

## Understanding Question (Item) Level of Difficulty

WNYRIC's **Assessment Analysis – Item Difficulty** report provides information regarding the level of difficulty for every regents item and every released 3-8 Math and ELA item. The report is provided in Excel format and indicates the learning standard and regional p-value for each item as well as cut point performance information.

	A	B	C	D	E	K
1	Assessment Date	Assessment Name	Question	Question Level	Identifier	Regional p-value
2	2014Jun	Common Core Algebra I	II-28 - CR	LEVEL 5	F.BF.3	0.11
3	2014Jun	Common Core Algebra I	II-31 - CR	LEVEL 5	A.SSE.2	0.15
4	2014Jun	Common Core Algebra I	III-34 - CR	LEVEL 5	A.CED.1	0.09
5	2014Jun	Common Core Algebra I	I-08 - MC	LEVEL 4	A.REI.4a	0.40
6	2014Jun	Common Core Algebra I	I-10 - MC	LEVEL 4	A.REI.4b	0.33
7	2014Jun	Common Core Algebra I	I-11 - MC	LEVEL 4	S.ID.8	0.49
8	2014Jun	Common Core Algebra I	I-17 - MC	LEVEL 4	F.IF.2	0.49
9	2014Jun	Common Core Algebra I	I-18 - MC	LEVEL 4	F.IF.6	0.26
10	2014Jun	Common Core Algebra I	I-20 - MC	LEVEL 4	F.IF.1	0.53
11	2014Jun	Common Core Algebra I	I-23 - MC	LEVEL 4	A.CED.4	0.26
12	2014Jun	Common Core Algebra I	I-24 - MC	LEVEL 4	F.LE.2	0.27
13	2014Jun	Common Core Algebra I	II-26 - CR	LEVEL 4	F.LE.5	0.42
14	2014Jun	Common Core Algebra I	II-27 - CR	LEVEL 4	A.REI.3	0.33
15	2014Jun	Common Core Algebra I	II-30 - CR	LEVEL 4	F.IF.1	0.45
16	2014Jun	Common Core Algebra I	III-33 - CR	LEVEL 4	A.REI.4b	0.25
17	2014Jun	Common Core Algebra I	I-01 - MC	LEVEL 3	A.REI.1	0.58

### Item difficulty—an explanation

New York State assessments are, by design, composed of items of varying difficulty. That is done to identify student performance levels. A process called bookmarking is used when developing the assessments. NYSED has a well-defined and rather involved procedure to do that, but it boils down to this:

1. Questions are arranged in order of difficulty using the p-values measured in field testing.
2. Teacher evaluation committees determine the levels 2, 3, and 4 cut points (and level 5 for the common core regents exams) by identifying the most difficult question that two-thirds of the students just at the cusp of each proficiency level would be expected to answer correctly.
3. Cut scores are calculated by adding up all the points from the easiest question (the one with the highest p-value) to the question identified to be at each of the cut points (from step 2).

That is all done before the fact, i.e., prior to the assessment administration. NYSED does not indicate the question levels in any of its documentation, but we can look at aggregated student score results to reverse engineer the level setting.

### Reverse engineering the level setting

WNYRIC's level calculations are based on the premise that a student **at** the cut point will have a better than two out of three chance of correctly answering all of those questions to and including the difficulty level of the question at the cut point. This report shows the aggregated results of all of the students in our regional data set who scored **exactly at** the cut points. While it's not exactly the same thing, we're looking at the population of students scoring at the cut point and identifying the questions that two-thirds of them answered correctly. A student performing at level 5 would be expected to be able to answer all of the questions up to and including the question **at** the level 5 cut-point at least 67% of the time. (Note that this does not mean that a student whose score reached the cut point would be able to answer all of the level 5 questions correctly 67% of the time...just the easiest of them; the question **at** the cut point.)