



**ATP**  
**Innovations  
in Testing**  
Orlando, Florida

**2010**

# Building a Better Mousetrap: Using Discrete Options to Improve the Multiple-Choice Question

Neal Kingston, University of Kansas

David Foster, Kryterion Inc.

Hal L. Miller, Jr., Brigham Young University

Gail Tiemann, University of Kansas

# Overview

- Introducing Discrete Option Multiple-Choice
- Research design
- Test Design
- Testlet Statistics
- Generalizability
- Confirmatory factor analysis
- Classical item statistics
- Gender DIF
- Survey responses
- Future research directions



# Discrete Option Multiple Choice: Example

Traditional Multiple Choice	Foster Item
<p data-bbox="92 611 571 651">Which is a prime number?</p> <ul data-bbox="166 722 282 911" style="list-style-type: none"><li data-bbox="166 722 282 753">a. 39</li><li data-bbox="166 773 282 805">b. 75</li><li data-bbox="166 825 282 856">c. 57</li><li data-bbox="166 876 282 908">d. 29</li></ul>	<p data-bbox="989 611 1418 651">Is this a prime number?</p> <p data-bbox="989 722 1031 753">29</p> <div data-bbox="989 819 1495 888"><input data-bbox="989 819 1236 888" type="button" value="No"/> <input data-bbox="1240 819 1495 888" type="button" value="Yes"/></div>

Demonstration:

[http://fosteritem.com/index.php?option=com\\_fi&view=tests&Itemid=53](http://fosteritem.com/index.php?option=com_fi&view=tests&Itemid=53)



# Testlet Specification

Benchmark/Standard	Item Description
Uses basic and advanced procedures while performing the processes of computation	Multiplication of two three digit numbers where the first digit is greater than or equal to 5 and the second digit is 0
	Solve a numeric equation requiring knowledge of order of operations. Problem will include one of each of the following operations: parentheses, exponentiation, multiplication, division, and subtraction and will be written on one horizontal line.
	Reduce an equation to simplest form. Equation will include two variables and two constants represented algebraically.
	Simple combination word problem.
Understands and applies basic and applied properties of the concepts of measurement	Relationship between area and volume.
	Solve a simple word problem using velocity.
Understands and applies basic and applied properties of probability	Determines probability using a counting procedure, enumerating equally likely events and dividing events of interest by total events.
	Determines conditional probabilities from a three by three contingency table.
Understands and applies basic and applied properties of functions and algebra	Identify the correct graph of a pair of inequalities that have an area of overlap
	Solve a system of linear equations with two variables

# Research Design

Forms were to be assigned randomly to students at two universities, but it went slightly awry – thank goodness for the inclusion of strong covariates

Group Taking Form	n	Raw Mean		.95 Confidence Interval	
		DOMC Set 1	MC Set 2	DOMC Set 1	MC Set 2
A1	249	6.2	7.0	5.95-6.45	6.79-7.21
A2	281	6.5	7.2	6.26-6.74	7.00-7.40
B1	141	6.1	6.9	5.77-6.43	6.62-7.18
B2	131	6.0	6.8	5.66-6.34	6.51-7.09

DOMC Set 1      SD = 2.01

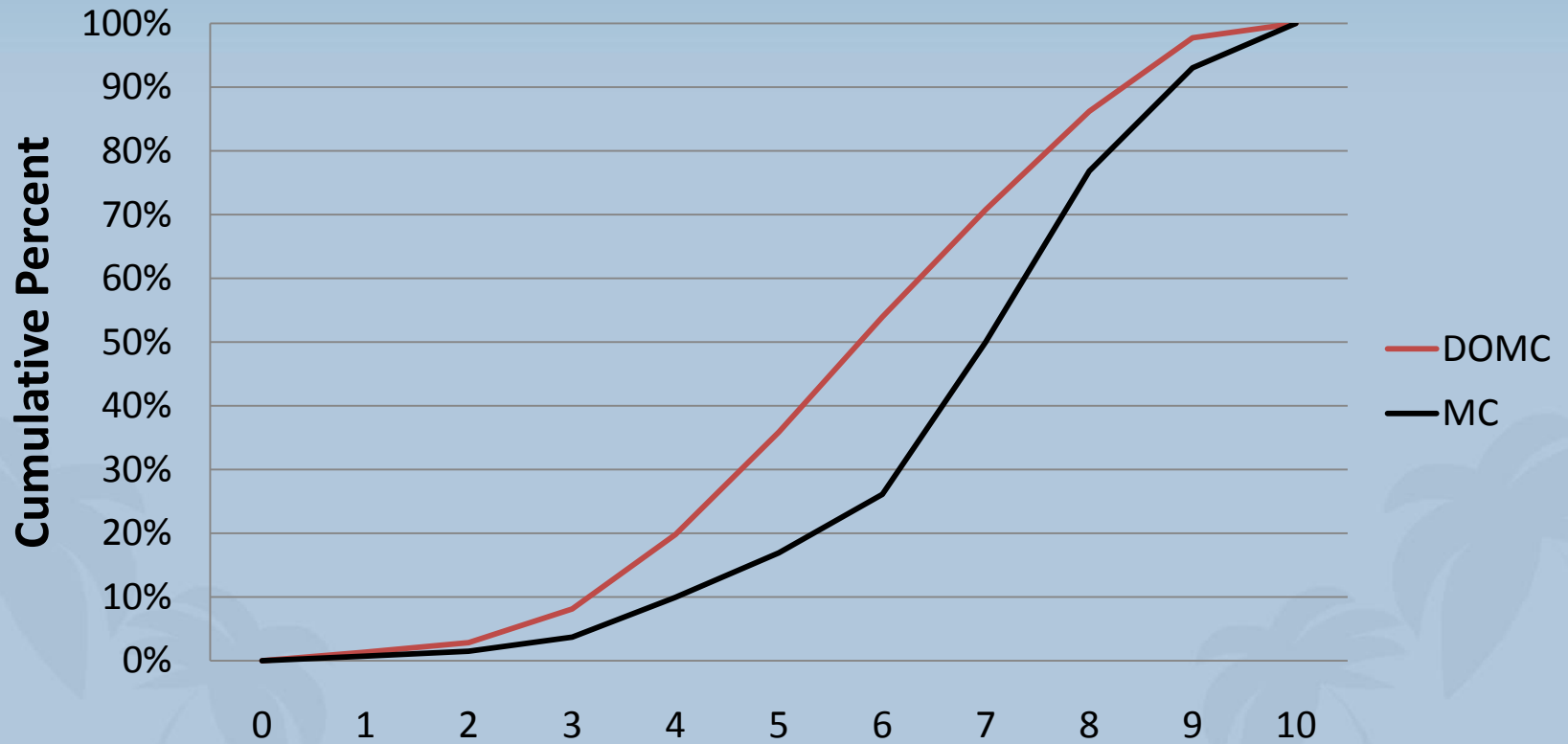
MC Set 2        SD = 1.67

# Testlet Statistics

Set 3 - Raw Frequencies (Groups are not random)				
Score	Count		Cum. Percent	
	DOMC	MC	DOMC	MC
0	0	0	0.0%	0.0%
1	7	2	1.3%	0.7%
2	8	2	2.8%	1.5%
3	28	6	8.1%	3.7%
4	62	17	19.8%	9.9%
5	85	19	35.8%	16.9%
6	96	25	54.0%	26.1%
7	89	65	70.8%	50.0%
8	82	73	86.2%	76.8%
9	61	44	97.7%	93.0%
10	12	19		
n	530	272		
Mean	6.23	7.21		
SD	1.95	1.80		

# Testlet Statistics

## Item Set 3

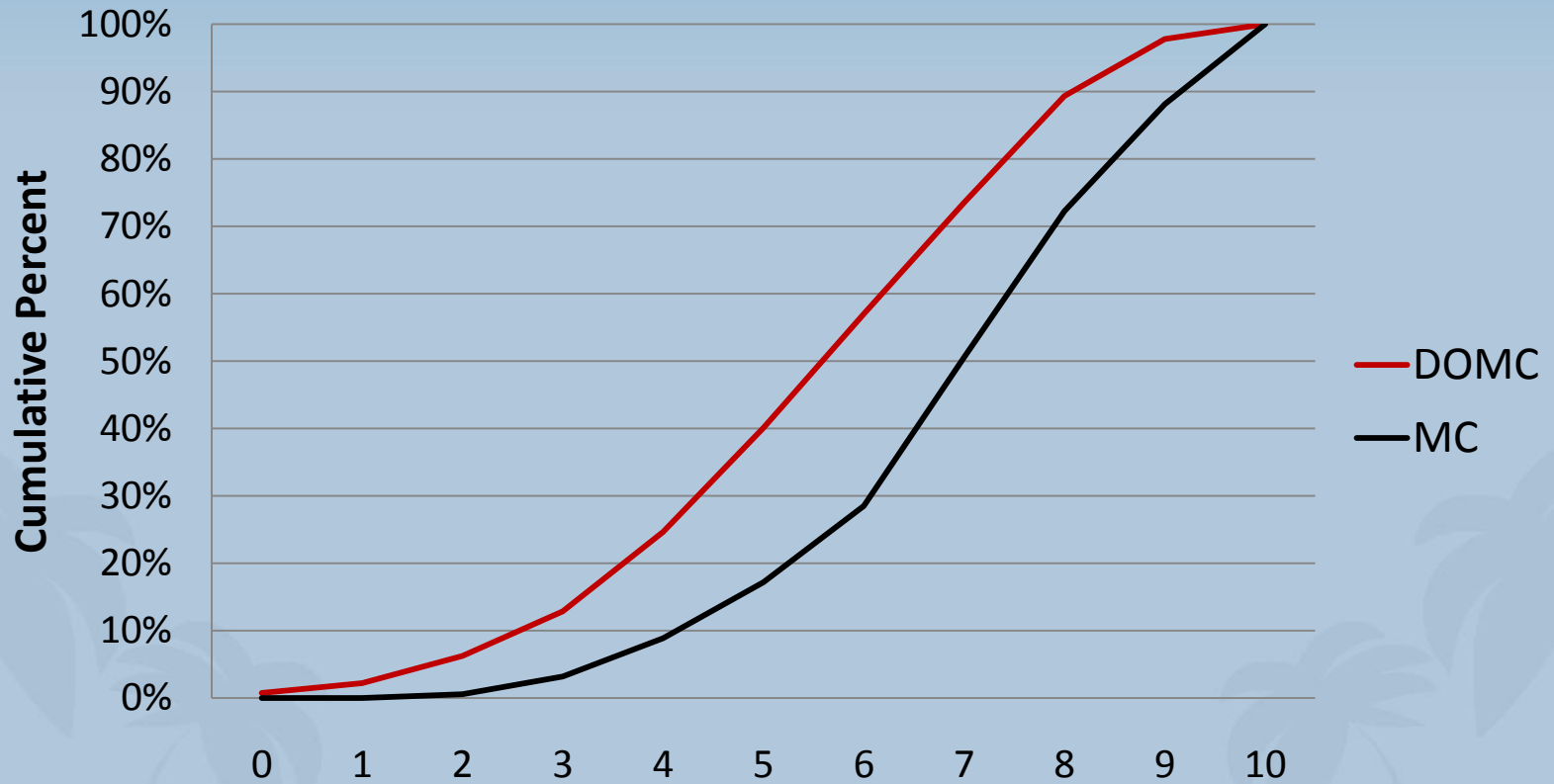


# Testlet Statistics

Set 4 - Raw Frequencies (Groups are not random)				
Score	Count		Cum. Percent	
	DOMC	MC	DOMC	MC
0	2	0	0.7%	0.0%
1	4	0	2.2%	0.0%
2	11	3	6.3%	0.6%
3	18	14	12.9%	3.2%
4	32	30	24.6%	8.9%
5	42	44	40.1%	17.2%
6	46	60	57.0%	28.5%
7	45	117	73.5%	50.6%
8	43	115	89.3%	72.3%
9	23	84	97.8%	88.1%
10	6	63		
n	272	530		
Mean	5.96	7.21		
SD	2.10	1.83		

# Testlet Statistics

## Item Set 4



# Testlet Statistics: MANCOVA

	Testlet 3		Testlet 4	
	DOMC	MC	DOMC	MC
Raw Means	6.23	7.21	5.96	7.31
Adjusted Means	6.17	7.34	6.10	7.24

	Testlet 3		Testlet 4	
	DOMC	MC	DOMC	MC
Raw S.E.E.	.08	.11	.13	.08
Adjusted S.E.E.	.07	.09	.09	.06

All mean differences are significant well beyond the .0001 level

# Generalizability: Internal Consistency

Cronbach Alpha				
Item Type	Item Set			
	1	2	3	4
DOMC	.56		.54	.59
MC		.47	.51	.57

# Generalizability: Intercorrelations

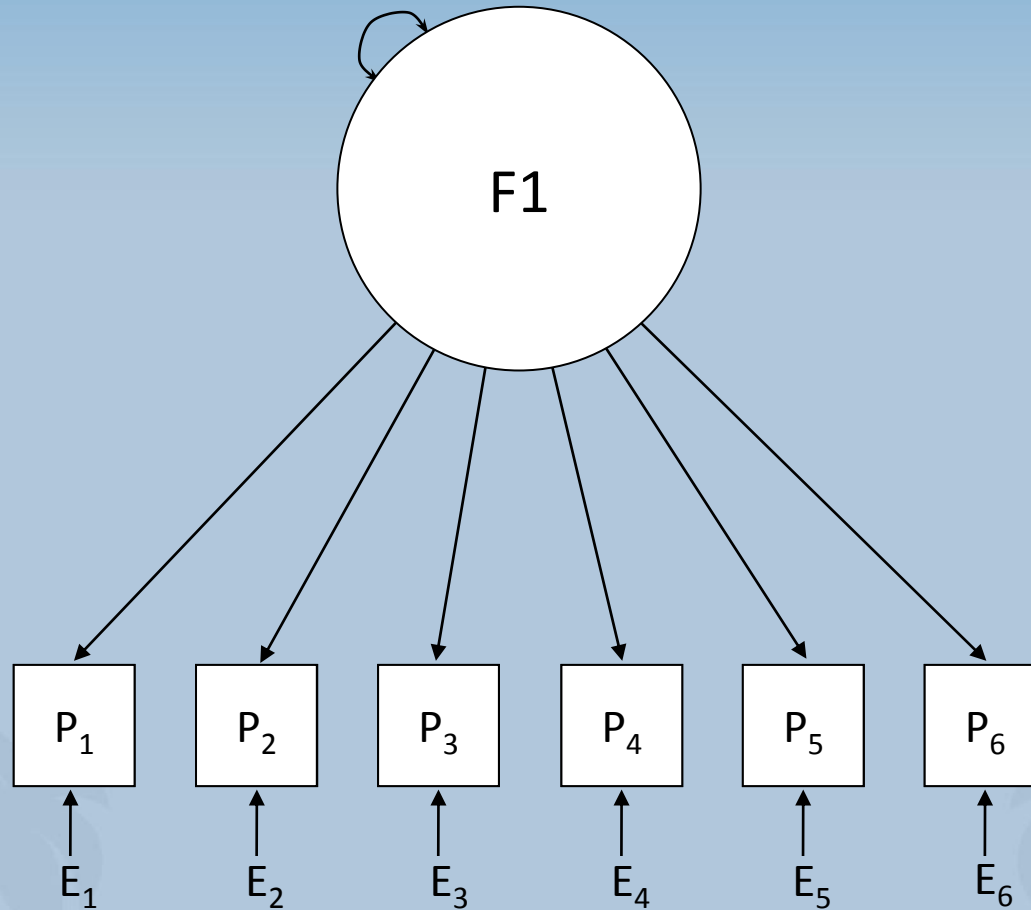
(Correlations in lower triangle, reliabilities on diagonal, corrected correlations in upper triangle)

	DOMC 1	DOMC 3	DOMC 4	MC 2	MC 3	MC 4
DOMC 1	.56	>1.00	.96	>1.00	.92	>1.00
DOMC 3	.61	.54	--	.99	--	>1.00
DOMC 4	.55	--	.59	>1.00	>1.00	--
MC 2	.55	.50	.56	.47	>1.00	>1.00
MC 3	.49	--	.59	.50	.51	--
MC 4	.61	.61	--	.54	--	.57

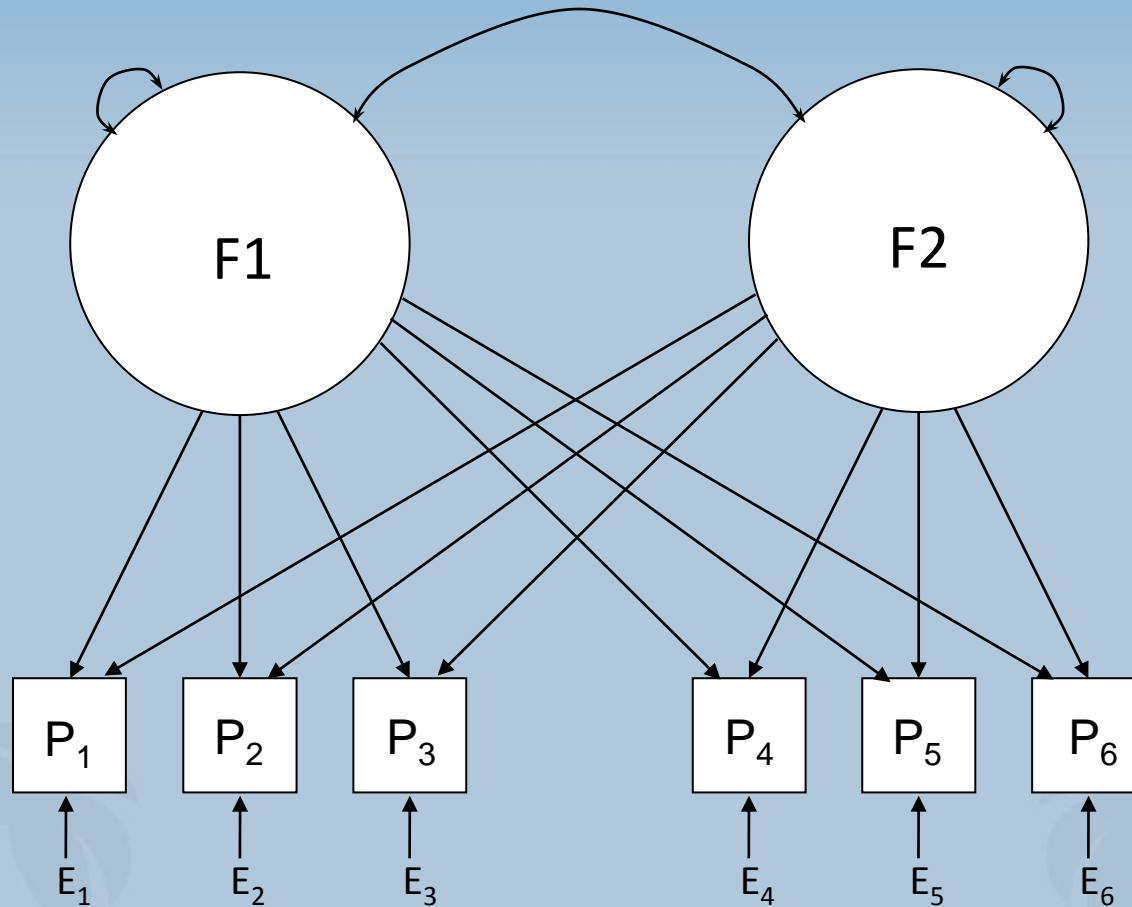
# Confirmatory Factor Analysis

- Divided Testlet 1 (DOMC) and Testlet 2 (MC) each into three item parcels (3, 3, and 4 items)
- Used SAS Proc Calis to fit two one-factor and one two-factor model

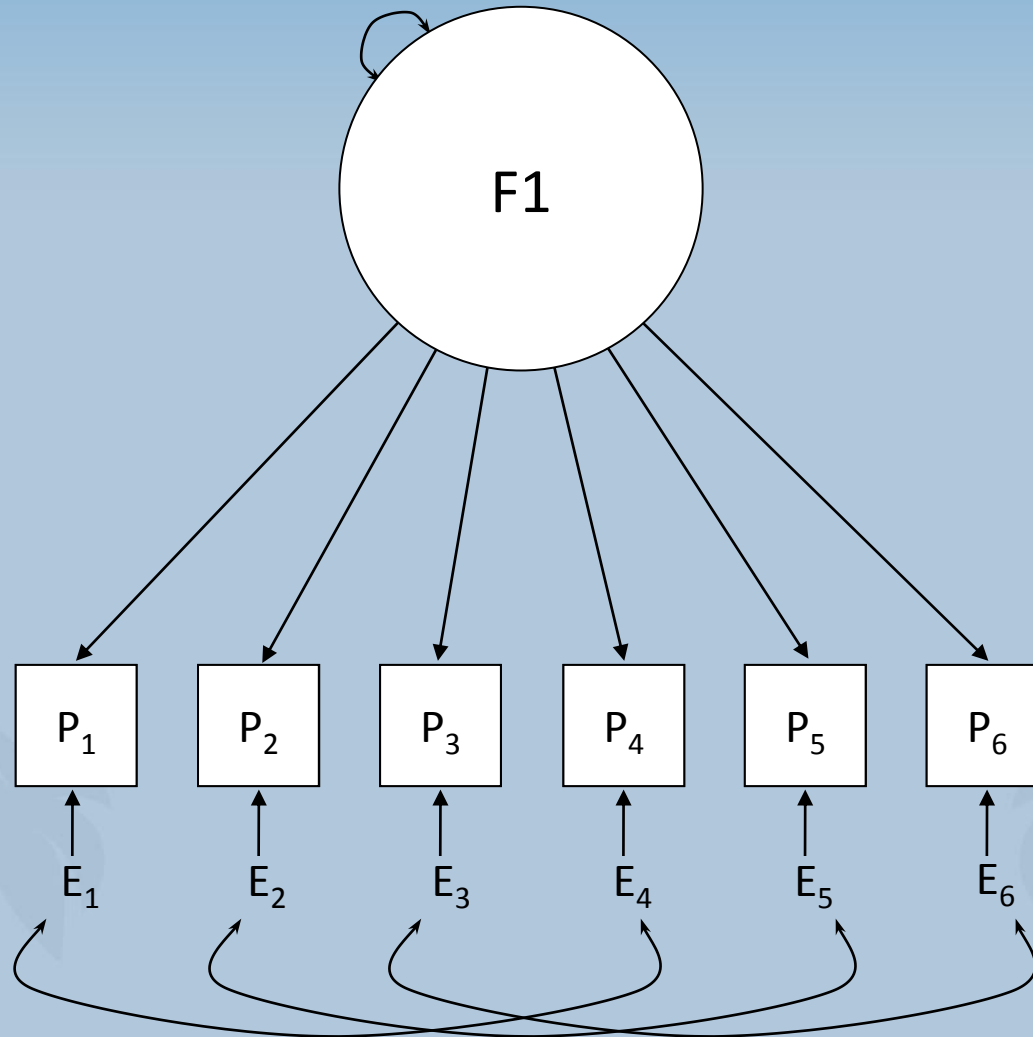
# One-Factor Model



# Two-Factor Model



# One-Factor Model with Pairwise Fixed Error Variances

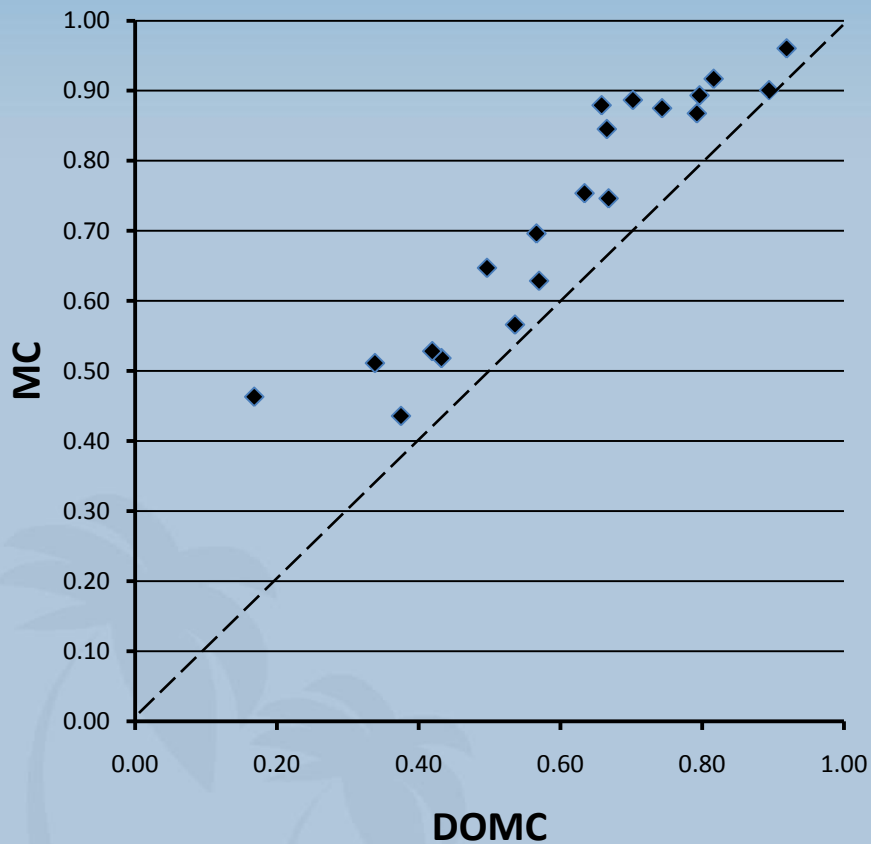


# Confirmatory Factor Analysis

Statistic	One Factor	One Factor with pairwise fixed error variances	2 Factor Model
Chi-square	51.2	12.9	5.81
df	9	6	2
p (influenced by sample size)	>.0001	.04	.05
Chi-square/df	5.7	2.2	2.9
RMSEA	.08	.04	.05
Bentler comparative fit index (should exceed .90)	.93	.99	.95
Bentler & Bonnet Non-normed index	.88	.97	.95

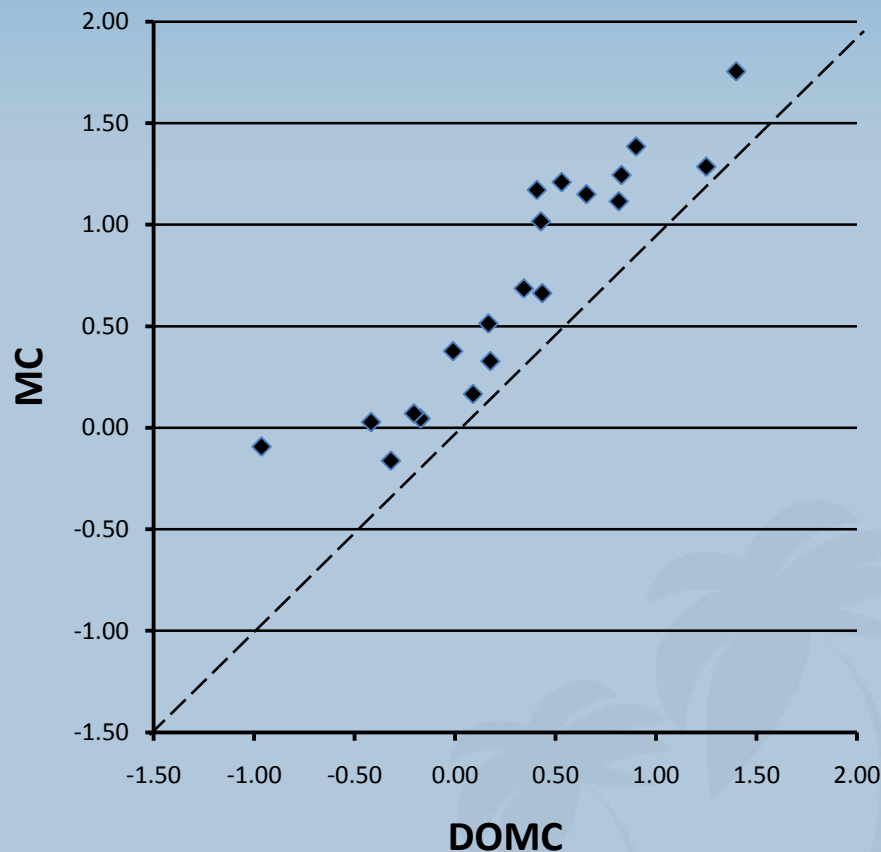
# Classical Item Statistics - Difficulty

## Item P-Values



Mean DOMC = .61  
Mean MC = .73

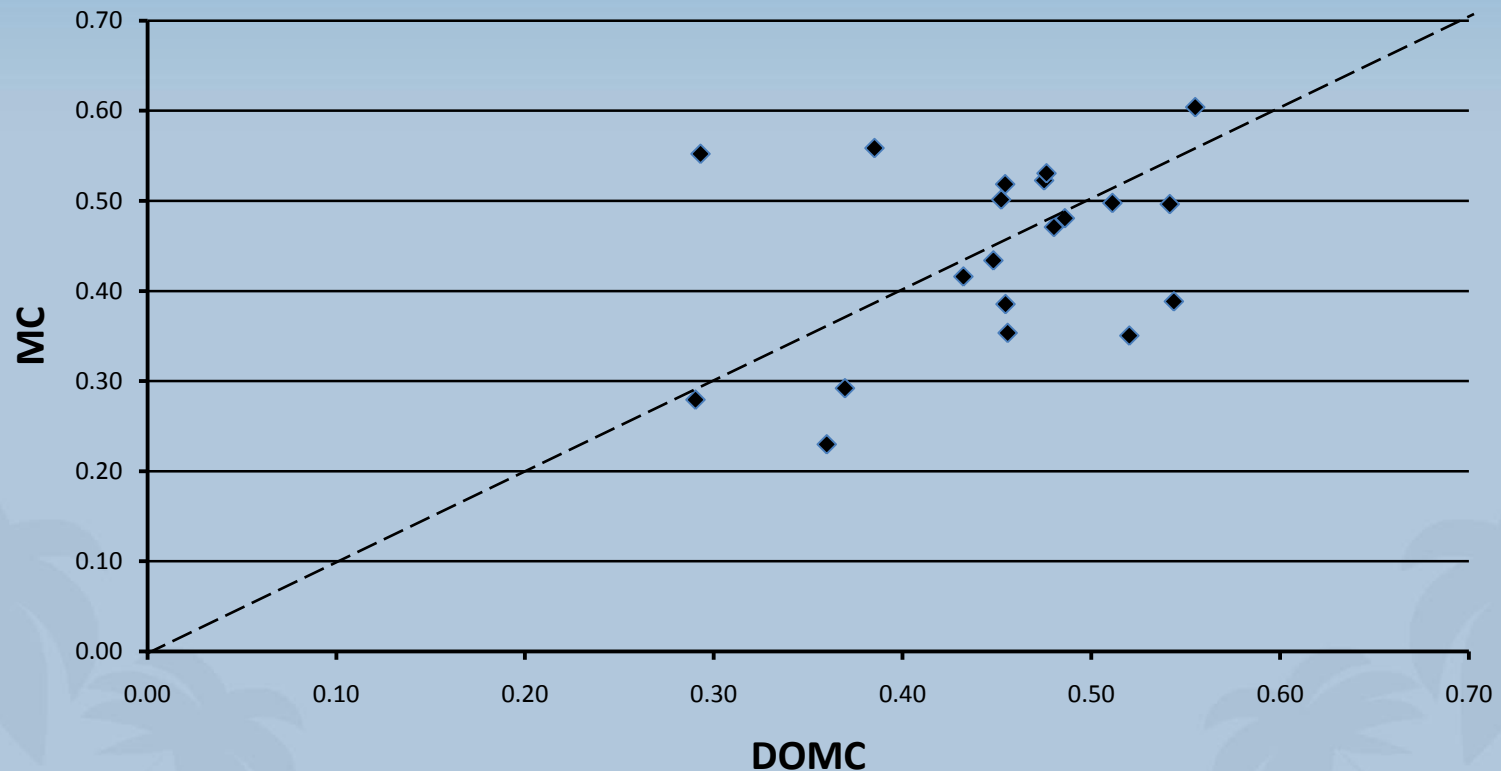
## Normalized Item Difficulties



Mean DOMC = .32  
Mean MC = .70

# Classical Item Statistics - Biserial

## Biserial Correlation



Mean DOMC = .45

Mean MC = .44

# Gender DIF

- Common Items 1-20
  - DOMC – 10 items
  - Multiple Choice – 10 items
- Reference Group: Males – 303
- Focal Group: Females - 407
- Analyzed using DIFAS 4.0 software

# Differential Item Functioning

- Combined Decision Rule

- DOMC                    3 Flags                    1(M), 3(M), 5(F)
- MC                        3 Flags                    1(M), 7(F), 9(F)

- ETS Categorization

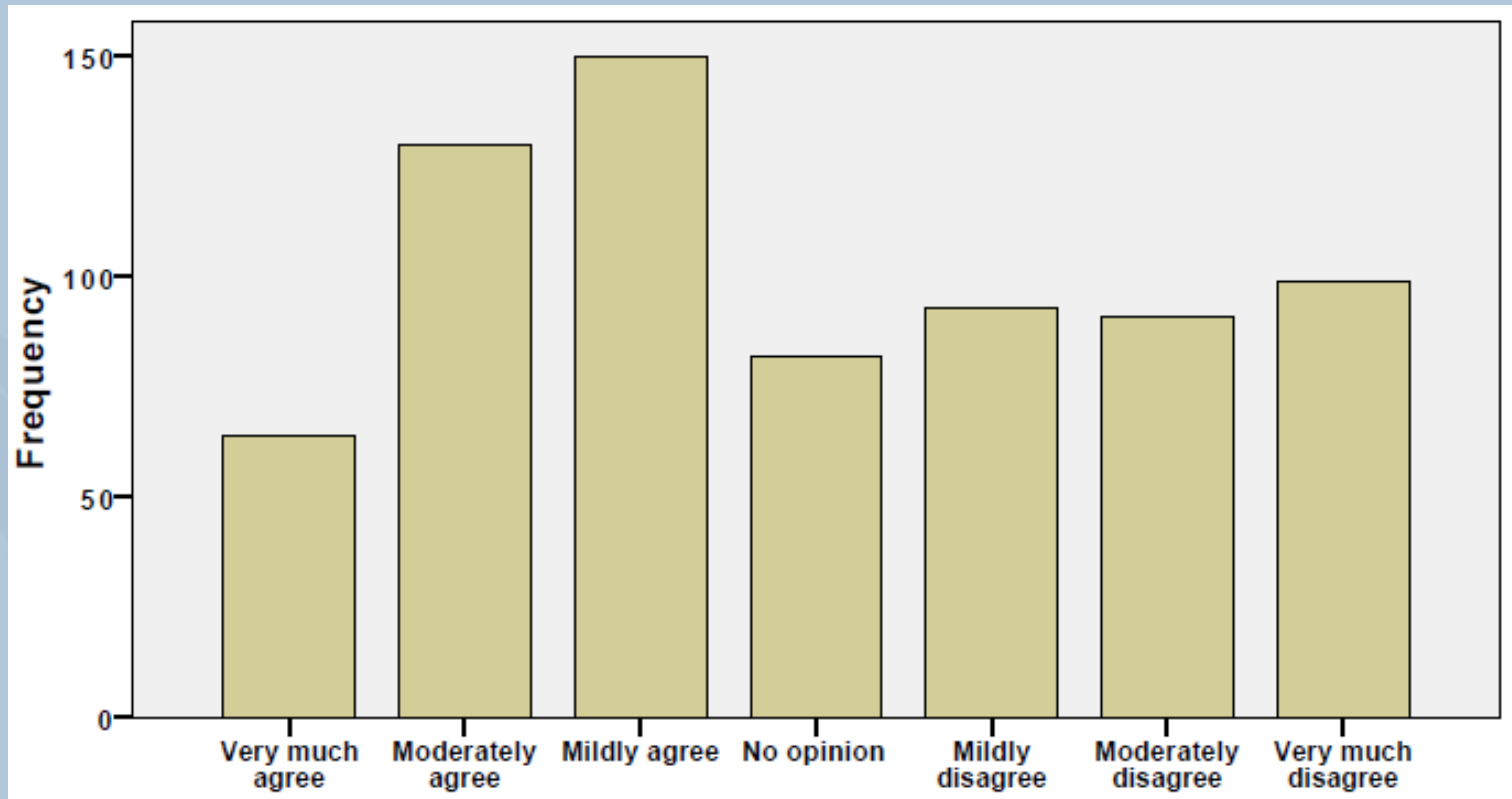
- DOMC                    1 Moderate                1(M)
- MC                        4 Moderate                1(M), 3(M), 7(F), 9(F)

M = Males do more poorly than expected  
F = Females do more poorly than expected

# Survey Results

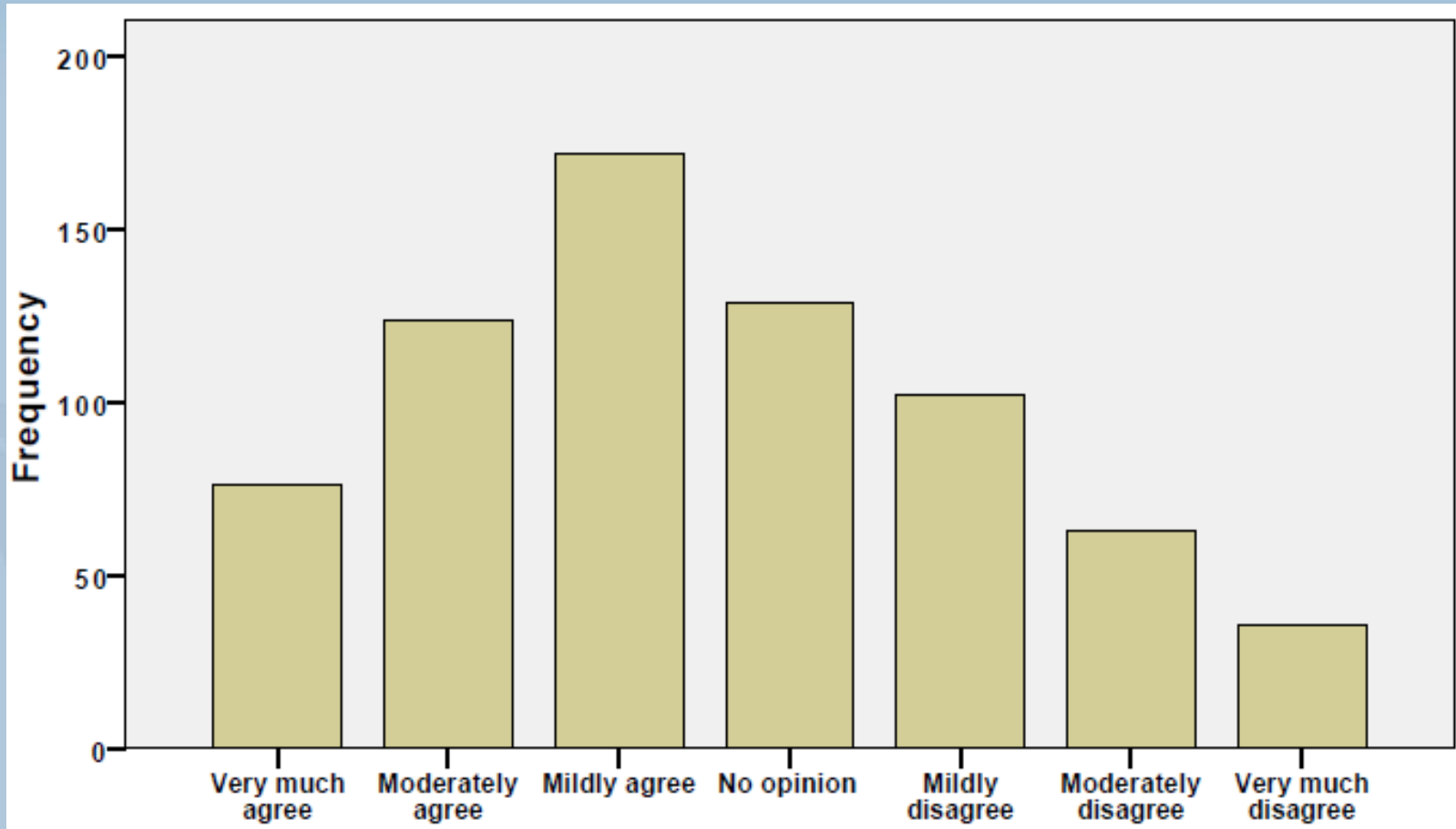
# Fairness

The DOMC question type is considered more fair to ALL students because it removes the advantage some students have who are better at taking tests that have multiple choice questions. Since test taking skills are skills considered to be unrelated to the course content of psychology knowledge, it is best to remove the effects of such skills from the assessment and judge everyone equally as to their knowledge of the subject matter.



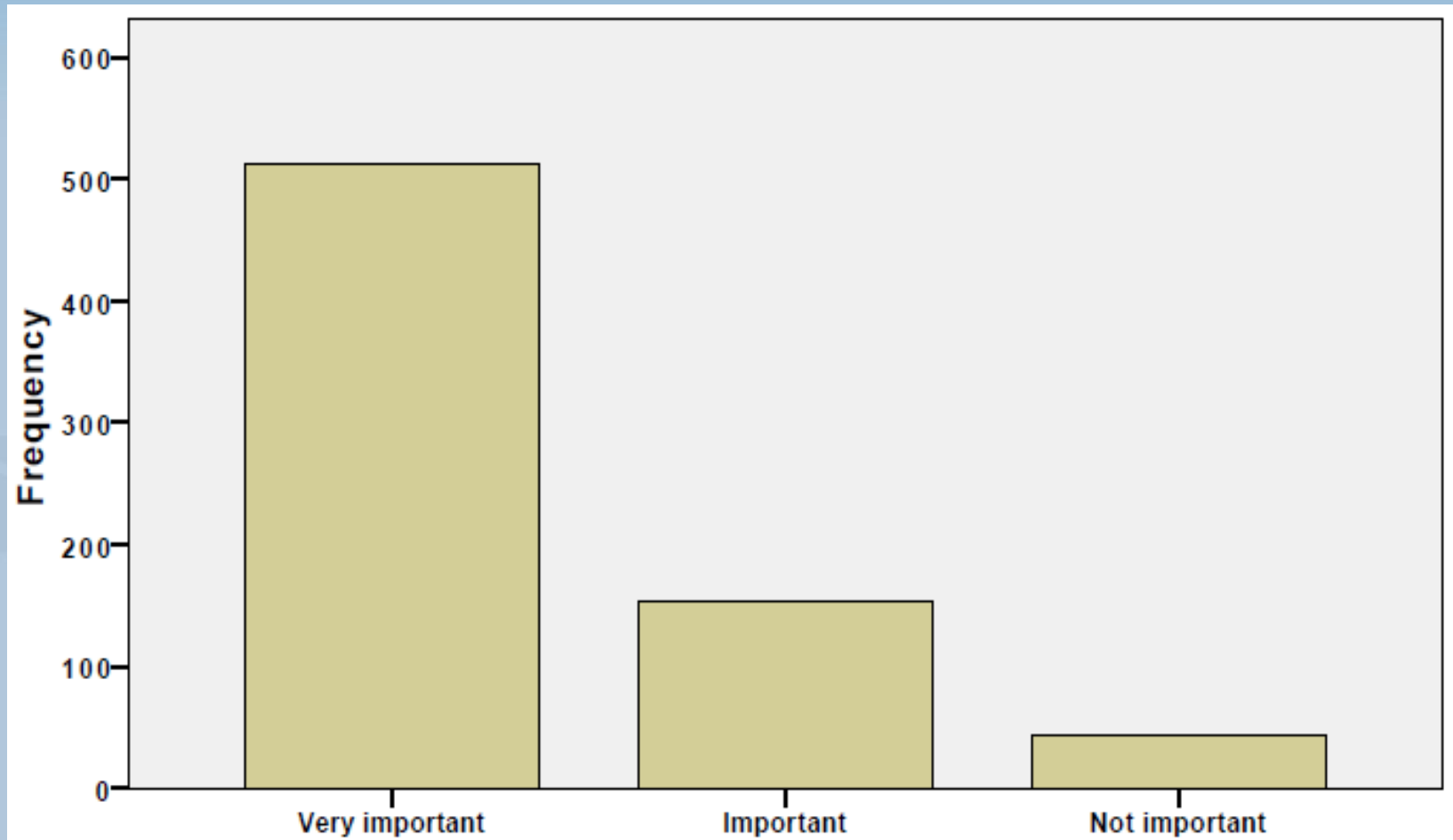
# Fairness

The DOMC question type removes the advantage some students have who are better at taking tests with multiple choice questions. If a person has those test taking skills he or she should be able to keep the advantage they have over students who don't have those skills.



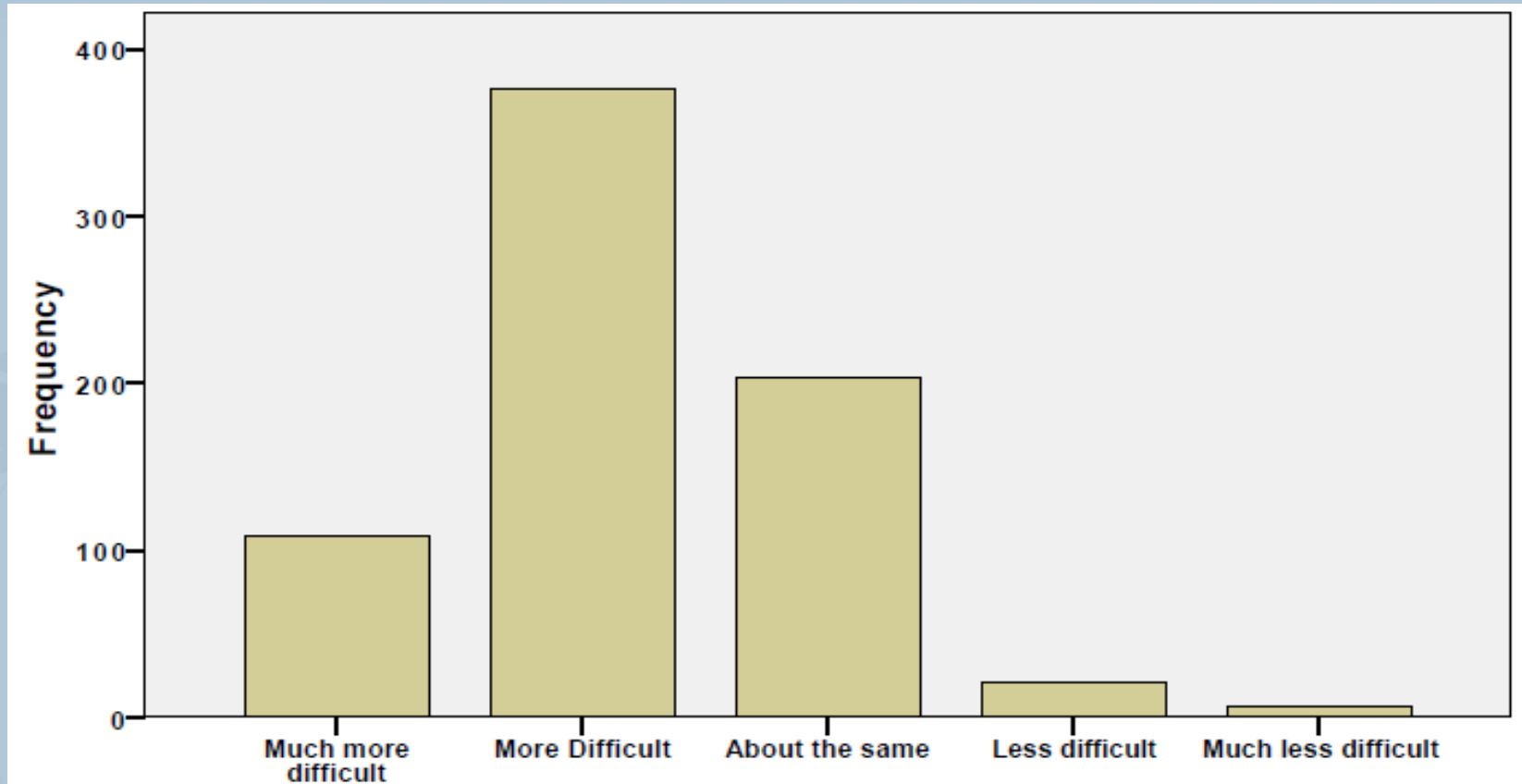
# Review Questions

When test questions are presented using the DOMC question type, it is not possible to allow students to go back, review previous questions, and change answers. How important is it for you to be able to review previously seen or answered questions?



# Reduces Cheating

One purpose of the DOMC format is to make it more difficult to cheat. Cheating is defined as using some pre-knowledge of the question by getting such information from students who have already taken the test. How much more difficult is it to effectively use such information on the DOMC format compared to the traditional MC question?



# Reasons for Reduced Cheating

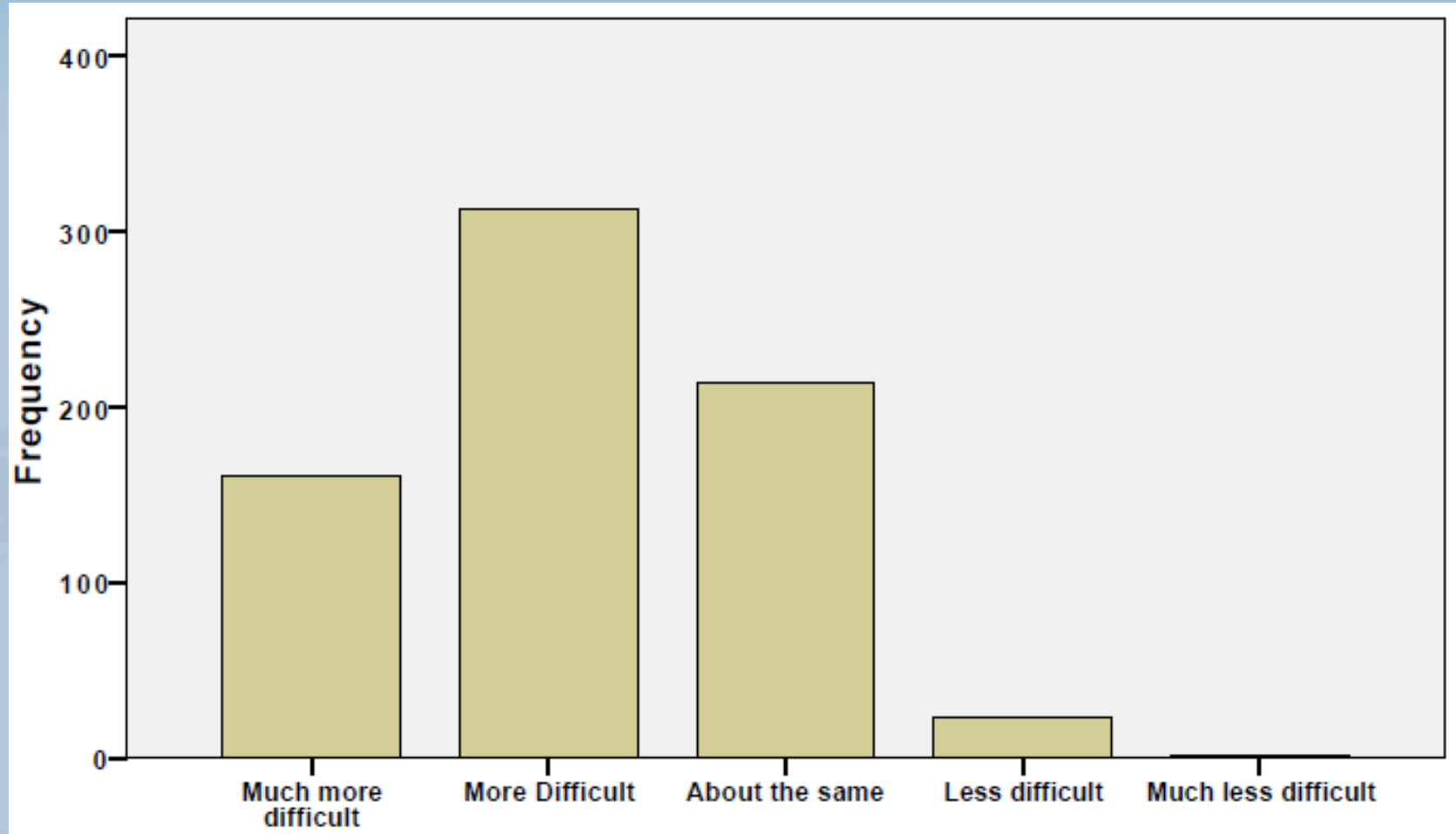
Students were asked to provide reasons why the DOMC question type makes it more difficult to cheat.

## Freq. Responses

- 99 Do not see all answer choices at the same time
- 83 Difficult to keep track of the answers
- 71 Hard to remember answer choices
- 46 Get info from answers, comparing the answer to others
- 29 Random order of questions and answers
- 22 No chance to look at answers again

# Reduces Test Theft

How much more difficult is it to memorize or copy test question content so it can be shared with others for the Foster question type compared to the traditional multiple choice question?



# Reasons for Reduced Theft

Students were asked to provide reasons why the DOMC question type makes it more difficult to steal test content.

<b>Freq.</b>	<b>Responses</b>
91	Do not see all answer choices at the same time
86	Hard to remember answer choices
54	Difficult to keep track of the answers
34	Random order of questions and answers
23	Hard to remember letters for both questions/answers
19	Hard to remember the question
18	Need to remember more info at once

# Acceptability

Given your brief experience with the DOMC question type, please provide any other opinions you have about whether or how this test should be used in college exams.

## **Freq. Responses**

98 Should not be used in college exams. Prefer the traditional multiple choice test type.

51 Should be used in college exams

41 Takes much more time than the traditional multiple-choice question

41 Forces test takers to solve the problem on their own instead of choosing the answer

# Acceptability

Given your brief experience with the DOMC question type, please provide any other opinions you have about whether or how this test should be used in college exams.

## **Freq. Responses**

- 34 Makes the test harder than with traditional multiple-choice questions
- 32 Serves the purpose of eliminating cheating and memorizing answers
- 31 Takes the same amount of time
- 30 Cannot go back and see previous answers
- 30 Hard to guess on choices

# Acceptability

Given your brief experience with the DOMC question type, please provide any other opinions you have about whether or how this test should be used in college exams.

## **Freq. Responses**

- 27 Cannot see all the answer choices
- 21 Cannot get info from answers; cannot compare the answers to others
- 17 Cannot eliminate answer choices
- 17 Easier than traditional multiple choice questions, takes less time
- 13 Effective in math tests
- 12 Not sure

# Future Research Directions

- IRT calibrations
- Analysis of response times
- Empirical Cheating Study