



Item Format and  
Test Security

Gorney &  
Wollack (2021)

Background

Purpose

Method

Results &  
Discussion

Classical Statistics

Item Response  
Theory (IRT)

Conclusion

References

# Does Item Format Affect Test Security?

Kylie N. Gorney and James A. Wollack

University of Wisconsin-Madison

National Council on Measurement in Education (NCME)  
Annual Meeting · 2021



# Background

Item Format and  
Test Security

Gorney &  
Wollack (2021)

Background

Purpose

Method

Results &  
Discussion

Classical Statistics

Item Response  
Theory (IRT)

Conclusion

References

- The **discrete-option multiple-choice (DOMC) item format** was proposed as an alternative to the traditional multiple-choice (MC) items (Foster & Miller, 2009)
  - Sequential presentation of answer options



# DOMC Example

## Item Format and Test Security

Gorney &  
Wollack (2021)

## Background

Purpose

Method

Results &  
Discussion

Classical Statistics

Item Response  
Theory (IRT)

Conclusion

References

Q: This is an example of a fine motor skill.

Balancing

- Yes
- No



# DOMC Example

## Item Format and Test Security

Gorney &  
Wollack (2021)

## Background

Purpose

Method

Results &  
Discussion

Classical Statistics

Item Response  
Theory (IRT)

Conclusion

References

Q: This is an example of a fine motor skill.

Balancing

- Yes
- **No**



# DOMC Example

## Item Format and Test Security

Gorney &  
Wollack (2021)

## Background

Purpose

Method

Results &  
Discussion

Classical Statistics

Item Response  
Theory (IRT)

Conclusion

References

Q: This is an example of a fine motor skill.

Crawling

- Yes
- No



# DOMC Example

## Item Format and Test Security

Gorney &  
Wollack (2021)

## Background

Purpose

Method

Results &  
Discussion

Classical Statistics

Item Response  
Theory (IRT)

Conclusion

References

Q: This is an example of a fine motor skill.

Crawling

- Yes
- No



# DOMC Example

## Item Format and Test Security

Gorney &  
Wollack (2021)

## Background

Purpose

Method

Results &  
Discussion

Classical Statistics

Item Response  
Theory (IRT)

Conclusion

References

Q: This is an example of a fine motor skill.

Writing

- Yes
- No



# DOMC Example

## Item Format and Test Security

Gorney &  
Wollack (2021)

## Background

Purpose

Method

Results &  
Discussion

Classical Statistics

Item Response  
Theory (IRT)

Conclusion

References

Q: This is an example of a fine motor skill.

Writing

- **Yes**
- **No**





# Background

## Item Format and Test Security

Gorney &  
Wollack (2021)

### Background

#### Purpose

#### Method

#### Results & Discussion

#### Classical Statistics

#### Item Response Theory (IRT)

#### Conclusion

#### References

- The **discrete-option multiple-choice (DOMC) item format** was proposed as an alternative to the traditional multiple-choice (MC) items (Foster & Miller, 2009)
  - Sequential presentation of answer options
  - Reduces the average number of answer options shown
- Previous research on item preknowledge did not separate the difficulty and security aspects of the item formats (Tiemann et al., 2014)



# Purpose

## Item Format and Test Security

Gorney &  
Wollack (2021)

## Background

## Purpose

## Method

## Results & Discussion

## Classical Statistics

## Item Response Theory (IRT)

## Conclusion

## References

The purpose of this study was to see whether the DOMC format is more or less affected by item preknowledge than the multiple-choice format.



# Participants

## Item Format and Test Security

Gorney &  
Wollack (2021)

## Background

## Purpose

## Method

## Results & Discussion

## Classical Statistics

## Item Response Theory (IRT)

## Conclusion

## References

- **Participants:** 150 UW-Madison students enrolled in a human development course in the Spring 2020 or Fall 2020 semesters
- **When:** approximately four weeks prior to the end of each semester
- **Where:** entirely virtual using the Secure Exam Interface (SEI) by Caveon



# Procedure

## Item Format and Test Security

Gorney &  
Wollack (2021)

## Background

## Purpose

## Method

## Results & Discussion

## Classical Statistics

## Item Response Theory (IRT)

## Conclusion

## References

- 1 Review the study guide for 50–60 minutes to prepare for the upcoming test
- 2 Take the test (70 minutes max)
- 3 Potentially win one of several \$40 cash prizes (must score in the top 50% of all test-takers)



# Human Development Test

## Item Format and Test Security

Gorney &  
Wollack (2021)

## Background

## Purpose

## Method

## Results & Discussion

## Classical Statistics

## Item Response Theory (IRT)

## Conclusion

## References

- 68-item test
- 6 item sets that differed in compromise status and item format
- 8 test forms

Test Form	Half 1		Half 2	
	Item Format	Item Sets	Item Format	Item Sets
A1	MC	1, 3, 5	DOMC	2, 4, 6
A2	MC	1, 3, 5	DOMC	2, 4, 6
B1	MC	1, 4, 6	DOMC	2, 3, 5
B2	MC	1, 4, 6	DOMC	2, 3, 5
C1	DOMC	2, 4, 6	MC	1, 3, 5
C2	DOMC	2, 4, 6	MC	1, 3, 5
D1	DOMC	2, 3, 5	MC	1, 4, 6
D2	DOMC	2, 3, 5	MC	1, 4, 6

*Note.* Secure item sets are indicated using black text, while compromised item sets are in red.



# Classical Statistics

## Item $p$ -values

### Item Format and Test Security

Gorney &  
Wollack (2021)

### Background

### Purpose

### Method

### Results & Discussion

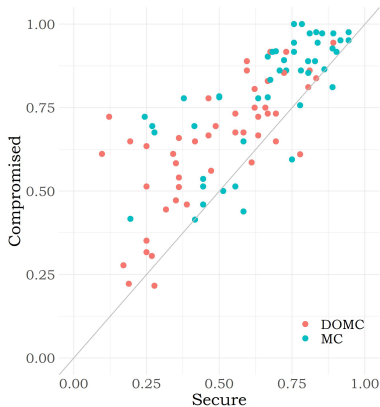
### Classical Statistics

### Item Response Theory (IRT)

### Conclusion

### References

- Secure items were generally more difficult than their compromised counterparts
- Similar pattern observed for both item formats



	DOMC	MC
Secure (Anchor)	0.50 (0.08)	0.61 (0.18)
Secure (Non-Anchor)	0.49 (0.21)	0.67 (0.20)
Compromised	0.64 (0.19)	0.80 (0.18)



# Classical Statistics

Item RTs (in Seconds)

Item Format and  
Test Security

Gorney &  
Wollack (2021)

Background

Purpose

Method

Results &  
Discussion

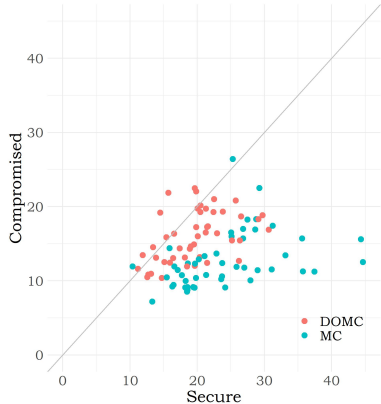
Classical Statistics

Item Response  
Theory (IRT)

Conclusion

References

- Compromised DOMC items answered 19% faster than secure DOMC counterparts
- Compromised MC items answered 47% faster than secure MC counterparts



	DOMC	MC
Secure (Anchor)	19.55 (5.23)	24.51 (5.35)
Secure (Non-Anchor)	19.57 (4.80)	24.07 (7.47)
Compromised	15.92 (3.43)	12.70 (3.77)



# Classical Statistics

## Item $p$ -values

### Item Format and Test Security

Gorney &  
Wollack (2021)

Background

Purpose

Method

Results &  
Discussion

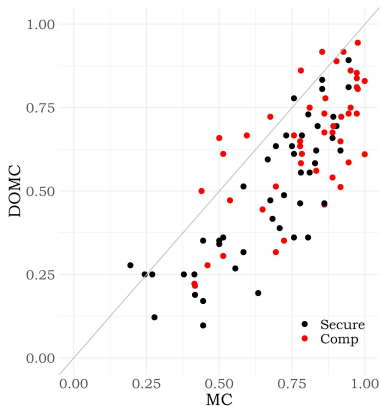
Classical Statistics

Item Response  
Theory (IRT)

Conclusion

References

- DOMC items were generally more difficult than their MC counterparts



	DOMC	MC
Secure (Anchor)	0.50 (0.08)	0.61 (0.18)
Secure (Non-Anchor)	0.49 (0.21)	0.67 (0.20)
Compromised	0.64 (0.19)	0.80 (0.18)





# Classical Statistics

Item RTs (in Seconds)

Item Format and  
Test Security

Gorney &  
Wollack (2021)

Background

Purpose

Method

Results &  
Discussion

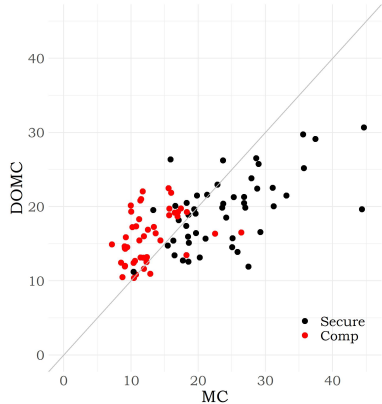
Classical Statistics

Item Response  
Theory (IRT)

Conclusion

References

- Secure DOMC items answered 19% faster than secure MC counterparts
- Compromised MC items answered 20% faster than compromised DOMC counterparts



	DOMC	MC
Secure (Anchor)	19.55 (5.23)	24.51 (5.35)
Secure (Non-Anchor)	19.57 (4.80)	24.07 (7.47)
Compromised	15.92 (3.43)	12.70 (3.77)



# Classical Statistics

## Key Position (DOMC Items)

### Item Format and Test Security

Gorney &  
Wollack (2021)

### Background

### Purpose

### Method

### Results & Discussion

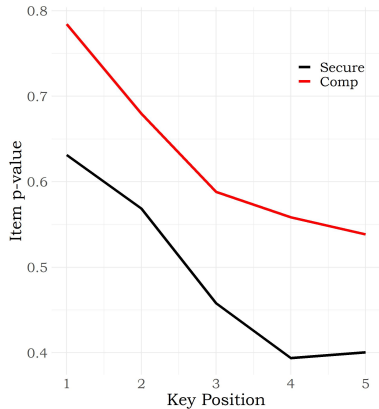
### Classical Statistics

### Item Response Theory (IRT)

### Conclusion

### References

- Items with later key positions tended to be more difficult
- This effect may level off after a certain point





# Classical Statistics

## Option Statistics (DOMC Items)

### Item Format and Test Security

Gorney &  
Wollack (2021)

### Background

### Purpose

### Method

### Results & Discussion

### Classical Statistics

### Item Response Theory (IRT)

### Conclusion

### References

- Secure and compromised distractors showed differences in RT, but not difficulty
- Secure and compromised keys showed differences in RT and difficulty

Compromise Status	Key	Option $p$ -value	Option RT (in Seconds)
Secure	No	0.83 (0.16)	5.13 (2.14)
Comp	No	0.81 (0.21)	3.75 (2.39)
Secure	Yes	0.70 (0.22)	4.84 (1.58)
Comp	Yes	0.86 (0.15)	3.18 (1.54)



# Item Response Theory (IRT)

## Item Format and Test Security

Gorney &  
Wollack (2021)

## Background

## Purpose

## Method

## Results & Discussion

## Classical Statistics

## Item Response Theory (IRT)

## Conclusion

## References

- Rasch model
- 1st calibration: secure items
  - True ability estimate
- 2nd calibration: compromised items, fix the secure item parameter estimates
  - Cheating DOMC ability estimate
  - Cheating MC ability estimate
- Compare cheating ability estimates to true ability estimates using bias, root mean squared difference (RMSD), and correlation



# Item Response Theory (IRT)

## Item Format and Test Security

Gorney &  
Wollack (2021)

## Background

## Purpose

## Method

## Results & Discussion

## Classical Statistics

## Item Response Theory (IRT)

## Conclusion

## References

## Bias

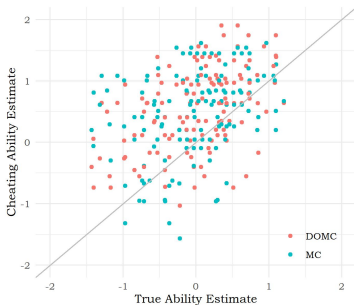
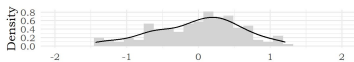
- DOMC: 0.55
- MC: 0.53

## RMSD

- DOMC: 0.72
- MC: 0.84

## Correlation

- DOMC: 0.47
- MC: 0.36





# Conclusions

## Item Format and Test Security

Gorney &  
Wollack (2021)

Background

Purpose

Method

Results &  
Discussion

Classical Statistics

Item Response  
Theory (IRT)

Conclusion

References

- Which item format is more affected by preknowledge?
- Similar score gains for both formats when items were compromised
- However, the DOMC format produced a lower RMSD and a higher correlation, thereby offering a slight advantage
- Future research
  - Additional real-data studies
  - Use existing preknowledge prevention and detection methods with DOMC items
  - Develop new methods specifically for DOMC items



# Thank you!

## Item Format and Test Security

Gorney &  
Wollack (2021)

Background

Purpose

Method

Results &  
Discussion

Classical Statistics

Item Response  
Theory (IRT)

Conclusion

References

For more information, please contact [kyliengorney@gmail.com](mailto:kyliengorney@gmail.com).



# References

## Item Format and Test Security

Gorney &  
Wollack (2021)

## Background

## Purpose

## Method

## Results & Discussion

## Classical Statistics

## Item Response Theory (IRT)

## Conclusion

## References

- 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.
- Tiemann, G., Miller, H., Kingston, N., & Foster, D. (2014, October 1–2). *Protecting item content via the discrete-option multiple-choice item type* [Oral presentation]. Conference on Test Security (COTS), Iowa City, IA.