

A Bayesian Analysis of Item Format and Test Security

Kylie N. Gorney

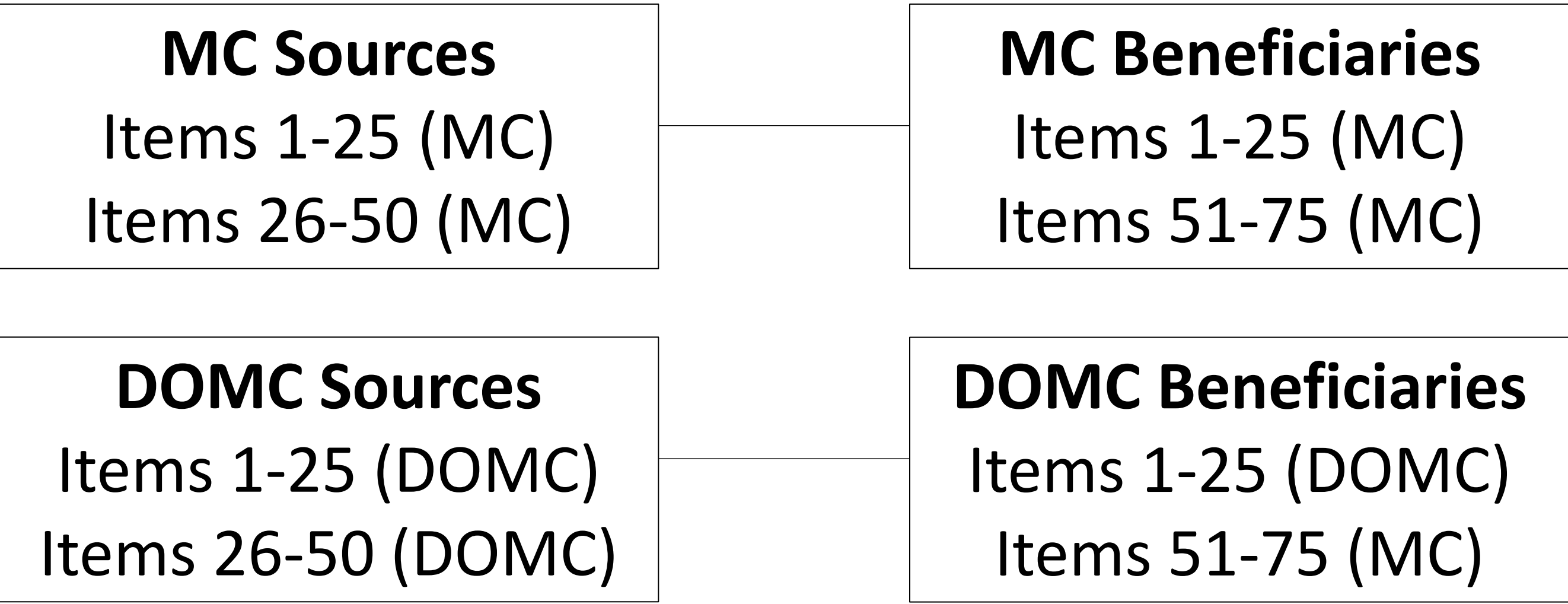
University of Wisconsin-Madison

Introduction

- While a **multiple-choice (MC)** item presents all answer options simultaneously, a **discrete-option multiple-choice (DOMC)** item delivers its options sequentially. After each option has been displayed, the examinee must decide whether they believe it to be correct or incorrect. Options continue to be presented until all have been exhausted, or the item has been scored.
- Those who hope to capture item information to assist future examinees may find this task to be more difficult if the DOMC item format is used since it is not guaranteed that all options will be displayed.

Method

- Participants:** 110 UW-Madison undergraduate students enrolled in a human development course in Spring 2019
- Materials:** 75 items, each with 5 options, measuring knowledge of human development
- Procedure:** Participants were randomly assigned to one of four groups. Sources took the test first, and once finished, each Source was paired with one Beneficiary to discuss the test and the items they had seen. Then, the Sources were dismissed, and the test was given to the Beneficiaries. Each Beneficiary was administered 25 compromised items and 25 secure items.



- Research Question:** How does item format affect the extent to which Beneficiaries benefit from item compromise?
- Hypothesis:** MC Beneficiaries will benefit more by scoring higher on compromised items than DOMC Beneficiaries.

Multiple Linear Regression

- Predictors:** item format (0 = MC, 1 = DOMC), Beneficiary's secure item score, Source's compromised item response time (RT) in mins, Beneficiary's major (0 = Human Development or Psychology, 1 = Other), Beneficiary's gender (0 = Male, 1 = Female)
- Outcome:** Beneficiary's compromised item score
- Results:** Item format and secure item score were the only predictors significant at the .05 level.

Bayesian Linear Regression

- Procedure:** Two chains were used with 1,000 iterations for adaptation, 5,000 burn-in samples, and 50,000 follow-up samples with a thinning interval of 10.
- Priors**
 - Intercept: $N(5, 5)$
 - Item format: $N(-3, 5)$
 - Secure score: $N(0.5, 1)$
 - Source RT: $N(0.5, 1)$
 - Major: $N(-3, 5)$
 - Gender: $N(0, 5)$
- Results:** Item format and secure item score were the only predictors effective with 95% confidence.

	Multiple Reg		Bayesian Reg		
	Estimate	SE	Mean	95% HPD Int	
Intercept	4.31	3.32	5.14	1.44	8.71
Item format	-2.65*	0.98	-2.69	-4.55	-0.99
Secure score	0.66*	0.18	0.64	0.39	0.88
Source RT	0.16	0.11	0.17	-0.03	0.36
Major	0.70	1.23	-0.12	-2.08	1.83
Gender	-0.89	1.13	-0.76	-2.67	1.12

Note: * $p < .05$

Bayesian Model Averaging

- Model Selection:** Though 10 models were returned as possible candidates, the posterior probabilities of Models 1 and 2 alone exceed 0.5.
- Results:** The coefficients for item format and secure item score returned the highest probabilities of being different from 0.

	p != 0	Mean	SD	Model 1	Model 2
Intercept	1.00	5.74	2.69	6.62	4.31
Item format	0.76	-1.77	1.31	-2.19	-2.48
Secure score	1.00	0.62	0.17	0.61	0.67
Source RT	0.31	0.05	0.10		0.17
Major	0.12	0.09	0.49		
Gender	0.17	-0.18	0.61		
R ²				0.30	0.34
BIC				-8.82	-7.75
Posterior probability				0.33	0.19

Discussion

- It is very likely that item format affects performance on compromised items.
- MC Beneficiaries benefitted more from the coaching sessions than did the DOMC Beneficiaries, suggesting that the DOMC item format may be an effective tool used to increased test security.

Reference

- Foster, D., & Miller, H. L. (2009). A new format for multiple-choice testing: Discrete-Option Multiple-Choice. Results from early studies. *Psychology Science Quarterly*, 51, 355–369.