

## Analysis of Embedded Field Test Items in Book1, Day 1 ELA

Prepared by Fred Smith, member of Change the Stakes and retired testing specialist for the New York City public schools

Here's a breakdown of the ELA tests in which Pearson has embedded multiple-choice items. My notes follow. The manner in which SED and Pearson have embedded items was unworkable from the outset (2012). Pearson, acting as a "consultant contractor," had to know the foreseeable problems inherent in contract specifications that could not possibly yield enough items needed to fulfill the objective of field testing, namely, the development of sound (i.e., reliable, valid, defensible) exams.

In the coming days, I will present evidence of defects and anomalies in the statewide exams, taken by 1.2 million students every year in grades 3 to 8, consequent on their poor design and development--defects that SED has hidden by operating without the level of transparency needed to analyze the quality of the instruments it has been developing and administering at the same time--in its hasty transition to the Common Core and its "core-aligned" exams.

What was done calls for an independent investigation of the contract, its execution and the relationship between the State Education Department and Pearson.

ELA Field Test (FT) Parameters - Embedded Multiple-Choice Items How Many Try-Out Passages and Items? - (Book 1, Day 1)						
2012						
Grade	# of FT Passages	# of FT Items	Item Placement	# of Items that Count	Total # of M-C items	% of Field Test Items
3	1	6	16-21	30	36	16.7%
4	1	6	16-21	31	37	16.2%
5	1	6	17-22	33	39	15.4%
6	1	6	17-22	33	39	15.4%
7	1	6	17-22	33	39	15.4%
8	1	6	17-22	27*	39	15.4%

\*In 2012, one passage (The Pineapple and the Hare) and six items were disqualified.

2013						
Grade	# of FT Passages	# of FT Items	Item Placement	# of Items that Count	Total # of M-C items	% of Field Test Items
3	1	6	7-12	24	30	20.0%
4	1	6	7-12	24	30	20.0%
5	2	14	8-13; 28-35	28	42	33.3%
6	2	14	8-14; 28-34	28	42	33.3%
7	2	14	8-14; 28-34	28	42	33.3%
8	2	14	8-14; 29-35	28	42	33.3%

2014 (Projected from 2013 and 2014 Teacher's Directions)						
Grade	# of FT Passages	# of FT Items	Item Placement	# of Items that Count	Total # of M-C items	% of Field Test Items

3	1	6	tbd	24	30	20.0%
4	1	6	tbd	24	30	20.0%
5	2	14	tbd	28	42	33.3%
6	2	14	tbd	28	42	33.3%
7	2	14	tbd	28	42	33.3%
8	2	14	tbd	28	42	33.3%

*Prepared by Fred Smith 3-29-14*

- Notes: 1 - Pearson/SED started embedding m-c items in NYS exams in 2012. They did this in 2013 and yesterday (4/1) and will do it again at the end of April (math). They continue to use only four forms of the test for embedding. This has yielded insufficient numbers of items for assembling subsequent exams. Do the math: Multiply # of FT items by 4 and these field tests generate results for between 24 and 56 items per grade--an inadequate pool from which to identify and draw good items for future tests.**
- 2 - This deficiency has led Pearson and SED to do stand-alone field testing in June, as they plan to do again this year. It is a flawed approach to test construction that has been accomplished by not notifying parents of their children's unknowing participation in a commercial test research and development project.**
- 3 - It is also extra work that was never specified in the \$32M contract.**
- 4 - The proportion of field test items has increased in relation to the total number of items on the test. FT items now comprise from one-fifth to one-third of the test (Book 1, Day 1), but still not enough to grow the next test crop.**
- 5 - And because of their placement in the middle of the test books, the embedded FT items suppress student performance on the items that count. Children are forced to spend their efforts on unrefined trial items likely**