

A Different DIF Study: Psychometric Examination of a Technology Enhanced Item Type

- ▶ Presented by Cameron Clyne, M.A.
- ▶ Center for Educational Testing and Evaluation
- ▶ University of Kansas

Technology Enhanced Items (TE)

- ▶ What are Technology Enhanced Items?
 - ▶ Parshall et al. (2002)
 - ▶ “items that depart from the traditional, discrete, text-based, multiple-choice format.”
- ▶ Potential Benefits of TE Items
 - ▶ Have increased fidelity
 - ▶ Reduced guessing
 - ▶ More construct less real estate

Are TE Items Better?

- ▶ Huff and Sireci (2001)
 - ▶ Too much focus on finding new TE types, rather than validating the types we have.
- ▶ Not enough focus on comparison with conventional item types.
- ▶ Issue
 - ▶ Viewing TE items as a single item type, instead of focusing on the differences between different types of technology enhanced items.

Editing TE Item Prototype

<http://media.cete.us/dlm/cpass/>



Method

- ▶ Steps
 - ▶ Two TE editing items were created.
 - ▶ “Equivalent” multiple choice items were created.
 - ▶ Items were piloted to CTE students.
 - ▶ Total N= 870

Method

- ▶ BILOG-MG was run using a 2-PL model.
- ▶ DIFAS
 - ▶ Categories were created every 0.5 theta, ranging from -3 to 3.
 - ▶ MH and ETS DIF stats were reviewed.
- ▶ SPSS
 - ▶ Logistic regression was run to further test for DIF.

Results

- ▶ No clear picture!
 - ▶ Overall, the items were not similar in difficulty.
 - ▶ Length may have played a role in the difficulty level.
 - ▶ First item had 193 words
 - ▶ Second item had 134 words
 - ▶ But...there is DIF!
 - ▶ DIF occurred in most items, and normally in the same direction (favoring MC).
 - ▶ Certain items appeared to be overly difficult, causing issues with item statistics.

DIFAS

Item Name	MH CHI	MH LOR	ETS
Item 1a	5.6517	-0.4486	B
Item 1b	182.9132	3.4479	C
Item 1c	195.5438	3.9768	C
Item 1d	37.4449	1.2190	C
Item 1e	207.3861	3.2411	C
Item 2a	13.4004	-0.7259	B
Item 2b	178.4939	3.1185	C
Item 2c	60.0974	1.6322	C
Item 2d	160.6889	3.1669	C
Item 2e	3.0008	0.3616	A

Logistic Regression

Item Name	B	Log Odds	Nagelkerke R ²	P-Value	DIF
Item 1a	.487	1.627	.175	0.00755**	No DIF
Item 1b	-3.564	.028	.509	<0.001***	Uniform
Item 1c	-3.815	.022	.592	<0.001***	Uniform
Item 1d	-1.245	.288	.302	<0.001***	Uniform
Item 1e	-3.477	.031	.525	<0.001***	Uniform
Item 2a	.650	1.915	.189	0.001	Uniform
Item 2b	-3.039	.048	.473	<0.001***	Uniform
Item 2c	-1.636	.195	.363	<0.001***	Uniform
Item 2d	-3.272	.038	.518	<0.001***	Uniform
Item 2e	-.332	.717	.291	0.074	No DIF

Results

Both of these items assessed a students ability to identify a missing “s” within the sentence.

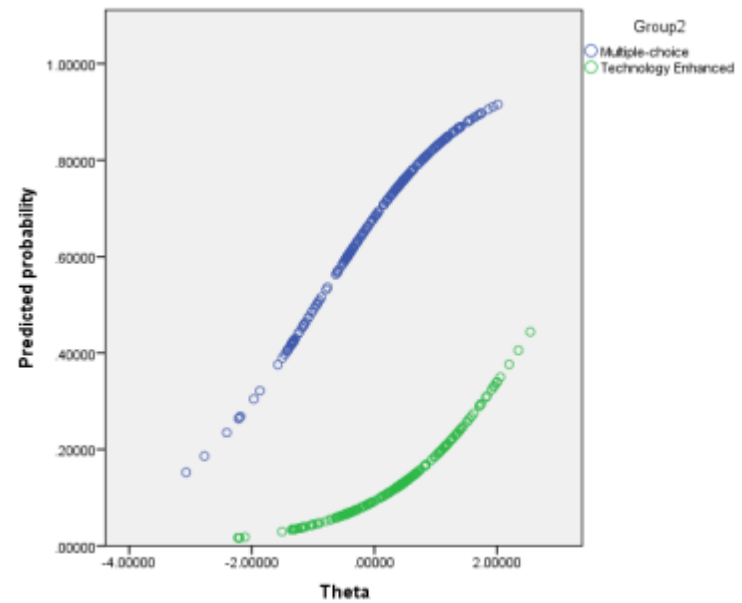
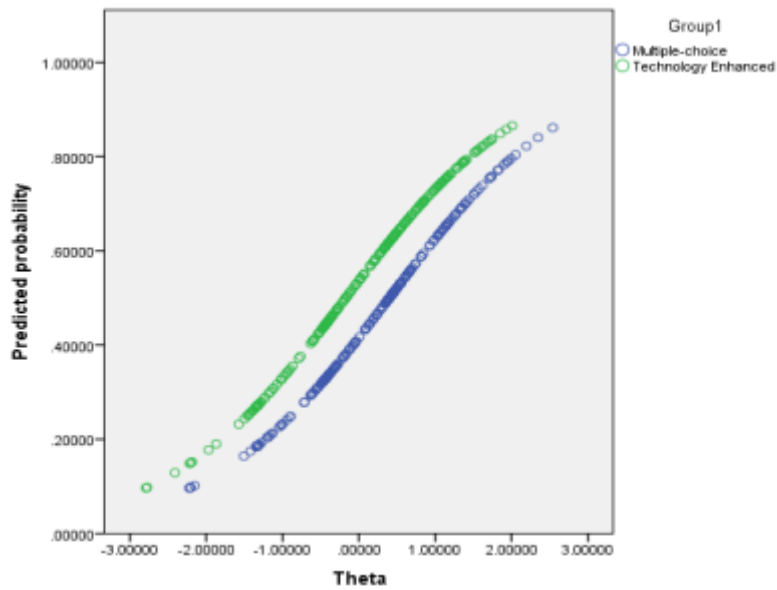
Item 1a: Grammar Item

Type	P-value	Item-total Corr	a	b
MC	43.3	.251	.395	.481
TE	53.4	.371	.545	-.154
DIF	Statistic	p-Value		
MH LOR	-.04486			
ETS	B			
Logistic Regression	1.627	0.00755**		

Item 2b: Grammar Item

Type	P-value	Item-total Corr	a	b
MC	66.1	.268	.431	-1.019
TE	11.4	.219	.505	2.715
DIF	Statistic	p-Value		
MH LOR	3.1185			
ETS	C			
Logistic Regression	.048	< 0.001***		

Grammar Items Compared



Results

Both of these items assessed a students ability to identify a missing capitalization within a sentence.

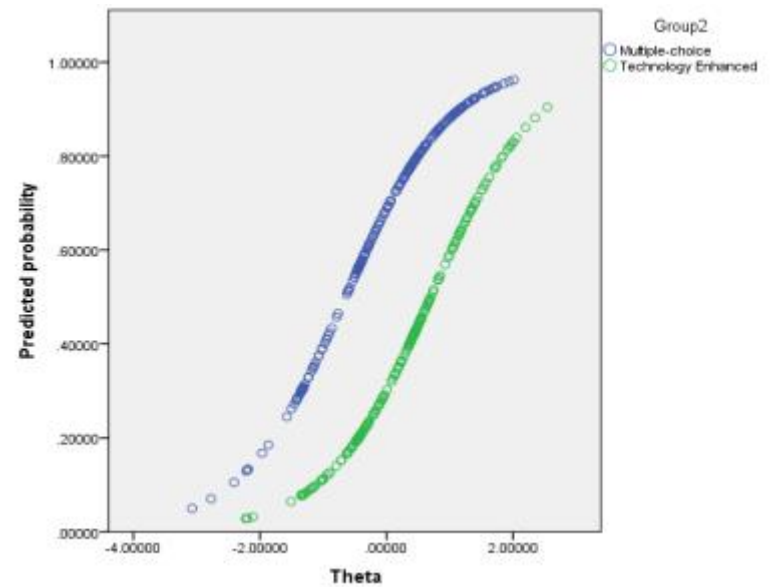
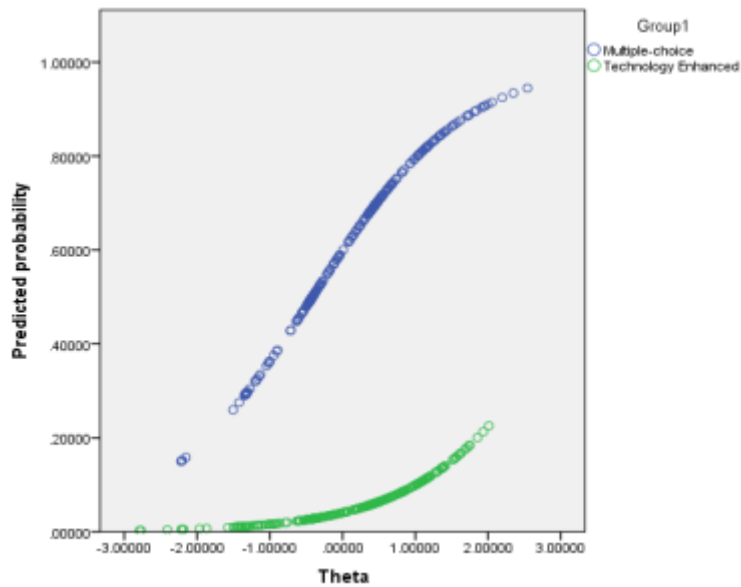
Item 1b: Capitalization

Type	P-value	Item-total Corr	a	b
MC	59.4	.322	.536	-.441
TE	5.6	.139	.475	3.817
DIF	Statistic	p-Value		
MH LOR	3.4479			
ETS	C			
Logistic Regression	.028	<0.001***		

Item 2c: Capitalization

Type	P-value	Item-total Corr	a	b
MC	65.0	.486	.794	-.621
TE	34.4	.324	.520	.878
DIF	Statistic	p-Value		
MH LOR	1.6322			
ETS	C			
Logistic Regression	.195	<0.001***		

Capitalization Items Compared



Results

Both of these items assessed a students ability to identify a spelling error within a sentence.

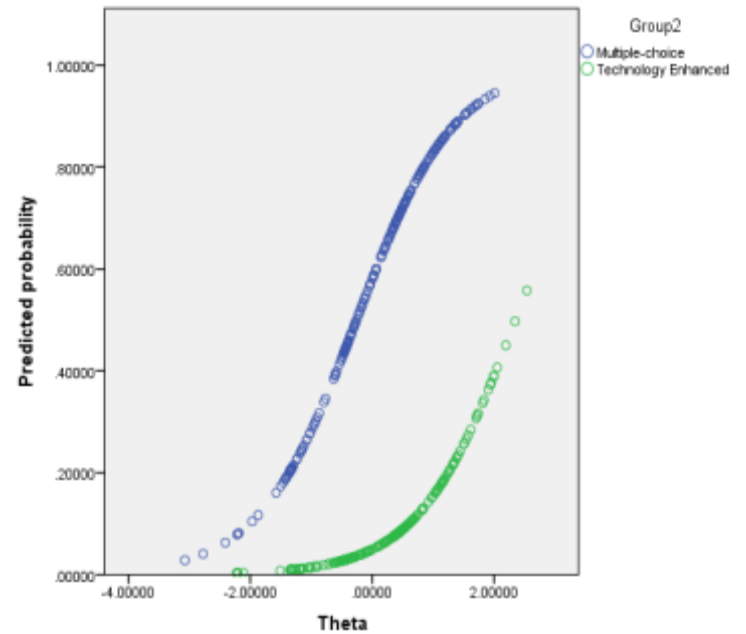
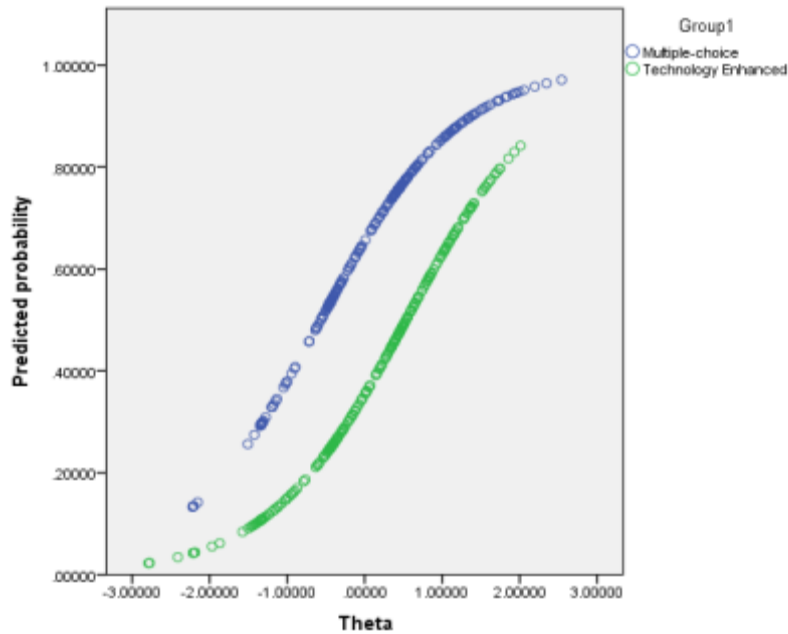
Item 1d: Spelling

Type	P-value	Item-total Corr	a	b
MC	63.8	.316	.549	-.665
TE	38.4	.406	.658	.558
DIF	Statistic	p-Value		
MH LOR	1.219			
ETS	C			
Logistic Regression	.288	<0.001***		

Item 2d: Spelling

Type	P-value	Item-total Corr	a	b
MC	56.5	.464	.754	-.256
TE	8.3	.157	.491	3.212
DIF	Statistic	p-Value		
MH LOR	3.1669			
ETS	C			
Logistic Regression	.038	<0.001***		

Spelling Items Compared



Results

Both of these items assessed a students ability to identify an error in punctuation within a sentence.

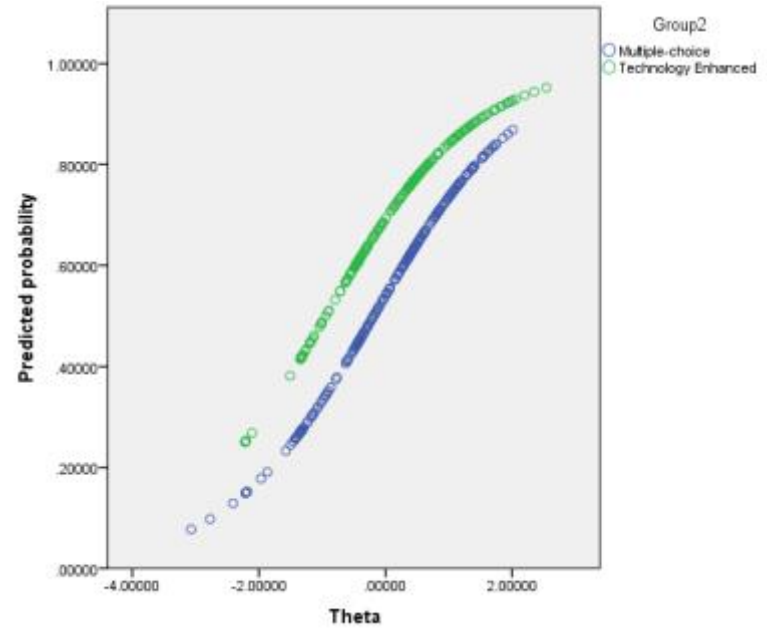
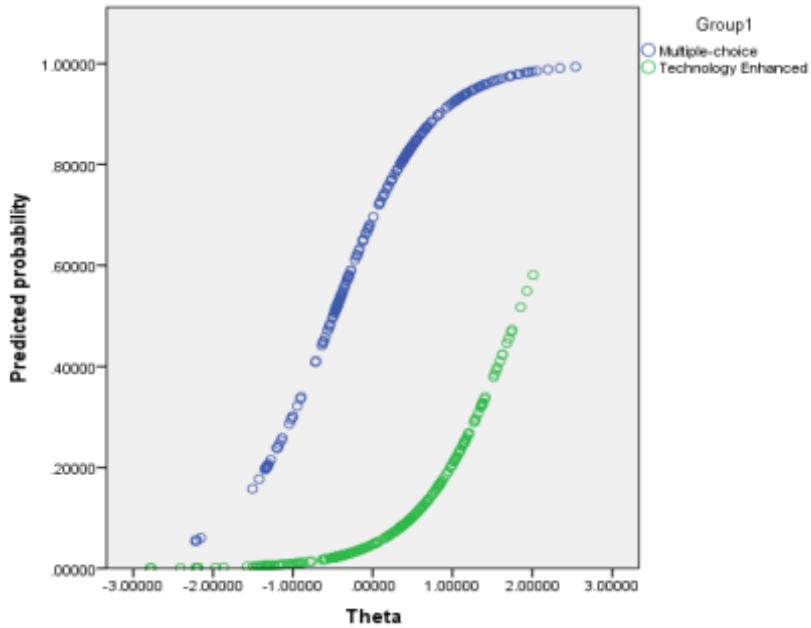
Item 1c: comma

Type	P-value	Item-total Corr	a	b
MC	67.8	.321	.574	-.851
TE	10.0	.232	.652	2.411
DIF	Statistic	p-Value		
MH LOR	3.2411			
ETS	C			
Logistic Regression	.022	< 0.001		

Item 2a: comma

Type	P-value	Item-total Corr	a	B
MC	65.0	.486	.794	-.621
TE	67.7	.344	.600	-.837
DIF	Statistic	p-Value		
MH LOR	-0.7259			
ETS	B			
Logistic Regression	.1915	.001		

Comma Items Compared



Results

Both of these items assessed a students ability to identify an error in punctuation within a sentence.

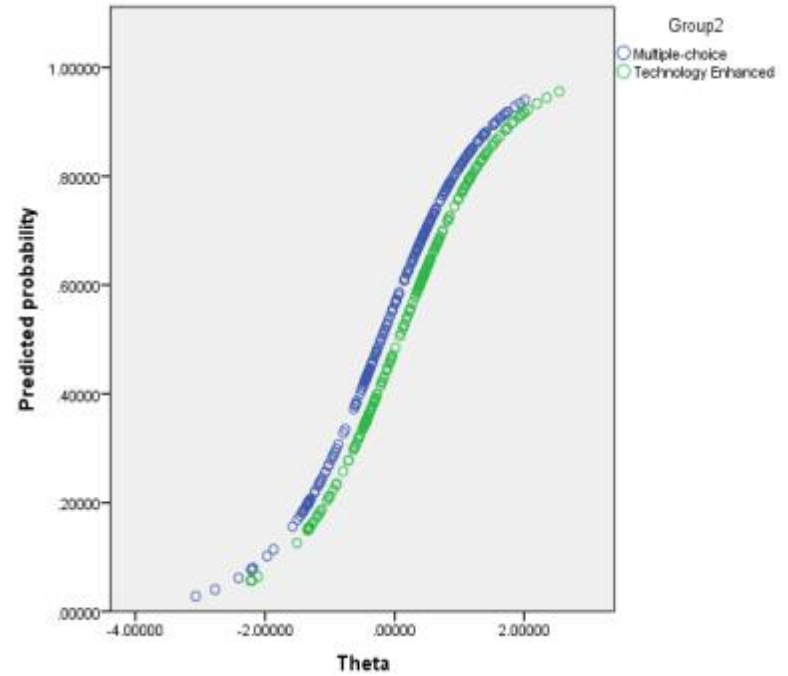
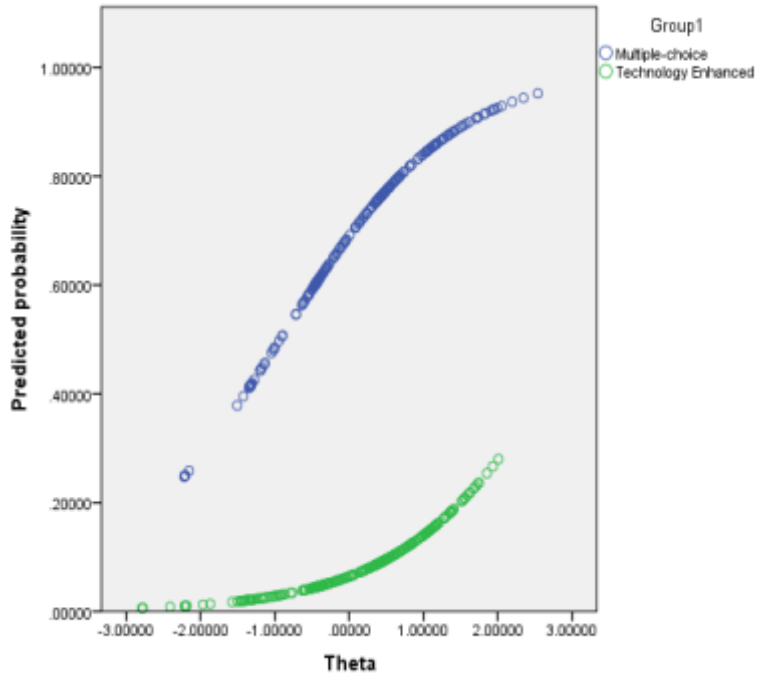
Item 1e: Remove Punctuation

Type	P-value	Item-total Corr	a	b
MC	67.8	.321	.574	-.851
TE	8.2	.073	.351	4.226
DIF	Statistic	p-Value		
MH LOR	3.2411			
ETS	C			
Logistic Regression	.031	< 0.001***		

Item 2e: Remove Punctuation

Type	P-value	Item-total Corr	a	b
MC	55.3	.393	.637	-.233
TE	49.3	.414	.677	.056
DIF	Statistic	p-Value		
MH LOR	0.3616			
ETS	A			
Logistic Regression	.717	0.074		

Remove Punctuation Items Compared



Conclusion

- ▶ DIF is occurring!
 - ▶ Why?
- ▶ Other factors that may be interfering
 - ▶ Item length
 - ▶ Balancing of MC
 - ▶ Difficulty of tech items
 - ▶ Guessing
 - ▶ Pilot test data
- ▶ Should we be using these item types?
 - ▶ It may depend on the construct of interest
 - ▶ Degree of fidelity needed

Final Thoughts

- ▶ Technology enhanced items
 - ▶ Increasing use in the educational field.
 - ▶ More research into their characteristics is needed!
 - ▶ Smaller sample size may have affected outcome.
 - ▶ Larger sample sizes may be beneficial for future studies (less missing data).

QUESTIONS?

