

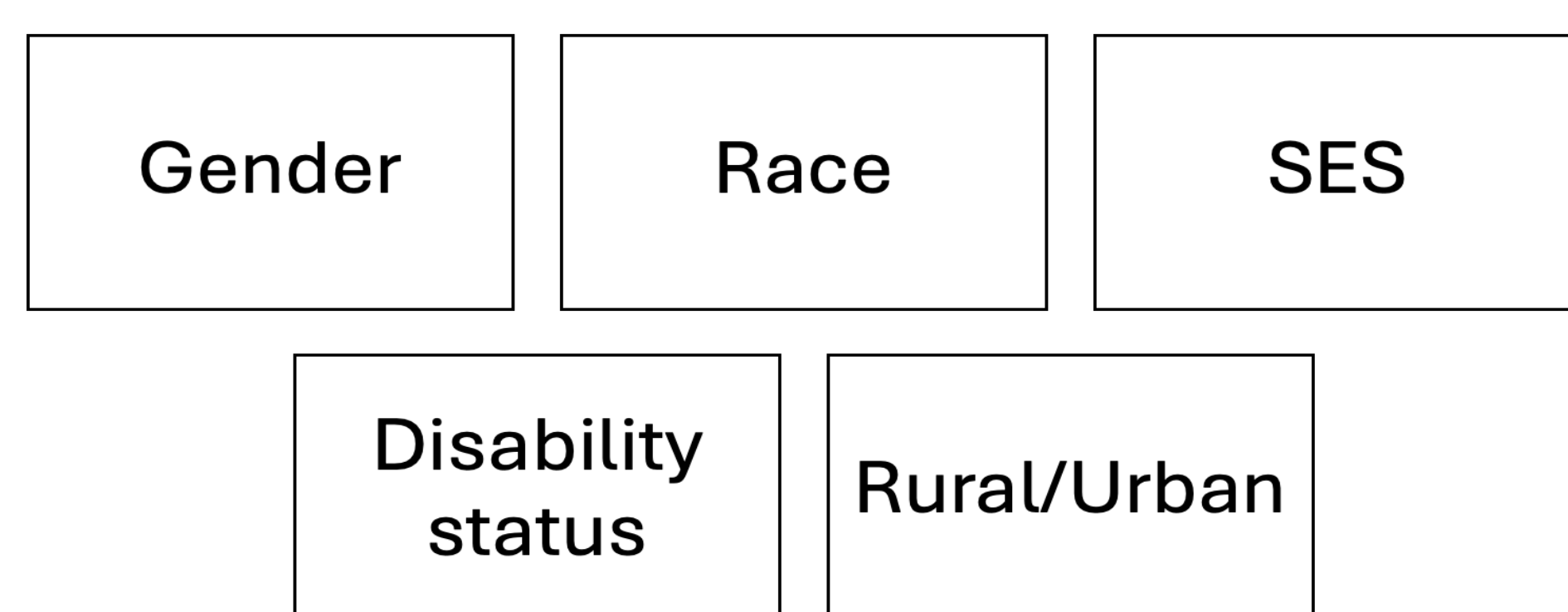
# A methodological review of intersectionality in differential item functioning, issues, and challenges.

Winifred Graham Wilberforce & Ann A. O’Connell, Ed.D., The Ohio State University

## Introduction

- Intersectionality, as explained by CRT literature examines how different social categories such as gender, race, sexuality, and class intersect and interact, influencing an individual’s experiences of privilege or oppression (Carbado et al., 2013).
- The adaptation of intersectionality in education research has been transformative, especially in how scholars and educators understand and address the complexities that come with identity, power, inequality, and privilege within education systems.
- The intersectional approach to differential item functioning (I-DIF) provides a new lens for psychometricians and quantitative researchers to contextualize the multiple identities of test-takers that may exist in their sample of interest.

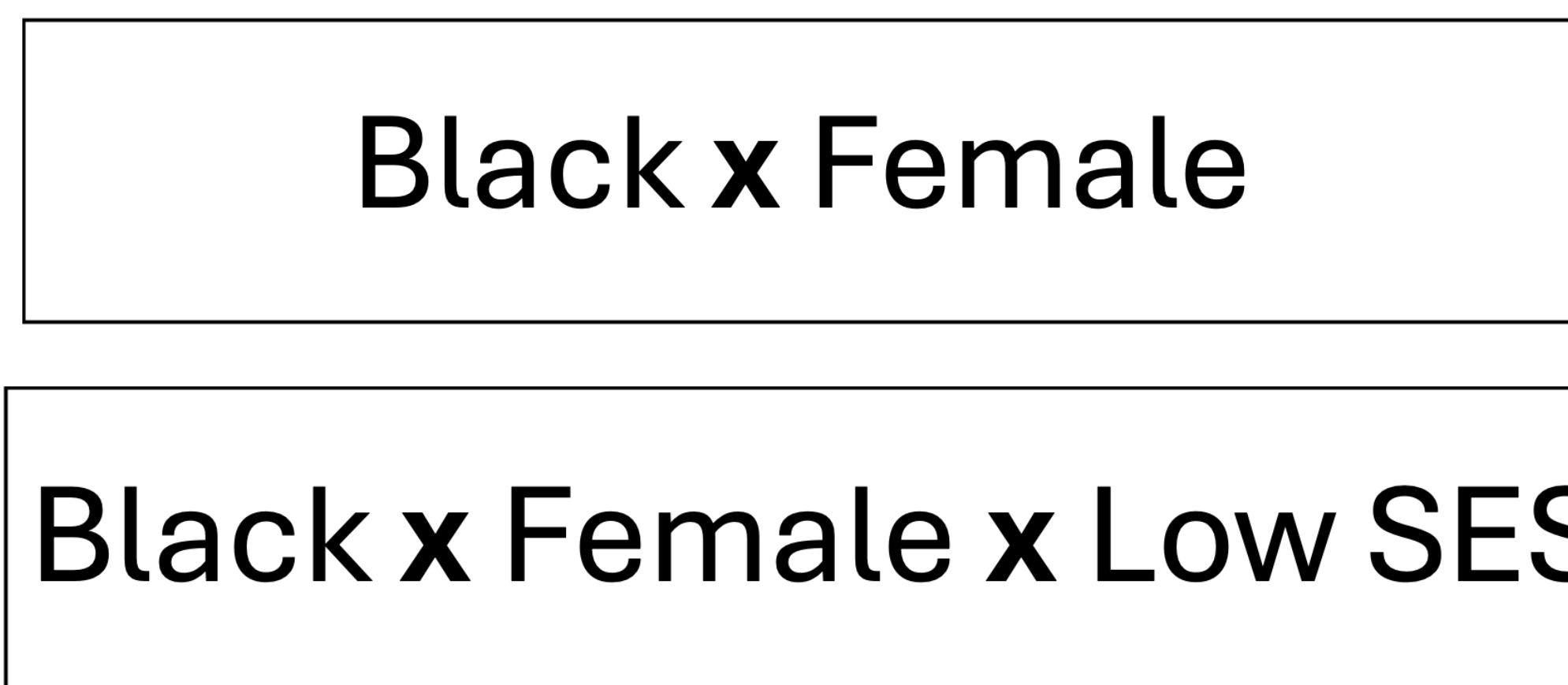
## Common Demographic groups used in DIF research



## Why I-DIF?

- Compared to traditional DIF methods which focus on isolated group differences, i.e. due to race or gender, the I-DIF approach allows for a broader and more accurate understanding of potential item bias, as it captures a more realistic intersection of test-taker identity and lived experience.
- Available literature suggests that the intersectionality idea is rooted in Critical Race Theory and operationalized in some Quantitative Critical Race Theory (QuantCrit) applications as an interaction between sub-groups of interest (Castillo & Babb, 2023).

## An example of intersectional groups in I-DIF Literature



## Research Questions

1. What is the intersectional approach to differential item functioning?
2. What are the limitations of the current methods used in I-DIF studies?

## Methods used I-DIF Literature

- Mantel-Haenszel (MH) (Adopted by Holland & Thayer, 1988; Holland & Wainer, 1993)
- Logistic Regression (LR) (Swaminathan & Rogers, 1990)
- SIBTEST (Shealy & Stout, 1993)
- Standardized-D (Dorans & Stout, 1993)

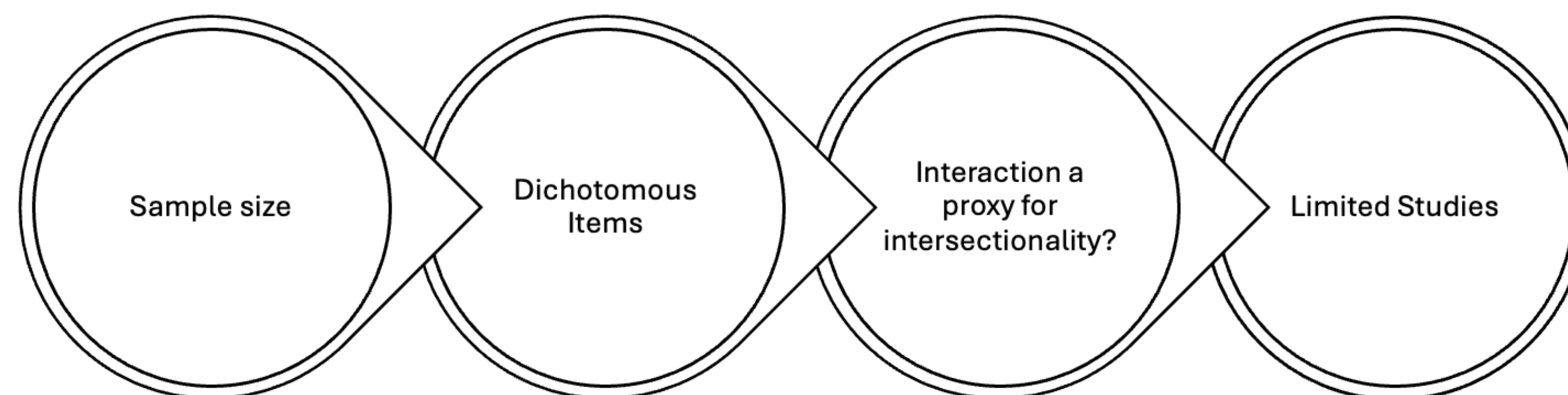
## Methodological & Practical Limitations Methods

- ❖ All four methods are sensitive to sample size (Russel et al., 2022; Albano et al., 2024).
- ❖ Logistic regression however is more sensitive, thereby producing results that are inconsistent with sample size fluctuation.
- ❖ The standardization method “does not account for chance increases in the rate of DIF detection” (Russel et al., 2023; Albano et al., 2024).
- ❖ The formation of an intersectional group poses statistical challenges (the likelihood of Type 1 error rates).

## What next.....

- ❖ Broaden current I-DIF approaches to include additional methods such as
  - Item Response Theory Models (IRT)
  - Generalized Linear Mixed Models (GLMM)
  - Bayesian Approaches
- ❖ Sample size guidelines for intersectional groups
- ❖ When does adjusting for Type I error rates become necessary and when is it redundant?
- ❖ How can test developers and researchers incorporate social justice perspectives such as intersectional identities in test/assessment development?

## Limitations



## Conclusion

To fully embody the tenets of QuantCrit, a reimagining of our instruments and procedures is necessary at all levels of educational testing and in research work. The integration of intersectionality and DIF research is a step in the right direction and the consideration of intersectional identities in measurement research will provide a comprehensive approach to examining and promoting test fairness.

E-mail [Wilberforce.2@osu.edu](mailto:Wilberforce.2@osu.edu) with suggestions.

## Bibliography

Albano, T., French, F., & Vo, T., t. (2024). Traditional vs International DIF Analysis: Considerations and a Comparison Using State Testing Data. *Applied Measurement in Education*.

Castillo, W., & Babb, N. (2023). Transforming the future of quantitative educational research: a systematic review of enacting quantCrit. *Race Ethnicity, and Education*. 27(1), 1-21.

Carbado, D., Crenshaw, K., Mays, V., & Tomlinson, B. (2018). Intersectionality: Mapping the Movement of a Theory. *Du Bois Rev*. Doi: 10.1017/S1742058X13000349

Crenshaw, K. (1989). Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics. *University of Chicago Legal Forum*: Vol. 1989, Article 8

Dorans, N. J., & Kulick, E. (1986). Demonstrating the utility of the standardization approach to assessing unexpected differential item performance on the Scholastic Aptitude. *Journal of educational measurement*, 23(4), 355-368.

Holland, P. W., & Wainer, H. (1993). *Differential item functioning*. New York, NY: Erlbaum.

Holland, P. W., & Thayer, D. T. (1988). Differential item performance Mantel-Haenszel procedure. In H. Wainer & H.I. Braun (Eds.), *Test validity* (pp.129-145). Hillsdale, NJ: Erlbaum.

Russel, M., & Kaplan, L. (2021). An Intersectional Approach to Differential Item review. *Functioning: Reflecting Configurations of Inequality. Practical Assessment, Research, and Evaluation*, 26, 21. DOI: <https://doi.org/10.7275/20614854>

Russel, M., Szendey, O., & Kaplan, L. (2021). An Intersectional Approach to DIF: Do initial findings hold across Tests? *Educational Assessment*, 26(4), 284-298. DOI: 10.1080/10627197.2021.1965473

Russel, M., Szendey, O., & Li, Z. (2022). An Intersectional Approach to DIF: Comparing outcomes across methods. *Educational Assessment*, 27(2), 115-135. DOI: 10.1080/10627197.2022.2094757

Shealy, R., & Stout, W. (1993). A model-based standardization approach that separates true bias/DIF from group ability differences and detects test bias/DTF as well as item bias/DIF. *Psychometrika*, 58, 159-194

Swaminathan, H., & Rogers, H. J. (1990). Detecting differential item functioning using logistic regression procedures. *Journal of Educational Measurement*, 27(4), 361-370