

EUROPE ATP
2012

Berlin, Germany

GROWING
TALENT
IN EUROPE:

*Gaining
Advantage
Through
Assessment*

Realise Advantage

**Policies and Procedures for the
Development of More Effective Item
Writing and Test Development**

*Andrew Wiley
The College Board*

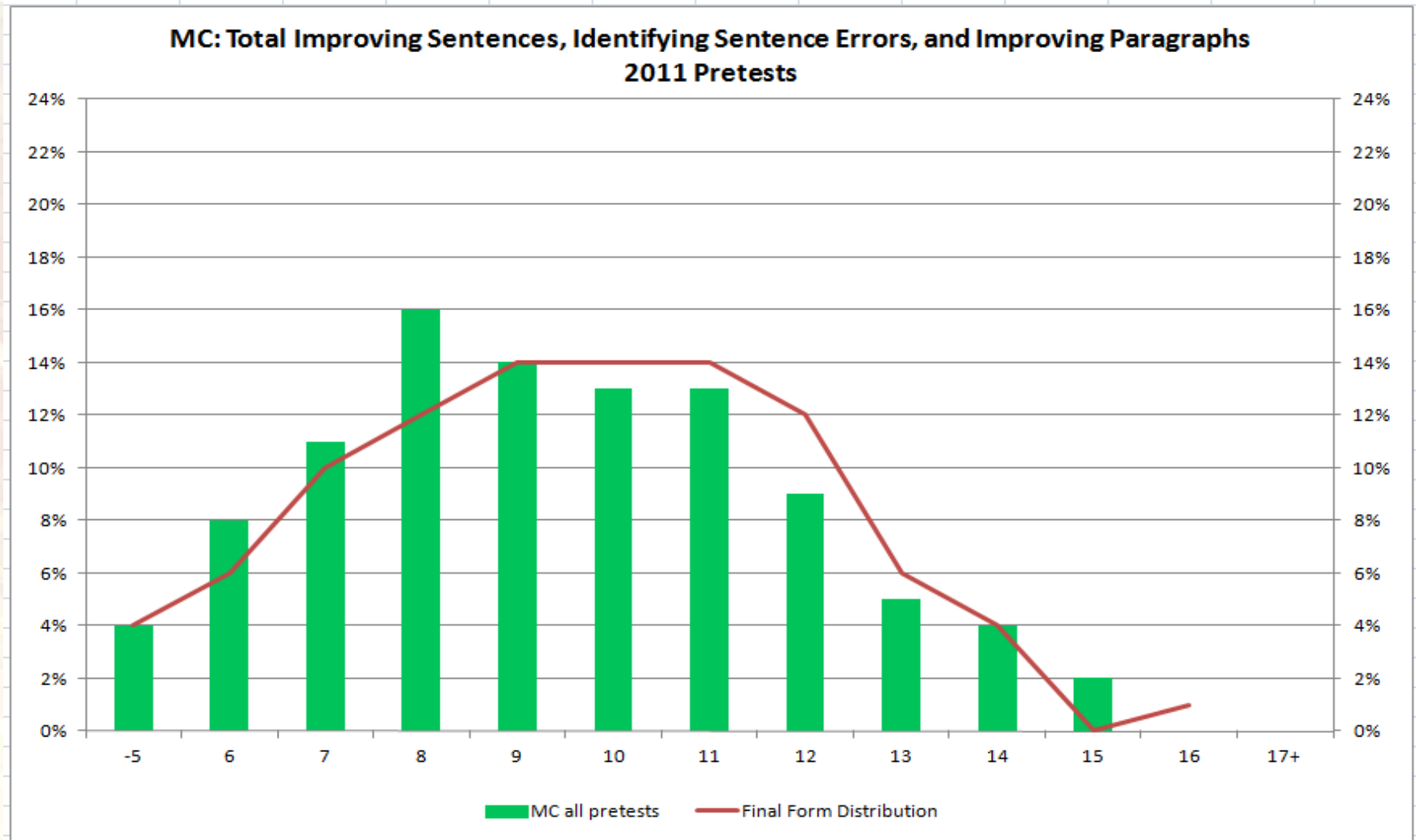
EDUCATE. SKILL. EMPLOY.



The Context

- High stakes assessment program
 - Multiple test administrations a year
 - Extensive item pools are required
- Extensive item development processes
 - item creation & review
 - item pretesting
- Significant labor cost
- Need for increased efficiency and improved quality

Snapshot of SAT Writing item pool compared to test difficulty distribution





Multiple Projects covering different perspectives

- Mathematics – development of automated item generation protocols with Mark Gierl
- Critical Reading – Application of Assessment Engineering principles with Ric Luecht
- Writing - Application of Principled Assessment design with internal staff



A Solution: Principled Assessment Design

- Principled assessment design:
 - Clearly defined measurement constructs and test purpose
 - Having an item attribute taxonomy and item models to describe how item attributes, components, and features interact to drive item difficulty
 - Develop a taxonomy of item attributes and item models
 - create detailed test specifications
 - guide item development
 - guide form assembly



Principled Assessment Design: Expected Outcomes

- Items are tightly anchored in test specifications, definition of construct, and purpose of test
- Item writers have better understanding on what drives item difficulty
- Item writers should be able to create a greater percentage of items that pass pretesting that address shallow areas in item pools



Pilot Project: Two Phases

Phase 1 – Analysis of current item pool

- Performance of current pretest items
- Ability to predict item difficulty

Phase 2 – Development of item models

- Systematic review by SMEs
- Identification of key item characteristics
- Evaluate ability to predict item difficulty



Phase 1 – Current item pools

- Acceptance rates of pretested SAT Writing items, disaggregated by estimated difficulty levels

	# of items	Acceptable	Do Not use	Review
Total	490	82.4%	10.6%	6.9%
Difficult Estimation				
Easy	115	90.4%	9.6%	0.0%
Medium	178	84.3%	9.6%	6.2%
Hard	105	77.1%	18.1%	4.8%

Phase 1 – Current item pools

- Distribution of difficulty level for pretested SAT Writing items, disaggregated by estimated difficulty level

	# of items	Mean (SD) Equated Delta	Actual Difficulty		
			Easy	Medium	Hard
Total	490	10.3 (2.5)	23.7%	46.5%	29.8%
Difficult Estimation					
Easy	115	9.0 (2.0)	40.9%	47.0%	12.2%
Medium	178	10.2 (2.3)	24.7%	45.5%	29.8%
Hard	105	11.9 (2.5)	10.5%	40.0%	49.5%



Phase 2:

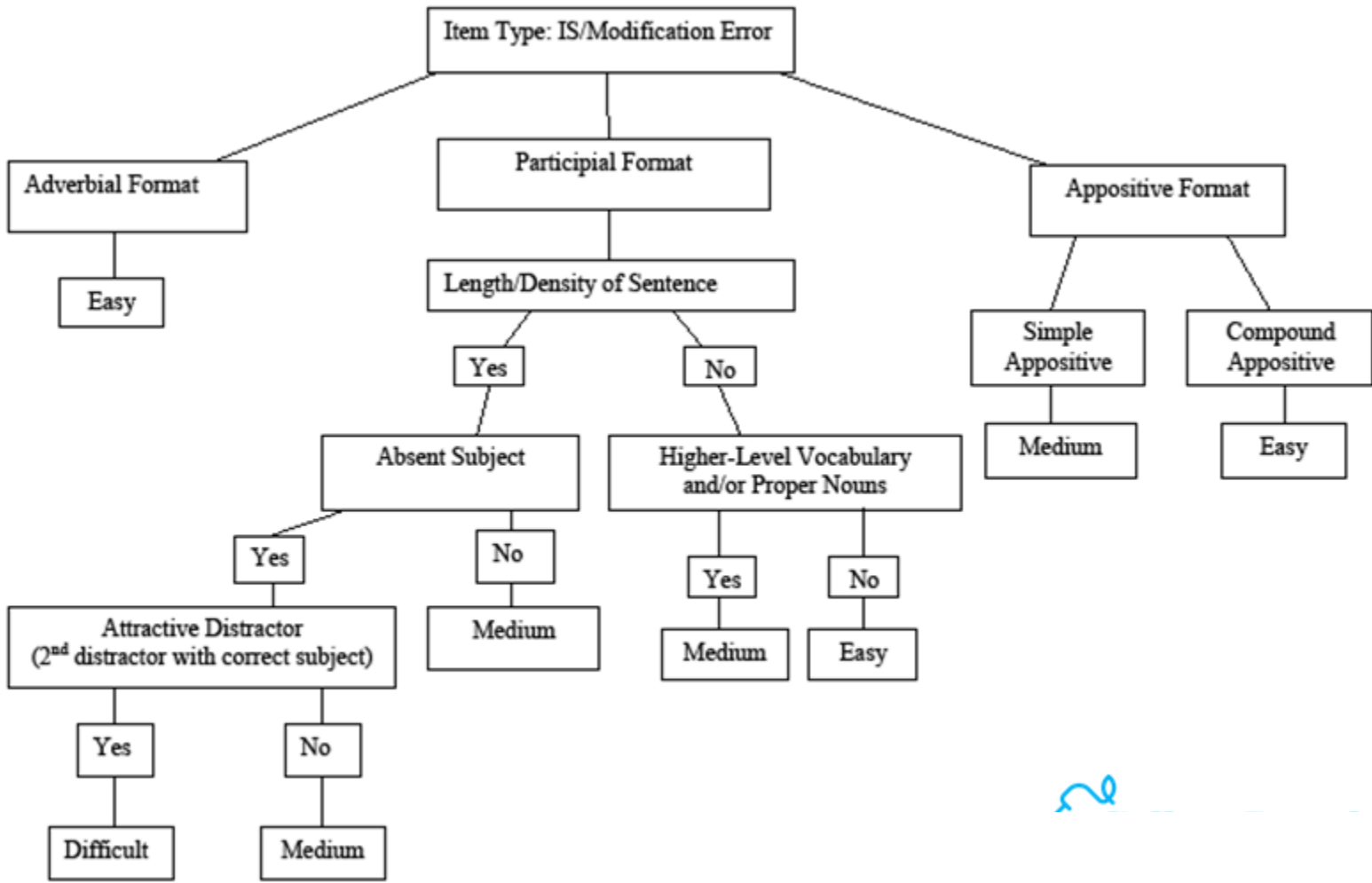
Development of item models

Our ELA content specialists:

- Identified common features and attributes hypothesized to drive item difficulty for a sample of items.
- Developed a theoretical model of how features and attributes interact & result in the level of difficulty.
- Reviewed 29 items & used the model to predict item difficulty.
- Consistency of predictions between specialists was evaluated.
- Accuracy of model based predictions were compared to subjective predictions already available



Pilot Project: Item model



EDUCATE. SKILL. EMPLOY.



Example Item & Item Model

Finding its way through sediment on the sea floor, the 400-700 whiskers on a walrus' snout are used to sense the presence of clams and other invertebrates on which it feeds

- A) the 400-700 whiskers on a walrus' snout are used
- B) the walrus uses the 400-700 whiskers on its snout
- C) walruses use the 400-700 whiskers on their snout
- D) a walrus' snout, having 400-700 whiskers, is used
- E) the snout of walruses use 400-700 whiskers



Example Item & Item Model

Finding its way through sediment on the sea floor, the 400-700 whiskers on a walrus' snout are used to sense the presence of clams and other invertebrates on which it feeds

Item Type: Improving Sentences
Modification error

- A) the 400-700 whiskers on a walrus' snout are used
- B) the walrus uses the 400-700 whiskers on its snout
- C) walruses use the 400-700 whiskers on their snout
- D) a walrus' snout, having 400-700 whiskers, is used
- E) the snout of walruses use 400-700 whiskers



Example Item & Item Model

Finding its way through sediment on the sea floor, the 400-700 whiskers on a walrus' snout are used to sense the presence of clams and other invertebrates on which it feeds

- A) the 400-700 whiskers on a walrus' snout are used
- B) the walrus uses the 400-700 whiskers on its snout
- C) walruses use the 400-700 whiskers on their snout
- D) a walrus' snout, having 400-700 whiskers, is used
- E) the snout of walruses use 400-700 whiskers

Item Type: Improving Sentences
Modification error



What is the format of the item?



Example Item & Item Model

Finding its way through sediment on the sea floor, the 400-700 whiskers on a walrus' snout are used to sense the presence of clams and other invertebrates on which it feeds

- A) the 400-700 whiskers on a walrus' snout are used
- B) the walrus uses the 400-700 whiskers on its snout
- C) walruses use the 400-700 whiskers on their snout
- D) a walrus' snout, having 400-700 whiskers, is used
- E) the snout of walruses use 400-700 whiskers

Item Type: Improving Sentences
Modification error

What is the format of the item?

The introductory clause is a participial phrase.



Example Item & Item Model

Finding its way through sediment on the sea floor, the 400-700 whiskers on a walrus' snout are used to sense the presence of clams and other invertebrates on which it feeds

Item Type: Improving Sentences
Modification error

↓
Participial format

- A) the 400-700 whiskers on a walrus' snout are used
- B) the walrus uses the 400-700 whiskers on its snout
- C) walruses use the 400-700 whiskers on their snout
- D) a walrus' snout, having 400-700 whiskers, is used
- E) the snout of walruses use 400-700 whiskers



Example Item & Item Model

Finding its way **through sediment** on the sea floor, the 400-700 whiskers on a walrus' snout are used to sense the presence of **clams** and other invertebrates **on which it feeds**

Item Type: Improving Sentences
Modification error

Participial format

Is the stem lengthy or dense?

- A) the 400-700 whiskers on a walrus' snout are used
- B) the walrus uses the 400-700 whiskers on its snout
- C) walruses use the 400-700 whiskers on their snout
- D) a walrus' snout, having 400-700 whiskers, is used
- E) the snout of walruses use 400-700 whiskers



Example Item & Item Model

Finding its way **through sediment** on the sea floor, the 400-700 whiskers on a walrus' snout are used to sense the presence of **clams** and other invertebrates **on which it feeds**

- A) the 400-700 whiskers on a walrus' snout are used
- B) the walrus uses the 400-700 whiskers on its snout
- C) walruses use the 400-700 whiskers on their snout
- D) a walrus' snout, having 400-700 whiskers, is used
- E) the snout of walruses use 400-700 whiskers

Item Type: Improving Sentences
Modification error

Participial format

Is the stem lengthy or dense?

Yes. Prepositional phrases add to the density of the stem.



Example Item & Item Model

Finding its way through sediment on the sea floor, the 400-700 whiskers on a walrus' snout are used to sense the presence of clams and other invertebrates on which it feeds

- A) the 400-700 whiskers on a walrus' snout are used
- B) the walrus uses the 400-700 whiskers on its snout
- C) walruses use the 400-700 whiskers on their snout
- D) a walrus' snout, having 400-700 whiskers, is used
- E) the snout of walruses use 400-700 whiskers

Item Type: Improving Sentences
Modification error

Participial format

Length/Density of Sentence - Yes



Example Item & Item Model

Finding its way through sediment on the sea floor, the 400-700 whiskers on a walrus' snout are used to sense the presence of clams and other invertebrates on which it feeds

- A) the 400-700 whiskers on a walrus' snout are used
- B) the walrus uses the 400-700 whiskers on its snout
- C) walruses use the 400-700 whiskers on their snout
- D) a walrus' snout, having 400-700 whiskers, is used
- E) the snout of walruses use 400-700 whiskers

Item Type: Improving Sentences
Modification error

Participial format

Length/Density of Sentence - Yes

Is the subject absent?



Example Item & Item Model

Finding its way through sediment on the sea floor, the 400-700 whiskers on a walrus' snout are used to sense the presence of clams and other invertebrates on which it feeds

- A) the 400-700 whiskers on a walrus' snout are used
- B) the walrus uses the 400-700 whiskers on its snout
- C) walruses use the 400-700 whiskers on their snout
- D) a walrus' snout, having 400-700 whiskers, is used
- E) the snout of walruses use 400-700 whiskers

Item Type: Improving Sentences
Modification error

Participial format

Length/Density of Sentence - Yes

Is the subject absent?

Yes. The intended subject "walrus" does not appear in the sentence as a noun. Students must form the intended subject from the possessive "walrus' snout."



Example Item & Item Model

Finding its way through sediment on the sea floor, the 400-700 whiskers on a walrus' snout are used to sense the presence of clams and other invertebrates on which it feeds

- A) the 400-700 whiskers on a walrus' snout are used
- B) the walrus uses the 400-700 whiskers on its snout
- C) walruses use the 400-700 whiskers on their snout
- D) a walrus' snout, having 400-700 whiskers, is used
- E) the snout of walruses use 400-700 whiskers

Item Type: Improving Sentences
Modification error

Participial format

Length/Density of Sentence - Yes

Absent Subject / Yes



Example Item & Item Model

Finding its way through sediment on the sea floor, the 400-700 whiskers on a walrus' snout are used to sense the presence of clams and other invertebrates on which it feeds

- A) the 400-700 whiskers on a walrus' snout are used
- B) the walrus uses the 400-700 whiskers on its snout
- C) walruses use the 400-700 whiskers on their snout
- D) a walrus' snout, having 400-700 whiskers, is used
- E) the snout of walruses use 400-700 whiskers

Item Type: Improving Sentences
Modification error

Participial format

Length/Density of Sentence - Yes

Absent Subject / Yes

How attractive are the distractors?



Example Item & Item Model

Finding its way through sediment on the sea floor, the 400-700 whiskers on a walrus' snout are used to sense the presence of clams and other invertebrates on which it feeds

- A) the 400-700 whiskers on a walrus' snout are used
- B) the walrus uses the 400-700 whiskers on its snout
- C) walruses use the 400-700 whiskers on their snout
- D) a walrus' snout, having 400-700 whiskers, is used
- E) the snout of walruses use 400-700 whiskers

Item Type: Improving Sentences
Modification error

Participial format

Length/Density of Sentence - Yes

Absent Subject / Yes

How attractive are the distractors?

None of the distractors start with the correct modified noun/subject "walrus."

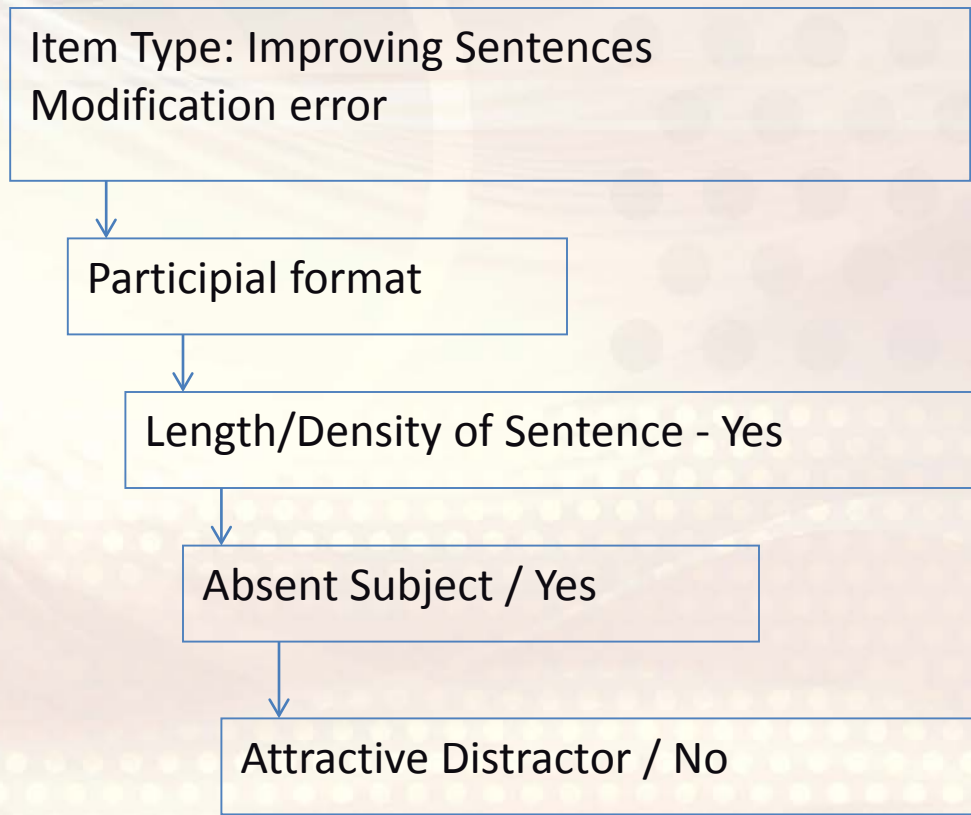




Example Item & Item Model

Finding its way through sediment on the sea floor, the 400-700 whiskers on a walrus' snout are used to sense the presence of clams and other invertebrates on which it feeds

- A) the 400-700 whiskers on a walrus' snout are used
- B) the walrus uses the 400-700 whiskers on its snout
- C) walruses use the 400-700 whiskers on their snout
- D) a walrus' snout, having 400-700 whiskers, is used
- E) the snout of walruses use 400-700 whiskers

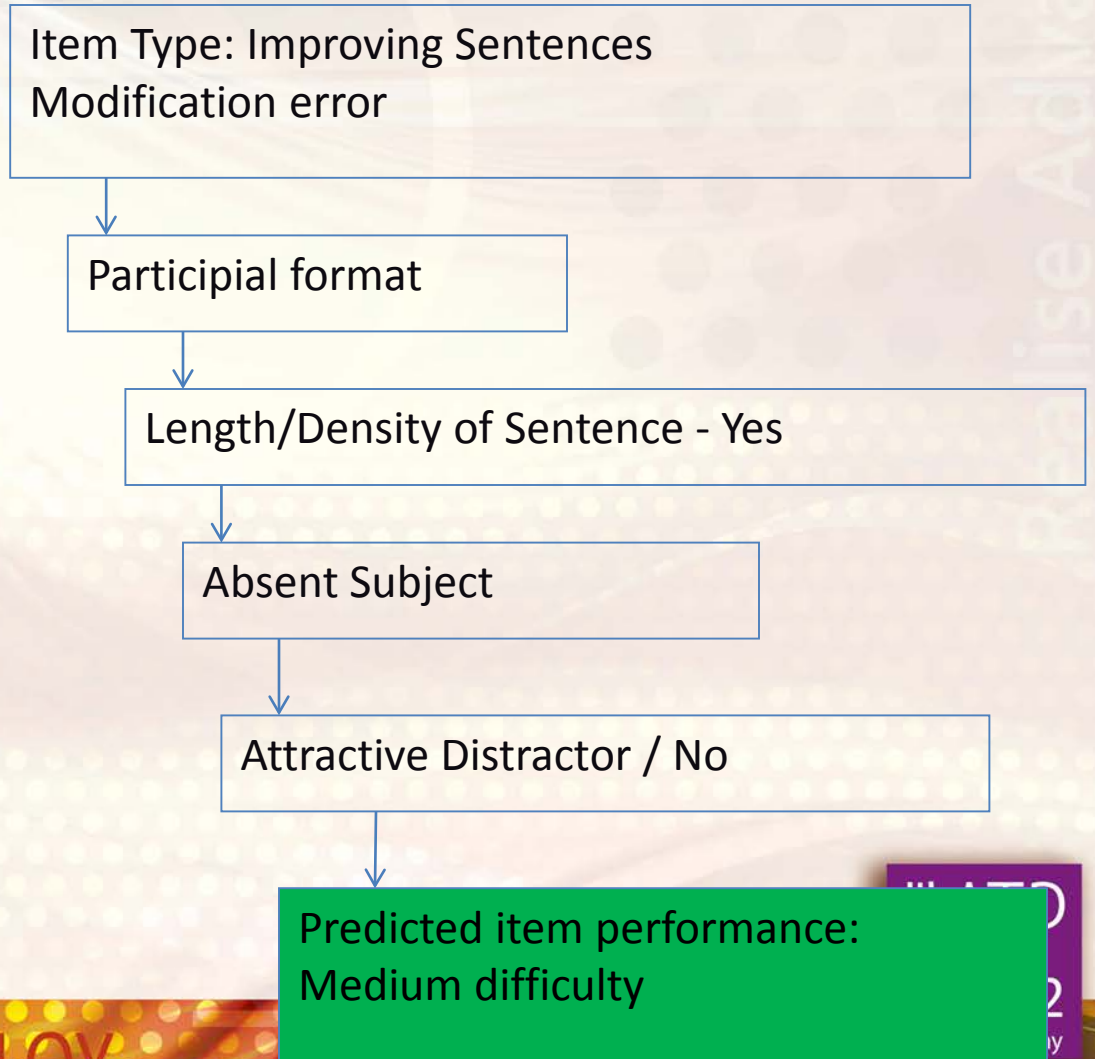




Example Item & Item Model

Finding its way through sediment on the sea floor, the 400-700 whiskers on a walrus' snout are used to sense the presence of clams and other invertebrates on which it feeds

- A) the 400-700 whiskers on a walrus' snout are used
- B) the walrus uses the 400-700 whiskers on its snout
- C) walruses use the 400-700 whiskers on their snout
- D) a walrus' snout, having 400-700 whiskers, is used
- E) the snout of walruses use 400-700 whiskers



Predictions of item difficulty

- Traditional item difficulty estimates and actual item difficulty classifications

		Actual Difficulty levels		
		Easy	Medium	Hard
	# of items			
Traditional Difficulty Estimates				
Easy	2	0.0%	100.0%	0.0%
Medium	22	18.2%	68.2%	13.6%
Hard	5	0.0%	60.0%	40.0%

Predictions of item difficulty

- Model based item difficulty estimates and actual item difficulty classifications

		Actual Difficulty levels		
	# of items	Easy	Medium	Hard
Model-Based Difficulty Estimates				
Easy	10	20.0%	70.0%	10.0%
Medium	17	5.9%	70.6%	23.5%
Hard	2	50.0%	50.0%	0.0%



Predictions of item difficulty

- Model based estimates do not appear to be any more accurate than the traditional difficulty estimates
 - Traditional – 68% of medium predictions were accurate
 - Model based – 70% of medium predictions were accurate
- Model based predictions were more likely to predict easy difficulty levels
 - 10 items predicted to be easy from the model based, as compared to only 5 in the traditional estimates



Some positive impacts

- Greater understanding of the items investigated
- Reviewers felt more qualified to conduct further reviews of these item types
- More qualified to try to build models for other item types
- Gained valuable insight to be used in the review of items with students



Some next steps

- Focus groups and “think alouds” with students
- Revise model based upon initial findings and information gained from student interviews
- Attempt to build other item models within SAT Writing
- Attempt to build models with other content areas (Reading and Mathematics)

Thank You!

**Please Complete Your Evaluation and
Hand to the Session Coordinator.**

**Andrew Wiley
Awiley@collegeboard.org**

EDUCATE. SKILL. EMPLOY.

