


# Moving beyond Efficiency to Allow CAT to Provide Better Diagnostic Information



**Brian D. Bontempo, Ph.D.**  
***Mountain Measurement, Inc.***  
***[www.mountainmeasurement.com](http://www.mountainmeasurement.com)***



# Background



- As CAT researchers and practitioners, our focus is typically on how to build more efficient CATs.
- For a moment, let's focus on what do we do with the CAT data.



# Common Test Score Report



John Doe

Oct 1, 2011

**PASS**



# The Future




<http://scores.espn.go.com/mlb/boxscore?gameId=310928130>

**Final in 12** Series: Game 3 of 3 ▼



**Yankees 7**  
(97-65, 45-36 away)



**Rays 8**  
(91-71, 47-34 home)

Check-in **463** Alerts 

**REAL STEEL** IN THEATRES AND IMAX 10.07.11

SEE HOW HP AND INTEL® BRING REAL STEEL TO LIFE

CHECK IT OUT




THE HP TOUCHSMART 520XT WITH VISIBLY SMART 2ND GEN INTEL® CORE™ PROCESSORS





©2011 DreamWorks II Distribution Co, LLC


7:00 PM ET, September 28, 2011  
Tropicana Field, St. Petersburg, Florida

	2	3	4	5	6	7	8	9	10	11	12	R	H	E	W: J. McGee (5-2)	L: S. Proctor (0-3)
<b>NY Yankees</b>	4	0	1	1	0	0	0	0	0	0	0	7	9	1		
<b>TB Rays</b>	0	0	0	0	0	0	6	1	0	0	1	8	10	1		

- Recap
- Box Score
- Play-By-Play
- Photos **88**
- Conversation 

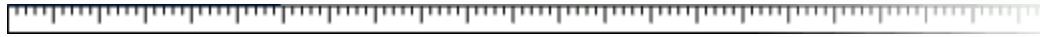
 **New York Yankees**

Hitters	AB	R	H	RBI	BB	SO	#P	AVG	OBP	SLG
<u>D Jeter</u> SS	3	1	0	0	1	1	23	.297	.355	.388
<u>R Pena</u> 2B	2	0	0	0	0	2	13	.100	.159	.175
<u>C Granderson</u> CF	3	1	1	0	0	1	7	.262	.364	.552
<u>a-G Golson</u> PH-CF-LF	3	0	1	0	0	0	15	.182	.250	.182
<u>M Teixeira</u> 1B	4	2	2	5	0	0	17	.248	.341	.494
<u>E Chavez</u> 3B	2	0	1	0	0	1	6	.263	.320	.356
<u>R Cano</u> DH	4	0	0	0	0	0	21	.302	.349	.533
<u>b-J Posada</u> PH-DH	2	0	0	0	0	1	7	.235	.315	.398
<u>N Swisher</u> RF	3	0	1	0	1	0	17	.260	.374	.449

 **Tampa Bay Rays**

Hitters	AB	R	H	RBI	BB	SO	#P	AVG	OBP	SLG
<u>D Jennings</u> LF	6	0	0	0	1	1	27	.259	.356	.449
<u>B Upton</u> CF	4	0	0	1	2	1	27	.243	.331	.429
<u>E Longoria</u> 3B	5	2	2	4	2	1	31	.244	.355	.495
<u>M Joyce</u> RF	3	0	0	0	0	2	9	.277	.347	.478
<u>b-R Canzler</u> PH	1	0	0	0	0	1	5	.333	.400	.333
<u>J Ruggiano</u> RF	0	0	0	0	0	0	0	.248	.273	.400
<u>c-J Jaso</u> PH	1	0	1	0	0	0	1	.224	.298	.354
<u>E Johnson</u> PR	0	0	0	0	0	0	0	.194	.257	.338
<u>J Lobaton</u> C	1	0	0	0	0	1	3	.118	.231	.147

# Types of Reports



- Examinee Reports
- Group Reports (School, Company, Jurisdiction)
- Psychometric Reports



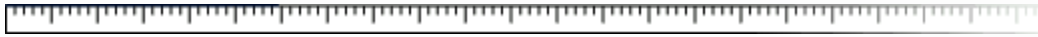
# Evolution of Reporting - Examinee Content

---

- Pass/Fail Decision -> Scores
- Achievement -> Growth
- Raw/Percent Scores -> Universal Ability Estimates
- Overall Score -> Content Subscores



# Issues in CAT Examinee Reports



- When modeled unidimensionally, subscores have limited utility
- Error associated with subscore ability estimates is high because there are not enough items in each content area
- Item selection is not optimal because it is based on efficiency of overall score
- Person fit is useless



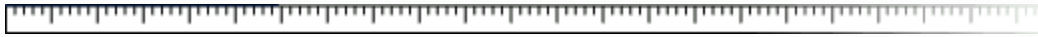
# Solutions to CAT Examinee Reports



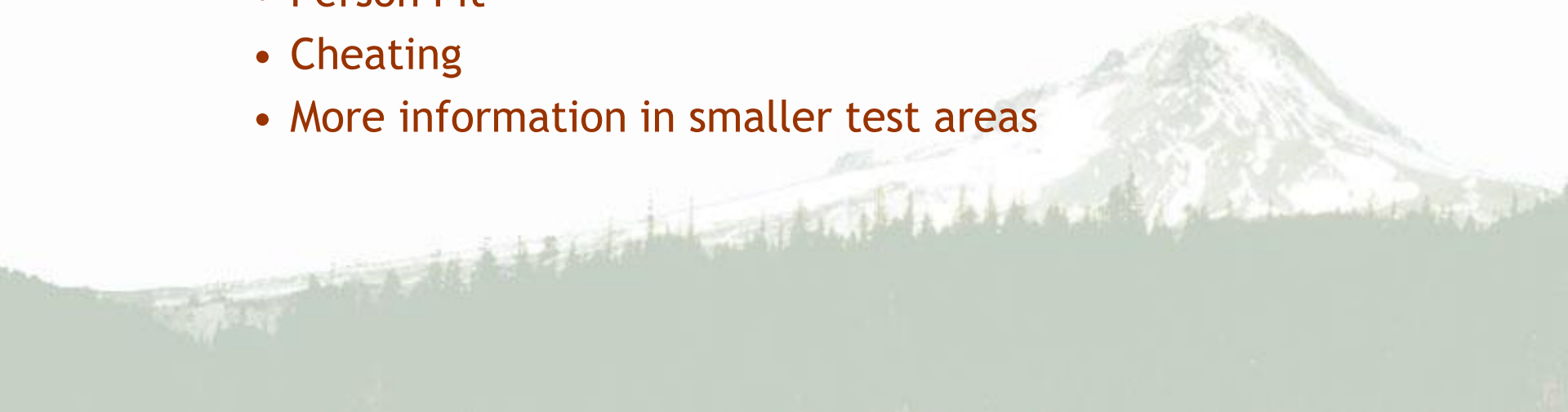
- When modeled unidimensionally, content area scores have limited utility
  - Increase Randomization in Item Selection
  - MIRT for subscores
  - Multiple unidimensional scales for subscores
- Not enough items in each content area
  - Non-mutually exclusive content categorization
  - Implement Multiple Content “Schema”



# Solutions to CAT Examinee Reports



- Item selection is not optimal because it is based on efficiency of overall score
- Person fit is useless
  - Smart CATs which add some items that are intentionally selected to
    - ISV
    - Person Fit
    - Cheating
    - More information in smaller test areas



# Research Questions for CAT Examinee Reports

---

- What are the strengths and weaknesses associated with each model for adding more items?
  - Use all items in overall score (But this violates the test blueprint)
  - Select item responses at random from each content area to meet blueprint
  - Select items responses intentionally (highest ISV) from each content area to meet blueprint
  - Select the first set of items administered that meet blueprint

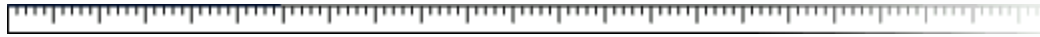
# Evolution of Reporting - Group Content

---

- Passing Rates -> Mean Scores
- Descriptive Statistics -> Distribution
- Normative -> Criterion Referenced
- Percentile Ranks -> Distance to Goal (e.g., cutscore)
- Comparisons -> Profiling



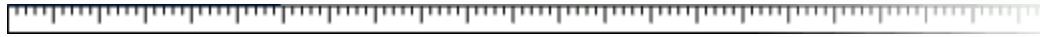
# Issues in CAT Group Reports



- Individual subscore ability estimates typically have large amounts of error
- Group sizes may be small



# Solutions to CAT Group Reports



- Aggregate the data over the group and calculate subscore ability estimates for the group directly



# Research Questions for CAT Group Reports

- What are the parameters around “group ability estimates” being more accurate than the median of individual ability estimates when the sample sizes are small?
  - How many items per person or SEM per person
  - How many examinees per group



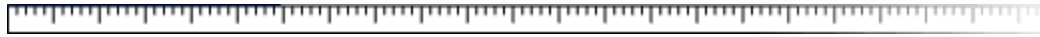
# Evolution of Reporting - Psychometric Content

---

- Error
- Fit
- Cheating
- Speed/Time
- Motivation
- Validity



# Issues in CAT Psychometric Reports



- Traditional item quality statistics have limited usefulness





# Solutions to CAT Psychometric Reports

- Implement item selection algorithms that increase the dispersion of the ability of the examinees taking each item



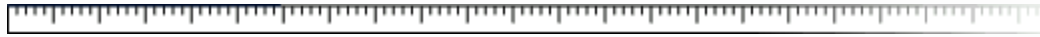
# Research Questions for CAT Psychometric Reports

---

- How can the progressive item selection Algorithm be used to improve dispersion?
  - Bontempo, Kingsbury & Zara (2010)
- Are there other item selection algorithm approaches that might improve dispersion?



# The Future of Test Reporting

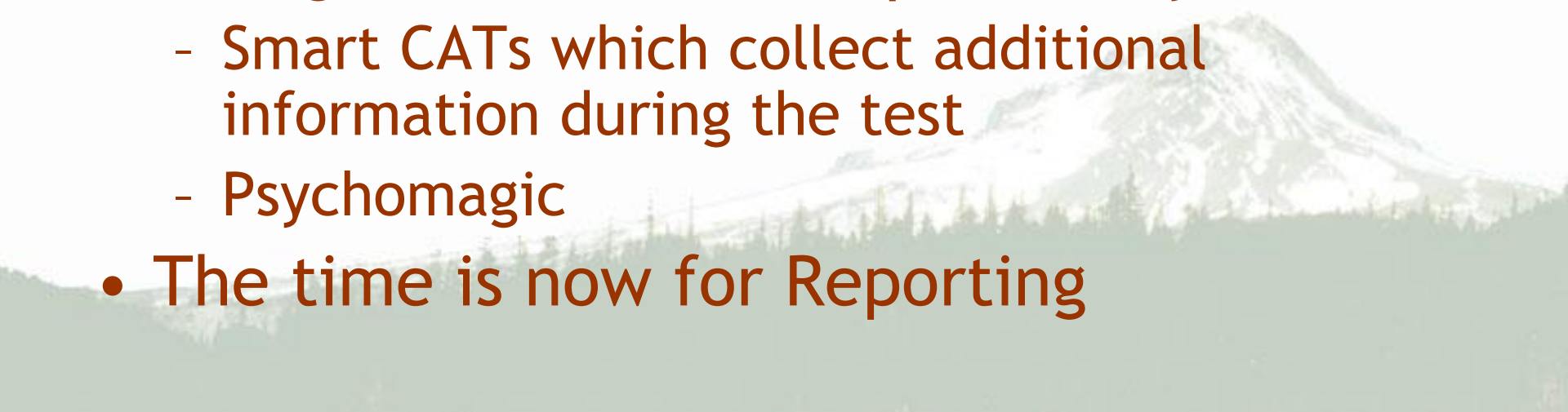


- Data Visualization
- Interactivity
- Data Driven Interpretation
- MULTImedia
- Integration



# Conclusion



- Future CATs will provide better diagnostic information to
    - Examinees
    - Regulators, Educators, Employers
    - Test Developers
  - This goal will be accomplished by
    - Smart CATs which collect additional information during the test
    - Psychomagic
  - The time is now for Reporting
- 



**Thank you for your kind  
attention**

**[brian@mountainmeasurement.com](mailto:brian@mountainmeasurement.com)**

