Met
		M	SD	M	SD	M	SD	M	SD	
Living Organisms	21.83	16%	0.09	77%	0.08	7%	0.04	84%	0.09	Yes
Ecology	13.00	29%	0.06	69%	0.07	1%	0.03	71%	0.06	Yes
Scientific Inquiry	0	0	0	0	0	0	0	0	0	No
Strands with CLEs Assessed Appropriately										2 of 3

[a] Mean Items per Big Idea = Mean number of items matched to strand (Categorical Concurrence)
[b] Items Below = Percentage of items below DOK level of CLEs per strand.
[c] Items Same = Percentage of items with same DOK level as CLEs per strand.
[d] Items Above = Percentage of items above DOK level of CLEs per strand.
[e] Items At and Above = Percentage of items, when added, with DOK at the Same level and Above level of CLE.
[f] 50% or More Items = At least half of items assessing strand matched DOK level of corresponding CLEs.

Table C-5. DOK Consistency for Biology 2010 Test Form: Mean Percentage of Multiple-Choice Items Below, At, or Above Corresponding CLEs

Strand	2010									
	[a] Mean Items per Strand	[b] Percent Items Below		[c] Percent Items At		[d] Percent Items Above		[e] Percent Items At/ Above DOK of CLE		[f] DOK Target Met
		M	SD	M	SD	M	SD	M	SD	
Living Organisms	21.64	12%	0.12	81%	0.13	6%	0.09	88%	0.12	Yes
Ecology	13.46	19%	0.11	81%	0.11	0	0	81%	0.11	Yes
Scientific Inquiry	0	0	0	0	0	0	0	0	0	No
Strands with CLEs Assessed Appropriately										2 of 3

[a] Mean Items per Big Idea = Mean number of items matched to strand (Categorical Concurrence)
[b] Items Below = Percentage of items below DOK level of CLEs per strand.
[c] Items Same = Percentage of items with same DOK level as CLEs per strand.
[d] Items Above = Percentage of items above DOK level of CLEs per strand.
[e] Items At and Above = Percentage of items, when added, with DOK at the Same level and Above level of CLE.
[f] 50% or More Items = At least half of items assessing strand matched DOK level of corresponding CLEs.

Table C-6. DOK Consistency for Biology 2011 Test Form: Mean Percentage of Multiple-Choice Items Below, At, or Above Corresponding CLEs

Strand	2011									
	[a] Mean Items per Strand	[b] Percent Items Below		[c] Percent Items At		[d] Percent Items Above		[e] Percent Items At/ Above DOK of CLE		[f] DOK Target Met
		M	SD	M	SD	M	SD	M	SD	
Living Organisms	22.00	22%	0.10	69%	0.11	9%	0.05	78%	0.10	Yes
Ecology	13.00	21%	0.11	78%	0.09	1%	0.03	79%	0.11	Yes
Scientific Inquiry	0	0	0	0	0	0	0	0	0	No
Strands with CLEs Assessed Appropriately										2 of 3

[a] Mean Items per Big Idea = Mean number of items matched to strand (Categorical Concurrence)
[b] Items Below = Percentage of items below DOK level of CLEs per strand.
[c] Items Same = Percentage of items with same DOK level as CLEs per strand.
[d] Items Above = Percentage of items above DOK level of CLEs per strand.
[e] Items At and Above = Percentage of items, when added, with DOK at the Same level and Above level of CLE.
[f] 50% or More Items = At least half of items assessing strand matched DOK level of corresponding CLEs.

Tables C-7 and C-10 summarize the same data in a different way by focusing on the percentage of CLEs assessed appropriately. These tables display the mean percentage of *standards* (CLEs) per Strand assessed at the appropriate DOK level (item DOK and standard DOK are the same), as well as the number of standards assessed below and above the level expected. At least 50% of items must be At or Above the DOK level of the corresponding CLE in order for the assessment of that CLE to be judged minimally appropriate.

Table C-7. DOK Consistency for Biology 2009 Test Form: Mean Percentage of CLEs Assessed by Multiple-Choice Items and Performance Events Below, At, or Above Expected DOK Level.

Strand	2009									
	[a] Mean Items per Strand	[b] Percent Items Below		[c] Percent Items At		[d] Percent Items Above		[e] Percent Items At/ Above DOK of CLE		[f] DOK Target Met
		M	SD	M	SD	M	SD	M	SD	
Living Organisms	17	16%	0.07	83%	0.06	1%	0.03	84%	0.07	Yes
Ecology	8	5%	0.72	5%	0.05	0	0	72%	0.05	Yes
Scientific Inquiry	15	13%	0.15	13%	0.12	0	0	15%	0.13	No
Strands with CLEs Assessed Appropriately										2 of 3

[a] Number of CLEs = Per Course Level Expectations 2.0, the number of CLEs per Strand.
[b] CLEs Assessed Below DOK = Percentage of CLEs per Strand assessed below consensus DOK level.
[c] CLEs Assessed At DOK = Percentage of CLEs per Strand assessed at the consensus DOK level.
[d] CLEs Assessed Above DOK = Percentage of CLEs per Strand assessed above consensus DOK level.
[e] CLEs Assessed At/Above DOK = Combined percentage of CLEs per Strand assessed at or above consensus DOK level.
[f] 50% or More CLEs = At least half of CLEs per strand were assessed by items at the appropriate DOK level.

Table C-8. DOK Consistency for Biology 2009 Test Form: Mean Percentage of CLEs Assessed By Multiple-Choice Items Below, At, or Above Expected DOK Level.

Strand	2009									
	[a] Mean Items per Strand	[b] Percent Items Below		[c] Percent Items At		[d] Percent Items Above		[e] Percent Items At/ Above DOK of CLE		[f] DOK Target Met
		M	SD	M	SD	M	SD	M	SD	
Living Organisms	17	16%	0.07	83%	0.06	1%	0.03	84%	0.07	Yes
Ecology	8	28%	0.05	72%	0.05	0	0	72%	0.05	Yes
Scientific Inquiry	15	0	0	0	0	0	0	0	0	No
Strands with CLEs Assessed Appropriately										2 of 3

[a] Number of CLEs = Per Course Level Expectations 2.0, the number of CLEs per Strand.
[b] CLEs Assessed Below DOK = Percentage of CLEs per Strand assessed below consensus DOK level.
[c] CLEs Assessed At DOK = Percentage of CLEs per Strand assessed at the consensus DOK level.
[d] CLEs Assessed Above DOK = Percentage of CLEs per Strand assessed above consensus DOK level.
[e] CLEs Assessed At/Above DOK = Combined percentage of CLEs per Strand assessed at or above consensus DOK level.
[f] 50% or More CLEs = At least half of CLEs per strand were assessed by items at the appropriate DOK level.

Table C-9. DOK Consistency for Biology 2010 Test Form: Mean Percentage of CLEs Assessed By Multiple-Choice Items Below, At, or Above Expected DOK Level.

Strand	[a] Mean Items per Strand	[b] Percent Items Below		[c] Percent Items At		[d] Percent Items Above		[e] Percent Items At/ Above DOK of CLE		[f] DOK Target Met
		2010								
		M	SD	M	SD	M	SD	M	SD	
Living Organisms	17	17%	0.13	76%	0.15	6%	0.09	83%	0.13	Yes
Ecology	8	19%	0.11	81%	0.11	0	0	81%	0.11	Yes
Scientific Inquiry	15	0	0	0	0	0	0	0	0	No
Strands with CLEs Assessed Appropriately										2 of 3

[a] Number of CLEs = Per Course Level Expectations 2.0, the number of CLEs per Strand.
[b] CLEs Assessed Below DOK = Percentage of CLEs per Strand assessed below consensus DOK level.
[c] CLEs Assessed At DOK = Percentage of CLEs per Strand assessed at the consensus DOK level.
[d] CLEs Assessed Above DOK = Percentage of CLEs per Strand assessed above consensus DOK level.
[e] CLEs Assessed At/Above DOK = Combined percentage of CLEs per Strand assessed at or above consensus DOK level.
[f] 50% or More CLEs = At least half of CLEs per strand were assessed by items at the appropriate DOK level.

Table C-10. DOK Consistency for Biology 2011 Test Form: Mean Percentage of CLEs Assessed By Multiple-Choice Items Below, At, or Above Expected DOK Level.

Strand	[a] Mean Items per Strand	2011								
		[b] Percent Items Below		[c] Percent Items At		[d] Percent Items Above		[e] Percent Items At/ Above DOK of CLE		[f] DOK Target Met
		M	SD	M	SD	M	SD	M	SD	
Living Organisms	17	21%	0.13	70%	0.10	9%	0.06	79%	0.13	Yes
Ecology	8	18%	0.11	80%	0.07	2%	0.06	82%	0.11	Yes
Scientific Inquiry	15	0	0	0	0	0	0	0	0	No
Strands with CLEs Assessed Appropriately										2 of 3

[a] Number of CLEs = Per Course Level Expectations 2.0, the number of CLEs per Strand.
[b] CLEs Assessed Below DOK = Percentage of CLEs per Strand assessed below consensus DOK level.
[c] CLEs Assessed At DOK = Percentage of CLEs per Strand assessed at the consensus DOK level.
[d] CLEs Assessed Above DOK = Percentage of CLEs per Strand assessed above consensus DOK level.
[e] CLEs Assessed At/Above DOK = Combined percentage of CLEs per Strand assessed at or above consensus DOK level.
[f] 50% or More CLEs = At least half of CLEs per strand were assessed by items at the appropriate DOK level.

Range-of-Knowledge Correspondence

Tables C-7 and C-8 present the range-of-knowledge correspondence results. The table includes the mean number (and percentage) of CLEs matched to at least one item per Strand. For acceptable range, a minimum of 50% of CLEs within each Strand should be matched to at least one item.

Table C-7. Range-of-Knowledge for Biology 2009 Test Form: Mean CLEs per Strand Linked with Multiple-Choice and Performance Event Items

	2009 Test Form						
Strand	Number of CLEs	Mean Items per Strand	Number of CLEs Assessed		% CLEs Assessed		50% or More CLEs
			M	S.D.	M	S.D.	
Living Organisms	17	21.83	14.00	0.89	82%	0.05	Yes
Ecology	8	13.00	7.00	0.00	88%	0.00	Yes
Scientific Inquiry	15	14.83	5.83	0.75	39%	0.05	No
Strands with CLEs Matched to At Least One Item							2 of 3

Table C-8. Range-of-Knowledge for Biology 2009, 2010, and 2011 Test Form: Mean CLEs per Strand Linked with Multiple-Choice Items

		2009 Test Form						2010 Test Form						2011 Test Form					
Strand	Number of CLEs	Mean Items per Strand	Number of CLEs Assessed		% CLEs Assessed		50% or More CLEs	Mean Items per Strand	Number of CLEs Assessed		% CLEs Assessed		50% or More CLEs	Mean Items per Strand	Number of CLEs Assessed		% CLEs Assessed		50% or More CLEs
			M	S.D.	M	S.D.			M	S.D.	M	S.D.			M	S.D.	M	S.D.	
Living Organisms	17	21.83	14.00	0.89	82%	0.05	Yes	21.64	11.80	0.75	70%	0.04	Yes	22.00	13.70	0.52	80%	0.03	Yes
Ecology	8	13.00	7.00	0.00	88%	0.00	Yes	13.46	7.83	0.41	98%	0.05	Yes	13.00	5.83	0.75	73%	0.09	Yes
Scientific Inquiry	15	0	0	0	0	0	No	0	0	0	0	0	No	0	0	0	0	0	No
Strands with CLEs Matched to At Least One Item							2 of 3						2 of 3						2 of 3

Balance-of-Knowledge Representation

Tables C-9 and C-10 display the balance indices for each Biology strand. This index is based on the mean number of items matched to each CLE. The minimum acceptable balance index is 70 out of 100. The table also includes the percentage of items linked to each Strand per strand.

Table C-9. Balance-of-Knowledge Representation for Biology 2009 Test Form: Balance Index per Strand

Strand	CLEs per Strand	2009					
		Mean CLEs Linked with Items	Mean Items per Strand	% of Items Linked to Strand		Balance Index	Balance Index Target Met
				M	S.D.		
Living Organisms	17	14.00	21.83	63%	0.35	0.80	Yes
Ecology	8	7.00	13.00	37%	0.00	0.84	Yes
Scientific Inquiry	15	5.83	14.83	30%	0.64	0.76	Yes
Balance Index Met							

Table C-10. Balance-of-Knowledge Representation for Biology 2009, 2010, and 2011 Test Forms: Balance Index per Strand

Strand	CLEs per Strand	2009						2010						2011					
		Mean CLEs Linked with Items	Mean Items per Strand	% of Items Linked to Strand		Balance Index	Balance Index Target Met	Mean CLEs Linked with Items	Mean Items per Strand	% of Items Linked to Strand		Balance Index	Balance Index Target Met	Mean CLEs Linked with Items	Mean Items per Strand	% of Items Linked to Strand		Balance Index	Balance Index Target Met
				M	S.D.					M	S.D.					M	S.D.		
Living Organisms	17	14.00	21.83	63%	0.35	0.80	Yes	11.80	21.64	62%	0.64	0.77	Yes	13.70	22.00	63%	0.00	0.82	Yes
Ecology	8	7.00	13.00	37%	0.00	0.84	Yes	7.83	13.46	38%	0.36	0.91	Yes	5.83	13.00	37%	0.00	0.88	Yes
Scientific Inquiry	15	0	0	0	0	0	N/A	0	0	0		0	N/A	0	0	0	0	0	N/A
Balance Index Met							2 of 3						2 of 3						2 of 3

Note: N/A indicates that no balance index was calculated for the Scientific Inquiry strand because reviewers did not match items to this strand.

Consensus DOK Ratings on CLEs

Table C-11 presents DOK ratings established through group consensus for each Biology CLE. Column 1 lists the Strand acronym, Big Idea number, Concept letter, and CLE letter. Column 2 displays the CLE content description. Column 3 indicates the DOK rating assigned to the CLE by the group.

Table C-11. Biology: Group Consensus Ratings on DOK Level per CLE

Strand, Big Idea, Concept, CLE	CLE Description	DOK
LO1Ba	Recognize cells both increase in number and differentiate, becoming specialized in structure and function, during and after embryonic development	1
LO1Cb	Describe the structure of cell parts (e.g., cell wall, cell membrane, cytoplasm, nucleus, chloroplast, mitochondrion, ribosome, vacuole) found in different types of cells (e.g., bacterial, plant, skin, nerve, blood, muscle) and the functions they perform (e.g., structural support, transport of materials, storage of genetic information, photosynthesis and respiration, synthesis of new molecules, waste disposal) that are necessary to the survival of the cell and organism	1
LO2Ac	Explain physical and chemical interactions that occur between organelles (e.g. nucleus, cell membrane, chloroplast, mitochondrion, ribosome) as they carry out life processes	2
LO2Ba	Explain the interrelationship between the processes of photosynthesis and cellular respiration (e.g., recycling of oxygen and carbon dioxide), comparing and contrasting photosynthesis and cellular respiration reactions	2
LO2Bb	Determine what factors affect the processes of photosynthesis and cellular respiration (i.e., light intensity, availability of reactants, temperature)	2
LO2Fa	Explain the significance of the selectively permeable membrane to the transport of molecules	2
LO2Fb	Predict the movement of molecules across a selectively permeable membrane (i.e., diffusion, osmosis, active transport) needed for a cell to maintain homeostasis given concentration gradients and different sizes of molecules	2
LO2Fc	Explain how water is important to cells (e.g., is a buffer for body temperature, provides soluble environment for chemical reactions, serves as a reactant in chemical reactions, provides hydration that maintains cell turgidity, maintains protein shape)	2
LO3Ba	Describe the chemical and structural properties of DNA (e.g., DNA is a large polymer formed from linked subunits of four kinds of nitrogen bases; genetic information is encoded in genes based on the sequence of subunits; each DNA molecule in a cell forms a single chromosome)	1
LO3Bb	Recognize that DNA codes for proteins, which are expressed as the heritable characteristics of an organism	1

LO3Be	Identify possible external causes (e.g., heat, radiation, certain chemicals) and effects of DNA mutations (e.g., altered proteins which may affect chemical reactions and structural development)	2
LO3Ca	Recognize the chromosomes of daughter cells, formed through the processes of asexual reproduction and mitosis, the formation of somatic (body) cells in multicellular organisms, are identical to the chromosomes of the parent cell	1
LO3Cb	Recognize that during meiosis, the formation of sex cells, chromosomes are reduced to half the number present in the parent cell	1
LO3Cc	Explain how fertilization restores the diploid number of chromosomes	2
LO3Da	Describe the advantages and disadvantages of asexual and sexual reproduction with regard to variation within a population	2
LO3Ea	Explain how genotypes (heterozygous and homozygous) contribute to phenotypic variation within a species	2
LO3Eb	Predict the probability of the occurrence of specific traits, including sex-linked traits, in an offspring by using a monohybrid cross	2
EC1Aa	Explain the nature of interactions between organisms in predator/prey relationships and different symbiotic relationships (i.e., mutualism, commensalisms, parasitism)	2
EC1Ab	Explain how cooperative (e.g., symbiotic) and competitive (e.g., predator/prey) relationships help maintain balance within an ecosystem	2
EC1Ba	Identify and explain the limiting factors (biotic and abiotic) that may affect the carrying capacity of a population within an ecosystem	2
EC1Da	Predict the impact (beneficial or harmful) a natural or human caused environmental event (e.g., forest fire, flood, volcanic eruption, avalanche, acid rain, global warming, pollution, deforestation, introduction of an exotic species) may have on the diversity of different species in an ecosystem	2
EC2Ac	Predict how the use and flow of energy will be altered due to changes in a food web	2
EC3Bb	Explain the importance of reproduction to the survival of a species (i.e., the failure of a species to reproduce will lead to extinction of that species)	2
EC3Ca	Identify examples of adaptations that may have resulted from variations favored by natural selection (e.g., long-necked giraffes, long-eared jack rabbits) and describe how that variation may have provided populations an advantage for survival	2
EC3Cc	Explain how environmental factors (e.g., habitat loss, climate change, pollution, introduction of non-native species) can be agents of natural selection	2
IN1Aa	Formulate testable questions and hypotheses	3
IN1Ab	Analyzing an experiment, identify the components (i.e., independent variable, dependent variables, control of constants, multiple trials) and explain their importance to the design of a valid experiment	3
IN1Ac	Design and conduct a valid experiment	4

IN1Ad	Recognize it is not always possible, for practical or ethical reasons, to control some conditions (e.g., when sampling or testing humans, when observing animal behaviors in nature)	2
IN1Ag	Evaluate the design of an experiment and make suggestions for reasonable improvements	3
IN1Bb	Measure length to the nearest millimeter, mass to the nearest gram, volume to the nearest milliliter, force (weight) to the nearest Newton, temperature to the nearest degree Celsius, time to the nearest second	1
IN1Bc	Determine the appropriate tools and techniques to collect, analyze, and interpret data	2
IN1Bd	Judge whether measurements and computation of quantities are reasonable	2
IN1Be	Calculate the range, average/mean, percent, and ratios for sets of data	1
IN1Ca	Use quantitative and qualitative data as support for reasonable explanations (conclusions)	3
IN1Cb	Analyze experimental data to determine patterns, relationships, perspectives, and credibility of explanations (e.g., predict/extrapolate data, explain the relationship between the independent and dependent variable)	3
IN1Cc	Identify the possible effects of errors in observations, measurements, and calculations, on the validity and reliability of data and resultant explanations (conclusions)	2
IN1Cd	Analyze whether evidence (data) and scientific principles support proposed explanations (laws/principles, theories/models)	3
IN1Da	Communicate the procedures and results of investigations and explanations through: oral presentations, drawings and maps, data tables (allowing for the recording and analysis of data relevant to the experiment such as independent and dependent variables, multiple trials, beginning and ending times or temperatures, derived quantities), graphs (bar, single, and multiple line), equations and writings	3
IN1Dc	Explain the importance of the public presentation of scientific work and supporting evidence to the scientific community (e.g., work and evidence must be critiqued, reviewed, and validated by peers; needed for subsequent investigations by peers; results can influence the decisions regarding future scientific work)	2

Biology Item Alignment to CLEs

Table B-12 provides the mean alignment rating per item based on the Overall Alignment rating scale (from '1=not aligned to any CLE' to '4=fully aligned to CLE; exemplary'). This rating serves as a confidence measure of the extent to which an item targets selected CLEs. The Biology panel included four reviewers.

Table B-12. Mean Overall Alignment Rating per Item for Each Biology Test Form

Form Year	Item Number	Degree of Alignment	
		M	S.D.
2009	1	3.33	0.52
2009	2	3.83	0.41
2009	3	3.67	0.52
2009	4	3.67	0.52
2009	5	3.83	0.41
2009	10	4.00	0.00
2009	11	3.83	0.41
2009	12	3.83	0.41
2009	13	3.50	0.55
2009	14	3.67	0.52
2009	15	3.83	0.41
2009	16	4.00	0.00
2009	17	3.17	0.75
2009	18	3.33	0.82
2009	19	3.83	0.41
2009	20	3.67	0.52
2009	21	3.67	0.52
2009	26	3.00	0.63
2009	27	3.67	0.52
2009	28	3.67	0.52
2009	29	3.67	0.52
2009	30	3.83	0.41
2009	31	3.83	0.41
2009	32	3.50	0.55
2009	33	3.83	0.41
2009	34	3.83	0.41
2009	35	3.33	0.52
2009	36	3.00	1.10
2009	37	3.50	0.55
2009	38	3.67	0.52
2009	43	3.83	0.41
2009	44	3.33	0.52
2009	45	3.33	0.52

Form Year	Item Number	Degree of Alignment	
		M	S.D.
2009	46	3.67	0.52
2009	47	3.67	0.52
2010	1	3.50	0.84
2010	2	3.67	0.52
2010	3	3.50	0.55
2010	4	3.67	0.52
2010	5	3.33	0.52
2010	10	3.50	0.55
2010	11	3.50	0.55
2010	12	3.67	0.52
2010	13	3.17	0.75
2010	14	3.83	0.41
2010	15	3.17	0.75
2010	16	3.17	0.75
2010	17	3.33	0.52
2010	18	3.67	0.52
2010	19	3.67	0.52
2010	20	3.67	0.52
2010	21	3.67	0.82
2010	26	3.83	0.41
2010	27	3.33	0.52
2010	28	2.83	0.98
2010	29	3.50	0.55
2010	30	3.00	0.89
2010	31	3.50	0.55
2010	32	3.50	0.84
2010	33	3.67	0.52
2010	34	3.67	0.52
2010	35	3.50	0.55
2010	36	3.00	0.89
2010	37	2.83	0.98
2010	38	3.50	0.84
2010	43	3.50	0.55
2010	44	3.67	0.52
2010	45	3.67	0.52
2010	46	3.83	0.41
2010	47	3.67	0.52
2011	1	3.67	0.52
2011	2	3.33	0.52
2011	3	2.83	0.98

Form Year	Item Number	Degree of Alignment	
		M	S.D.
2011	4	3.50	0.55
2011	5	3.33	0.82
2011	10	3.50	0.84
2011	11	3.33	0.82
2011	12	3.00	0.89
2011	13	3.67	0.52
2011	14	3.17	0.75
2011	15	3.00	0.89
2011	16	3.33	0.82
2011	17	3.50	0.55
2011	18	3.17	0.75
2011	19	3.67	0.52
2011	20	3.33	0.52
2011	21	3.50	0.55
2011	26	3.50	0.55
2011	27	3.33	0.82
2011	28	3.67	0.52
2011	29	3.50	0.55
2011	30	3.50	0.84
2011	31	3.83	0.41
2011	32	3.67	0.52
2011	33	3.33	0.52
2011	34	3.17	0.75
2011	35	3.50	0.55
2011	36	3.67	0.52
2011	37	3.67	0.52
2011	38	3.83	0.41
2011	43	3.50	0.55
2011	44	3.17	0.98
2011	45	3.00	0.89
2011	46	3.50	0.55
2011	47	3.50	0.55

Biology Item Quality Ratings

Table C-13 provides mean item quality ratings based on the Overall Item Quality rating scale (from '1= poor quality' to '4=exceptional quality'). This rating provides a global judgment on the format and clarity of items. The Biology panel included six reviewers.

Table C-13. Mean Overall Quality Rating per Item for Each Biology Test Form

Form Year	Item Number	Degree of Alignment	
		M	S.D.
2009	1	2.67	0.82
2009	2	3.67	0.52
2009	3	3.50	0.55
2009	4	3.67	0.52
2009	5	3.83	0.41
2009	10	3.83	0.41
2009	11	3.33	0.52
2009	12	3.83	0.41
2009	13	3.33	0.52
2009	14	3.33	0.52
2009	15	3.67	0.82
2009	16	3.83	0.41
2009	17	3.17	0.75
2009	18	3.50	0.55
2009	19	3.50	0.55
2009	20	3.67	0.52
2009	21	2.67	1.03
2009	26	3.17	0.75
2009	27	3.67	0.52
2009	28	3.17	0.41
2009	29	3.67	0.52
2009	30	3.50	0.55
2009	31	3.33	0.52
2009	32	3.67	0.52
2009	33	3.33	0.52
2009	34	3.83	0.41
2009	35	3.00	0.63
2009	36	3.50	0.55
2009	37	3.33	0.52
2009	38	3.50	0.55
2009	43	3.50	0.55
2009	44	3.00	0.63
2009	45	3.33	0.52

Form Year	Item Number	Degree of Alignment	
		M	S.D.
2009	46	3.17	0.41
2009	47	3.67	0.52
2010	1	3.67	0.52
2010	2	3.67	0.52
2010	3	3.67	0.52
2010	4	3.67	0.52
2010	5	3.33	0.52
2010	10	3.50	0.55
2010	11	3.50	0.55
2010	12	3.67	0.52
2010	13	3.33	0.52
2010	14	3.83	0.41
2010	15	3.33	0.52
2010	16	3.17	0.75
2010	17	3.50	0.55
2010	18	2.83	1.17
2010	19	3.67	0.52
2010	20	3.33	0.82
2010	21	3.83	0.41
2010	26	3.67	0.52
2010	27	3.17	0.75
2010	28	3.33	0.52
2010	29	3.67	0.52
2010	30	3.17	0.75
2010	31	3.33	0.52
2010	32	3.67	0.52
2010	33	3.50	0.55
2010	34	3.67	0.52
2010	35	3.33	0.52
2010	36	3.33	0.52
2010	37	3.33	0.82
2010	38	3.67	0.52
2010	43	3.50	0.55
2010	44	3.50	0.55
2010	45	3.67	0.52
2010	46	3.50	0.55
2010	47	3.67	0.52
2011	1	3.67	0.52
2011	2	3.67	0.52
2011	3	2.83	1.17

Form Year	Item Number	Degree of Alignment	
		M	S.D.
2011	4	3.33	0.82
2011	5	3.33	0.52
2011	10	3.50	0.55
2011	11	3.17	0.75
2011	12	3.17	0.98
2011	13	3.83	0.41
2011	14	3.33	0.52
2011	15	3.50	0.55
2011	16	3.50	0.55
2011	17	3.50	0.55
2011	18	3.50	0.55
2011	19	3.50	0.55
2011	20	3.00	0.89
2011	21	3.50	0.55
2011	26	3.50	0.55
2011	27	3.50	0.55
2011	28	3.67	0.52
2011	29	3.50	0.55
2011	30	3.67	0.52
2011	31	3.33	0.82
2011	32	3.67	0.52
2011	33	3.17	0.41
2011	34	2.33	1.21
2011	35	3.33	0.52
2011	36	3.50	0.55
2011	37	3.50	0.55
2011	38	3.50	0.55
2011	43	3.50	0.55
2011	44	3.00	0.89
2011	45	2.67	1.21
2011	46	3.67	0.52
2011	47	2.50	1.38

Panelist Comments on Biology Items

Tables C-14 through C-16 present panelists' comments on the individual items for the English II test forms. To maintain test security, no individual item identifiers are included.

Table C-14. Reviewer Comments on 2009 Test Form Items for Biology

Test Form	Item Number	Reviewer Comment
2009	1	Abiguous answers. Item does not describe the effect on the host organism (benefit, detriment or no effect) which will help to specifically identify the type of symbiosis
2009	1	Abiguous answers. Item does not describe the effect on the host organism (benefit, detriment or no effect) which will help to specifically identify the type of symbiosis
2009	1	unclear without prior knowledge that negative releationship
2009	13	It might be helpful to know the solute or water concentration inside of the red blood cell.
2009	13	It might be helpful to know the solute or water concentration inside of the red blood cell.
2009	15	the terminology somatic (Q)and parent(CLE) are dissimilar
2009	17	Specific term not in CLE
2009	17	term is not mentioned specifically on cle
2009	18	the cle is not asking for effect of factors, the question is
2009	21	DNA affected by heat?
2009	21	focus of Q seems to be on DNA while distractors do not include DNA. unclear what the level of heat is. Distractor focus on protien: Protein transcription/ translation enzymes denatured only if high enough, so no new protein produced . Transcription enzyme error rate is affected by heat. I don't think this is taught in most HS classrooms.
2009	26	term is not mentioned specifically on cle
2009	26	This CLE does not refer to levels of cell organization as asked in the question
2009	35	I find the distractors too distracting, difficult to see the correct choice at high school level.
2009	36	Can make the case for multiple CLE alignment
2009	36	Can make the case for multiple CLE alignment
2009	36	multiple CLEs
2009	36	the cle is not asking for effect of factors, the question is
2009	37	Answer choices are confusing
2009	37	Answer choices are confusing
2009	38	I think Q is OK, student needs to elimate distractors to find answer (easily).
2009	43	thirty days!
2009	44	would prefer link to genetic diversity

2009	47	multiple CLEs
2009	50	level of complexity does not warrant a DOK 3; students are simply identifying components of the presented data/experiment
2009	50	level of complexity does not warrant a DOK 3; students are simply identifying components of the presented data/experiment
2009	51	level of complexity does not warrant a DOK 3; students are simply identifying components of the presented data/experiment
2009	51	level of complexity does not warrant a DOK 3; students are simply identifying components of the presented data/experiment
2009	52	level of complexity does not warrant a DOK 3; students are simply identifying components of the presented data/experiment
2009	52	level of complexity does not warrant a DOK 3; students are simply identifying components of the presented data/experiment
2009	57	multiple CLEs

Table C-15. Reviewer Comments on 2010 Test Form Items for Biology

Test Form	Item Number	Reviewer Comment
2010	1	CLE refers to comparing "processes". No mention in the CLE regarding knowledge of the products of each reaction.
2010	13	the cle is not asking for effect of factors, the question is
2010	15	There is too much variation in the distractors
2010	16	Difficult to choose the CLE
2010	16	There is no clear CLE that this Q fits into.
2010	18	Size and "shape" are too similar. Also, water concentration affects this.
2010	18	size vs shape concepts too close, better to have charge as answer
2010	20	pain relief tied to destroying cells, difficult for students except in classrooms where this is text/teacher example.
2010	21	the cle is not asking for effect of factors, the question is
2010	27	Assumes prior knowledge about specific organism relationships
2010	28	no mention of event sequence of protein synthesis in cle
2010	28	The CLE's makes no reference to knowing how protein synthesis occurs.
2010	28	Weakly aligned to this benchmark
2010	30	cle does not address ongoing processes well
2010	30	Weakly aligned to this benchmark
2010	32	the cle is not asking for effect of factors, the question is
2010	35	distractors would be better with organelles and not photosynthesis products with this stem
2010	36	cle is not asking for effect of factors, the question is
2010	36	Can make the case for either one
2010	36	The Q doesn't fit into any CLE easily. There can be multiple CLE's it may fit into.

2010	37	Asks for specific steps in protein synthesis.
2010	37	no mention of event sequence of protein synthesis in cle
2010	37	No reference is made in the CLE regarding protein synthesis.
2010	37	Weakly aligned to this benchmark
2010	38	cle is not asking for effect of factors, the question is

Table C-16. Reviewer Comments on 2010 Test Form Items for Biology

2011	3	Asks a specific term instead of asking about chromosome number as stated in the CLE.
2011	3	terms are not mentioned in cles
2011	3	The Q is too asking too specific information regarding this CLE
2011	3	Weakly aligned to CLE
2011	4	movement is asked for, result is in answers
2011	5	Wording of the question is problematic.
2011	10	Weakly aligned to CLE
2011	11	thrive could mean stable or increase for pop A @ 25oC. No mention of incr temp being a negative affect on population. What choice? I see B and D with current wording.
2011	11	q is asking to predict, cle is not
2011	12	RNA structure in not mentioned, nor inferred, anywhere in the CLE's
2011	12	Specific question is about RNA but the CLE mainly refers to DNA.
2011	12	the process of protein synthesis is not in cle
2011	14	no mention of polarity in cle
2011	14	Q#14 & 40 are very similar. BOTH stems are the same, answers same - except 14 has two factors.
2011	15	Requires knowledge from both CLEs.
2011	16	q is asking to predict, cle is not
2011	18	q is asking to predict, cle is not
2011	20	think this fact fits better with natural selection
2011	20	This Q overall is a poor Q, maybe too high a level for average Bio students.
2011	27	q is asking to predict, cle is not
2011	30	q is asking to predict, cle is not
2011	31	reproductive cell could be a cell of an organ, not just gamete
2011	34	A strong case can be made that there are 2 possible answers, B and D.
2011	34	Answer choices could be B or D.
2011	34	Multiple answers; assumes a vast amount of prior knowledge
2011	34	some argument about protist interactions not being affected
2011	44	Refers to base pairing and the CLE specifically says NOT to memorize base pairing.
2011	44	students can find correct answer by eliminating distractors, however wording of answer is not clear.
2011	44	term complementary is in q, not cle

2011	44	The CLE directly says not to assess knowing the N base pairs. To answer the Q correctly a student must have the knowledge that N base pairs must match correctly. This seems in direct conflict with the CLE.
2011	45	Assumes specific prior knowledge
2011	45	Not all deserts are warm. In some cold deserts keeping warm may be an issue to animals. Students may have trouble answering this Q correctly since deserts are determined by amount of precipitation rather than temperature. Why do rabbits in MO have long ears?
2011	45	q is asking to predict, cle is not
2011	45	students unlikely to associate long ears with increased heat release.
2011	45	Unless the student has been specifically taught about ears releasing heat, this could not be known. Some deserts are cold.
2011	47	A case can be made that answer A is also correct. As photosyn rate increases, more ATP is produced in the light reactions.
2011	47	atp and oxygen would increase (two correct responses)
2011	47	stem wording or distractors could be changed. Perhaps stem, most likely to immediately increase.
2011	47	The wording of the stem makes choices A and C correct possibilities.

Biology CLEs Matched to Items

Table C-17 displays the Biology CLEs matched to items (by sequential item number) per test form by reviewers.

Table C-17. Items Matched to Biology CLEs by Test Form Year

Form Year	CLE	Item Number
2009	31BA	17
2009	31BA	26
2009	31CB	3
2009	31CB	4
2009	31CB	16
2009	31CB	20
2009	31CB	5
2009	32BA	12
2009	32BA	5
2009	32BA	16
2009	32BB	5
2009	32FA	13
2009	32FB	13
2009	32FB	14
2009	32FB	37
2009	32FC	43
2009	33BA	32
2009	33BB	10
2009	33BE	21
2009	33BE	45
2009	33CA	27
2009	33CA	29
2009	33CB	15
2009	33CC	27
2009	33CC	47
2009	33DA	33
2009	33EA	2
2009	33EB	2
2009	41AA	1
2009	41AA	34
2009	41AA	18
2009	41AB	18
2009	41AB	44
2009	41BA	11
2009	41BA	46

Form Year	CLE	Item Number
2009	41DA	30
2009	41DA	31
2009	41DA	35
2009	41DA	36
2009	42AC	28
2009	42AC	38
2009	43BB	44
2009	43CA	19
2009	43CA	36
2009	43CC	36
2009	71CA	37
2010	999	30
2010	31CB	4
2010	31CB	11
2010	31CB	35
2010	31CB	47
2010	32BA	1
2010	32BB	14
2010	32BB	21
2010	32BB	32
2010	32BB	38
2010	32FA	18
2010	32FB	18
2010	32FB	20
2010	33BA	5
2010	33BA	43
2010	33BA	37
2010	33BB	28
2010	33BB	37
2010	33BB	11
2010	33BB	33
2010	33BB	35
2010	33BE	15
2010	33BE	33
2010	33CA	30
2010	33CB	2
2010	33CB	12
2010	33CC	12
2010	33DA	34
2010	33EB	45
2010	41AA	27

Form Year	CLE	Item Number
2010	41AA	46
2010	41AA	16
2010	41AB	3
2010	41AB	16
2010	41AB	10
2010	41BA	13
2010	41BA	36
2010	41BA	29
2010	41DA	17
2010	41DA	19
2010	42AC	3
2010	42AC	10
2010	42AC	26
2010	43BB	31
2010	43CA	29
2010	43CA	44
2010	43CC	19
2010	43CC	36
2011	31BA	21
2011	31BA	3
2011	31CB	2
2011	31CB	26
2011	31CB	43
2011	32BA	2
2011	32BB	47
2011	32FA	26
2011	32FA	32
2011	32FA	14
2011	32FB	4
2011	32FB	14
2011	32FB	32
2011	32FB	46
2011	32FC	17
2011	32FC	38
2011	33BA	12
2011	33BA	44
2011	33BB	12
2011	33BB	35
2011	33BE	10
2011	33BE	33

Form Year	CLE	Item Number
2011	33CA	19
2011	33CB	5
2011	33CB	31
2011	33CC	3
2011	33DA	15
2011	33EA	15
2011	33EB	13
2011	41AA	1
2011	41AA	20
2011	41AA	28
2011	41AB	20
2011	41BA	11
2011	41BA	16
2011	41BA	18
2011	41BA	30
2011	41DA	18
2011	41DA	29
2011	41DA	34
2011	41DA	37
2011	42AC	30
2011	42AC	34
2011	42AC	18
2011	43CA	36
2011	43CA	45
2011	43CC	11
2011	43CC	27
2011	43CC	16
2011	43CC	29

Appendix D
Sample Alignment Review Materials

Panelists received a reference guide for making DOK ratings. Each content area received a different reference guide specific to its content review.

English

Reading DOK Levels

The reading levels are based on Valencia and Wixson (2000, pp. 909-935). The writing levels were developed by Marshá Horton, Sharon O'Neal, and Phoebe Winter.

Reading Level 1. Level 1 requires students to receive or recite facts or to use simple skills or abilities. Oral reading that does not include analysis of the text, as well as basic comprehension of a text, is included. Items require only a shallow understanding of the text presented and often consist of verbatim recall from text, slight paraphrasing of specific details from the text, or simple understanding of a single word or phrase. Some examples that represent, but do not constitute all of, Level 1 performance are:

- Support ideas by reference to verbatim or only slightly paraphrased details from the text.
- Use a dictionary to find the meanings of words.
- Recognize figurative language in a reading passage.

Reading Level 2. Level 2 includes the engagement of some mental processing beyond recalling or reproducing a response; it requires both comprehension and subsequent processing of text or portions of text. Inter-sentence analysis of inference is required. Some important concepts are covered, but not in a complex way. Standards and items at this level may include words such as summarize, interpret, infer, classify, organize, collect, display, compare, and determine whether fact or opinion. Literal main ideas are stressed. A Level 2 assessment item may require students to apply skills and concepts that are covered in Level 1. However, items require closer understanding of text, possibly through the item's paraphrasing of both the question and the answer. Some examples that represent, but do not constitute all of, Level 2 performance are:

- Use context cues to identify the meaning of unfamiliar words, phrases, and expressions that could otherwise have multiple meanings.
- Predict a logical outcome based on information in a reading selection.
- Identify and summarize the major events in a narrative.

Reading Level 3. Deep knowledge becomes a greater focus at Level 3. Students are encouraged to go beyond the text; however, they are still required to show understanding of the ideas in the text. Students may be encouraged to explain, generalize, or connect ideas. Standards and items at Level 3 involve reasoning and planning. Students must be able to support their thinking. Items may involve abstract theme identification, inference across an entire passage, or

students' application of prior knowledge. Items may also involve more superficial connections between texts. Some examples that represent, but do not constitute all of, Level 3 performance are:

- Explain or recognize how the author's purpose affects the interpretation of a reading selection.
- Summarize information from multiple sources to address a specific topic.
- Analyze and describe the characteristics of various types of literature.

Reading Level 4. Higher-order thinking is central and knowledge is deep at Level 4. The standard or assessment item at this level will probably be an extended activity, with extended time provided for completing it. The extended time period is not a distinguishing factor if the required work is only repetitive and does not require the application of significant conceptual understanding and higher-order thinking. Students take information from at least one passage of a text and are asked to apply this information to a new task. They may also be asked to develop hypotheses and perform complex analyses of the connections among texts. Some examples that represent, but do not constitute all of, Level 4 performance are:

- Analyze and synthesize information from multiple sources.
- Examine and explain alternative perspectives across a variety of sources.
- Describe and illustrate how common themes are found across texts from different cultures.

NOTE: Many on-demand assessment instruments will not include assessment activities that could be classified as Level 4. However, standards, goals, and objectives can be stated so as to expect students to perform thinking at this level. On-demand assessments that do include tasks, products, or extended responses would be classified as Level 4 when the task or response requires evidence that the cognitive requirements have been met. [added October 2009_LRT]

Writing DOK Levels

Writing Level 1. Level 1 requires the student to write or recite simple facts. The focus of this writing or recitation is not on complex synthesis or analysis, but on basic ideas. The students are asked to list ideas or words, as in a brainstorming activity, prior to written composition; are engaged in a simple spelling or vocabulary assessment; or are asked to write simple sentences. Students are expected to write, speak, and edit using the conventions of Standard English. This includes using appropriate grammar, punctuation, capitalization, and spelling. Students demonstrate a basic understanding and appropriate use of such reference materials as a dictionary, thesaurus, or Web site. Some examples that represent, but do not constitute all of, Level 1 performance are:

- Use punctuation marks correctly.
- Identify Standard English grammatical structures, including the correct use of verb tenses.

*Writing Level 2***.** Level 2 requires some mental processing. At this level, students are engaged in first-draft writing or brief extemporaneous speaking for a limited number of purposes and audiences. Students are expected to begin connecting ideas, using a simple organizational structure. For example, students may be engaged in note-taking, outlining, or simple summaries. Text may be limited to one paragraph. Some examples that represent, but do not constitute all of, Level 2 performance are:

- Construct or edit compound or complex sentences, with attention to correct use of phrases and clauses.
- Use simple organizational strategies to structure written work.
- Write summaries that contain the main idea of the reading selection and pertinent details.

*Writing Level 3***.** Level 3 requires some higher-level mental processing. Students are engaged in developing compositions that include multiple paragraphs. These compositions may include complex sentence structure and may demonstrate some synthesis and analysis. Students show awareness of their audience and purpose through focus, organization, and the use of appropriate compositional elements. The use of appropriate compositional elements includes such things as addressing chronological order in a narrative, or including supporting facts and details in an informational report. At this stage, students are engaged in editing and revising to improve the quality of the composition. Some examples that represent, but do not constitute all of, Level 3 performance are:

- Support ideas with details and examples.
- Use voice appropriate to the purpose and audience.
- Edit writing to produce a logical progression of ideas.

Writing Level 4. Higher-level thinking is central to Level 4. The standard at this level is a multi-paragraph composition that demonstrates the ability to synthesize and analyze complex ideas or themes. There is evidence of a deep awareness of purpose and audience. For example,

informational papers include hypotheses and supporting evidence. Students are expected to create compositions that demonstrate a distinct voice and that stimulate the reader or listener to consider new perspectives on the addressed ideas and themes. An example that represents, but does not constitute all of, Level 4 performance is:

- Write an analysis of two selections, identifying the common theme and generating a purpose that is appropriate for both.

Algebra I

DOK Levels

Level 1 (Recall) includes the recall of information such as a fact, definition, term, or a simple procedure, as well as performing a simple algorithm or applying a formula. That is, in Algebra I, a one-step, well-defined, and straight algorithmic procedure should be included at this lowest level. Other key words that signify Level 1 include "identify," "recall," "recognize," "use," and "measure." Verbs such as "describe" and "explain" could be classified at different levels, depending on what is to be described and explained.

Level 2 (Skill/Concept) includes the engagement of some mental processing beyond an habitual response. A Level 2 assessment item requires students to make some decisions as to how to approach the problem or activity, whereas Level 1 requires students to demonstrate a rote response, perform a well-known algorithm, follow a set procedure (like a recipe), or perform a clearly defined series of steps. Keywords that generally distinguish a Level 2 item include "classify," "organize," "estimate," "make observations," "collect and display data," and "compare data." These actions imply more than one step. For example, to compare data requires first identifying characteristics of objects or phenomena and then grouping or ordering the objects. Some action verbs, such as "explain," "describe," or "interpret," could be classified at different levels depending on the object of the action. For example, interpreting information from a simple graph, or reading information from the graph, also are at Level 2. Interpreting information from a complex graph that requires some decisions on what features of the graph need to be considered and how information from the graph can be aggregated is at Level 3. Level 2 activities are not limited only to number skills, but may involve visualization skills and probability skills. Other Level 2 activities include noticing or describing non-trivial patterns, explaining the purpose and use of experimental procedures; carrying out experimental procedures; making observations and collecting data; classifying, organizing, and comparing data; and organizing and displaying data in tables, graphs, and charts.

Level 3 (Strategic Thinking) requires reasoning, planning, using evidence, and a higher level of thinking than the previous two levels. In most instances, requiring students to explain their thinking is at Level 3. Activities that require students to make conjectures are also at this level. The cognitive demands at Level 3 are complex and abstract. The complexity does not result from the fact that there are multiple answers, a possibility for both Levels 1 and 2, but because the task requires more demanding reasoning. An activity, however, that has more than one possible answer and requires students to justify the response they give would most likely be at Level 3.

Other Level 3 activities include drawing conclusions from observations; citing evidence and developing a logical argument for concepts; explaining phenomena in terms of concepts; and deciding which concepts to apply in order to solve a complex problem.

Level 4 (Extended Thinking) requires complex reasoning, planning, developing, and thinking, most likely over an extended period of time. The extended time period is not a distinguishing factor if the required work is only repetitive and does not require applying significant conceptual understanding and higher-order thinking. For example, if a student has to take the water temperature from a river each day for a month and then construct a graph, this would be classified as a Level 2. However, if the student is to conduct a river study that requires taking into consideration a number of variables, this would be a Level 4. At Level 4, the cognitive demands of the task should be high and the work should be very complex. Students should be required to make several connections—relate ideas *within* the content area or *among* content areas—and have to select one approach among many alternatives on how the situation should be solved, in order to be at this highest level. Level 4 activities include designing *and* conducting experiments and projects; developing and proving conjectures, making connections between a finding and related concepts and phenomena; combining and synthesizing ideas into new concepts; and critiquing experimental designs.

> NOTE: Many on-demand assessment instruments will not include assessment activities that could be classified as Level 4. However, standards, goals, and objectives can be stated so as to expect students to perform thinking at this level. On-demand assessments that do include tasks, products, or extended responses would be classified as Level 4 when the task or response requires evidence that the cognitive requirements have been met. *[added October 2009_LRT]*

Biology

Biology DOK Levels

Please note that, in Biology, "knowledge" can refer both to content knowledge and knowledge of scientific processes. This meaning of knowledge is consistent with the *National Biology Education Standards* (NSES), which terms "Biology as Inquiry" as its first Content Standard.

Level 1 (Recall and Reproduction) requires the recall of information, such as a fact, definition, term, or a simple procedure, as well as performance of a simple Biology process or procedure. Level 1 only requires students to demonstrate a rote response, use a well-known formula, follow a set procedure (like a recipe), or perform a clearly defined series of steps. A "simple" procedure is well defined and typically involves only one step. Verbs such as "identify," "recall," "recognize," "use," "calculate," and "measure" generally represent cognitive work at the recall and reproduction level. Simple word problems that can be directly translated into and solved by a formula are considered Level 1. Verbs such as "describe" and "explain" could be classified at different DOK levels, depending on the complexity of what is to be described and explained.

A student answering a Level 1 item either knows the answer or does not: that is, the item does not need to be "figured out" or "solved." In other words, if the knowledge necessary to answer an item automatically provides the answer to it, then the item is at Level 1. If the knowledge needed to answer the item is not automatically provided in the stem, the item is at least at Level 2. Some examples that represent, but do not constitute all of, Level 1 performance are:

- Recall or recognize a fact, term, or property.
- Represent in words or diagrams a scientific concept or relationship.
- Provide or recognize a standard scientific representation for simple phenomenon.
- Perform a routine procedure, such as measuring length.

Level 2 (Skills and Concepts) includes the engagement of some mental processing beyond recalling or reproducing a response. The content knowledge or process involved is **more complex** than in Level 1. Items require students to make some decisions as to how to approach the question or problem. Keywords that generally distinguish a Level 2 item include "classify," "organize," "estimate," "make observations," "collect and display data," and "compare data." These actions imply **more than one step**. For example, to compare data requires first identifying characteristics of the objects or phenomena and then grouping or ordering the objects. Level 2 activities include making observations and collecting data; classifying, organizing, and comparing data; and organizing and displaying data in tables, graphs, and charts. Some action verbs, such as "explain," "describe," or "interpret," could be classified at different DOK levels, depending on the complexity of the action. For example, interpreting information from a simple graph, requiring reading information from the graph, is a Level 2. An item that requires interpretation from a complex graph, such as making decisions regarding features of the graph

that need to be considered and how information from the graph can be aggregated, is at Level 3. Some examples that represent, but do not constitute all of, Level 2 performance, are:

- Specify and explain the relationship between facts, terms, properties, or variables.
- Describe and explain examples and non-examples of Biology concepts.
- Select a procedure according to specified criteria and perform it.
- Formulate a routine problem, given data and conditions.
- Organize, represent, and interpret data.

Level 3 (Strategic Thinking) requires reasoning, planning, using evidence, and a higher level of thinking than the previous two levels. The cognitive demands at Level 3 are complex and abstract. The complexity does not result only from the fact that there could be multiple answers, a possibility for both Levels 1 and 2, but because the multi-step task requires more demanding reasoning. In most instances, requiring students to explain their thinking is at Level 3; requiring a very simple explanation or a word or two should be at Level 2. An activity that has more than one possible answer and requires students to justify the response they give would most likely be a Level 3. Experimental designs in Level 3 typically involve more than one dependent variable. Other Level 3 activities include drawing conclusions from observations; citing evidence and developing a logical argument for concepts; explaining phenomena in terms of concepts; and using concepts to solve non-routine problems. Some examples that represent, but do not constitute all of Level 3 performance, are:

- Identify research questions and design investigations for a scientific problem.
- Solve non-routine problems.
- Develop a scientific model for a complex situation.
- Form conclusions from experimental data.

Level 4 (Extended Thinking) involves high cognitive demands and complexity. Students are required to make several connections—relate ideas within the content area or among content areas—and have to select or devise one approach among many alternatives to solve the problem. Many on-demand assessment instruments will not include any assessment activities that could be classified as Level 4. However, standards, goals, and objectives can be stated in such a way as to expect students to perform extended thinking. “Develop generalizations of the results obtained and the strategies used and apply them to new problem situations,” is an example of a grade 8 objective that is a Level 4. Many, but not all, performance assessments and open-ended assessment activities requiring significant thought will be Level 4.

B-Level 4 requires complex reasoning, experimental design and planning, and probably will require an extended period of time either for the Biology investigation required by an objective, or for carrying out the multiple steps of an assessment item. However, the extended time period is not a distinguishing factor if the required work is only repetitive and does not require applying significant conceptual understanding and higher-order thinking. For example, if a student has to take the water temperature from a river each day for a month and then construct a graph, this would be classified as a Level 2 activity. However, if the student conducts a river study that requires taking into consideration a number of variables, this would be a Level 4. Some examples that represent, but do not constitute all of, a Level 4 performance are:

- Based on data provided from a complex experiment that is novel to the student, deduct the fundamental relationship between several controlled variables.
- Conduct an investigation, from specifying a problem to designing and carrying out an experiment, to analyzing its data and forming conclusions.

NOTE: Many on-demand assessment instruments will not include assessment activities that could be classified as Level 4. However, standards, goals, and objectives can be stated so as to expect students to perform thinking at this level. On-demand assessments that do include tasks, products, or extended responses would be classified as Level 4 when the task or response requires evidence that the cognitive requirements have been met. *[added October 2009_LRT]*

Panelists rated the depth-of-knowledge (DOK) level of the Missouri Course Level Expectations (CLEs) electronically using an Excel spreadsheet. This spreadsheet is the same as published on the DESE website with two changes: (a) the addition of a column in which to enter DOK ratings, and (b) elimination of locally assessed standards. A portion of the rating sheet used for English II is included below as an example of format.

Content_Area	Strand	Big_Idea	Concept	Grade	CLE_Code	CLE_Description	DOK
Communication Arts	R	1	E	E2	a	Develop vocabulary through text, using roots and affixes	
Communication Arts	R	1	E	E2	b	Develop vocabulary through text, using context clues	
Communication Arts	R	1	E	E2	c	Develop vocabulary through text, using glossary, dictionary and thesaurus	
Communication Arts	R	1	H	E2	a	Apply post-reading skills to comprehend, interpret, analyze, and evaluate text: identify and explain the relationship between the main idea and supporting details	
Communication Arts	R	1	H	E2	d	Apply post-reading skills to comprehend, interpret, analyze, and evaluate text: draw conclusions	
Communication Arts	R	1	H	E2	e	Apply post-reading skills to comprehend, interpret, analyze, and evaluate text: paraphrase	
Communication Arts	R	1	H	E2	f	Apply post-reading skills to comprehend, interpret, analyze, and evaluate text: summarize	
Communication Arts	R	1	I	E2	a	Compare, contrast, analyze and evaluate connections: text to text (information and relationships in various fiction and non-fiction works)	
Communication Arts	R	2	A	E2		Analyze and evaluate the text features in grade-level text	
Communication Arts	R	2	B	E2	a	Identify and explain literary techniques, in text emphasizing understatement	
Communication Arts	R	2	B	E2	b	Identify and explain literary techniques, in text emphasizing parallelism	
Communication Arts	R	2	B	E2	c	Identify and explain literary techniques, in text emphasizing allusion	
Communication Arts	R	2	B	E2	d	Identify and explain literary techniques, in text emphasizing analogy	
Communication Arts	R	2	B	E2	e	Identify and explain literary techniques, in text emphasizing analyze and evaluate literary techniques, sensory details, figurative language and sound devices previously introduced	

Panelists rated individual test form items also using an Excel spreadsheet. The format of the rating form was identical for each course test form reviewed. The graphic below demonstrates the format of the rating form fr English II on computer screen.

English II

Item Number	Depth Of Knowledge	CLE 1	CLE 2	Overall Alignment	Overall Item Quality	Explanation
(Number Listed in Test Form)	(Enter Level 1to 4)	(Enter HumRRO Code)	(Enter HumRRO Code)	(Enter Scale of 1 to 4)	(Enter Scale of 1 to 4)	Use ONLY IF you entered a low rating (a 1 or 2) on Overall Alignment or Overall Item Quality
1						
2						
3						
4						
5						