

**ILLUSTRATING AND ENACTING CRITICAL
QUANTITATIVE MEASUREMENT
WITH MIMIC MODELS**

MATTHEW A. DIEMER, MICHAEL B. FRISBY,
AIXA D. MARCHAND & EMANUELE BARDELLI

PROBLEMATIC HISTORY OF QUANTITATIVE METHODS

- Zuberi and Bonilla-Silva (2008) in *White Logic, White Methods: Racism & Methodology*, document this troubling history. A brief overview:
- "Francis Galton was a key intellectual power behind the modern statistical revolution in the social sciences....Galton used statistical analysis to make general statements regarding the superiority of different classes within England and of the European-origin race, statements that were consistent with his eugenic agenda." (p. 128)
- **Founded the ideology and scientific practice of eugenics**



PROBLEMATIC HISTORY OF QUANTITATIVE METHODS

- Galton's work was carried on by eugenicist Karl Pearson (1857-1936), who further developed the comparative method of the Pearson correlation coefficient to prove the intellectual superiority of the Aryan race. This has been well documented (Dixon-Román, 2019, p. 95)
- Mid-1990s saw the *Bell Curve* & modern-day attempts to establish (White) racial superiority, using quantitative methods (much of this measurement-focused)
- Note: qualitative inquiry criticized as colonial (Eve Tuck), racist (Smith), and/or exploitive (e.g., 'Sneaky Kid' & Harry Walcott)

PROBLEMATIC HISTORY OF PSYCHOMETRICS

These histories have led many to believe: “it remains difficult to utilize the statistical tools created by eugenicists to study R/E [race/ethnicity] in ways that do not lead to perpetuating that inequality.” (p. 304, Viano & Baker, 2020)

So, should we ‘throw the baby out with the bathwater’ and abandon quantitative methods if we want to center race, conduct anti-racist research, and/or maintain a critical perspective?



QUANTITATIVE METHODS: POSSIBILITIES

Let's not "throw the baby out with the bathwater"...

"quantitative methods can be, have been, and should be used by scientists (along with other methods) to achieve the goals of social justice." (p. 27, Cokley & Awad, 2013)

[These approaches] "will not dismantle systemic racism, however, it is a *tool* to begin to reimagine the role that research and data can play in an anti-racist society." (p. 3, Castillo & Gilborn, 2022)

But let's not pretend that our baby is perfect

While mindful that: "It follows that every attempt to 'measure' the social in relation to 'race' can only offer a crude approximation that risks fundamentally misunderstanding and misrepresenting the true nature of the social dynamics that are at play" (Gilborn et al., 2018; Helms et al., 2015)



ANNUNCIATION: CRITQUANT

Critical Quantitative (CritQuant, or CQ) perspective

A more general critical approach to quant methods, in solidarity with but not as paradigmatic/specific to CRT as QuantCRiT (Stage, 2007)

Guiding idea: “quantitative analysis can support other models of inquiry to serve as a valuable tool to challenge notions of race and racism in contemporary contexts” (p. 256, Knowles & Hawkman, 2020).

- What questions are asked & how they are posed, rather than specific and/or new methods

CRITICAL QUANTITATIVE METHODS: CORE ELEMENTS

1. High level understanding of critical theory & methods (Randall, 2021; Stage, 2007)

2. Knowing methods in order to critique, deconstruct, and reformulate them:

“[Alexander McQueen] learned the precision and skill in tailoring suits that later helped him deconstruct without losing structure or integrity of the garment. This approach of mastering the rules well enough to know how to break them stayed with me” (P. 93, Hernandez, 2018)

a) “mastery is required in order to demolish the rules but keep tradition” (p. 99, Hernández, 2015)

3. Numbers don't have more inherent truth or rigor:

“This does not mean that critical race theorists should dispense with quantitative approaches but that they should adopt a position of principled ambivalence, neither rejecting numbers out of hand nor falling into the trap of imagining that numeric data have any kind of enhanced status, value, or neutrality.” (p. 174, Gilborn et al., 2018)

CRITICAL QUANTITATIVE METHODS: CORE ELEMENTS

Acknowledgement of subjectivity/positionality (adopted from qualitative inquiry)

4. Perspective: scholarship/methods are political; no such thing as value-free or objective inquiry

- a.** “No such thing” as culture- or context-free assessment/measurement (Randall, 2021)

5. Researcher self-reflexivity across the research enterprise: “a researcher practicing self-reflexivity and making decisions with intention may be able to shape their use of those statistics toward an anti-racist agenda.” (p. 543, Suzuki et al., 2021)

5a. Use of researcher positionality statements (adopted from qual inquiry)

- a.** Acknowledgement of subjectivity and positionality is a strong departure from common assumption of objectivity/neutrality in quant methods

MIMIC MODELS

- Within the broader CQ perspective, MIMIC models are a strategy to identify and challenge racism and other forms of inequity in measurement
- **MIMIC models:** I suggest as one of the “pragmatic strategies for identifying or removing unfairness from individual test takers’ scores if construct-irrelevant variance is discovered” (Helms, 2006, p. 846)
 - > if that construct-irrelevant variance is *group-based measurement error*
 - > this is a way of assessing *construct bias* to ensure fairness in measurement (Helms, 2006)
- i.e., “random error variance or factors that are wholly unrelated to individual differences in the trait measured by the test” (Frisby, 1999, p. 264).
- MIMIC models would identify items that exhibit bias, in that they function differently across groups, one dimension of test fairness

WHAT DO MIMICS ESTABLISH?

“Strong Invariance” (a.k.a. Scalar invariance) -> presuppose configural (same ‘configuration’ of boxes to circles’) & metric (same ‘metric’ of loadings across groups’) invariance

Unstandardized intercepts between groups are constrained to be equal.

Two people from different groups with the same level on a certain factor have the same score on a given indicator

Scalar Invariance: Practical example: men and women with equal levels of depression self-report the same amount of binge eating

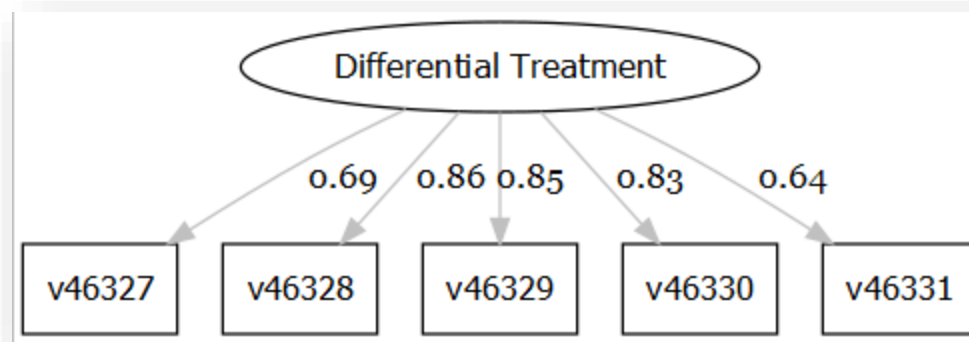
Vs Scalar variance (i.e., uniform DIF) :

At the same level of depression, women binge-eat twice per week and men binge-eat 4 times per week, even though the slopes (increases as depression increases) may be the same

Briefly, DIF detected via MIMICs equivalent to DIF via IRT: “IRT parameters converted from MIMIC models are comparable in magnitude to those produced by widely used IRT software” (p. 378, McIntosh & Hashim, 2003).

Let’s pivot to an empirical example ->

CFA: DIFFERENTIAL TREATMENT [MADICS]

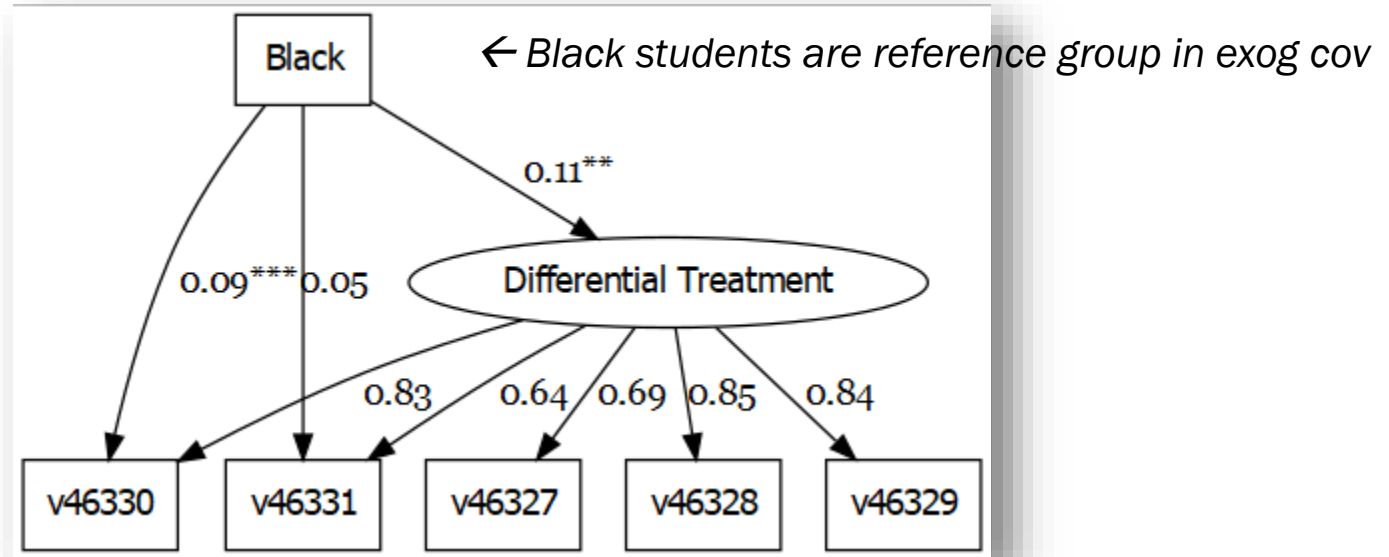


- MADICS dataset: Black and white respondents in 11th grade (W4), $N = 782$
- V46327: teachers call on you less because of your race
- v46328: teachers grade you harder because of your race
- v46329: you get disciplined more harshly because of your race
- v46330: teachers think you are less smart because of your race
- v46331: teachers discourage you from taking classes because of your race

Scaling: 1 (“never”) to 5 (“more than six times”), greater scores = more perceived differential treatment

- CFI = .98, TLI = .96, RMSEA = .06, SRMR = .03
 - Chi-square = 19.98, $p = .001$
- *All coefficients standardized*

MIMIC: DIFFERENTIAL TREATMENT



- Significantly higher latent mean for DiffTreat ($\beta = .11, p = .001$)
- DIF: (v46330) Black students more frequently felt teachers found them less smart because of their race ($\beta = .09, p < .001$)
- Approaching sig DIF: (v46311) Black students more frequently felt teachers discourage them from taking classes because of race ($\beta = .05, p = .06$)
- **Even after** adjusting for Black students' perceived higher levels of racial discrimination (significant latent mean differences by race)
 - 'undermeasuring racism'

- CFI = .98, TLI = .95, RMSEA = .05, SRMR = .03
 - Chi-square = 24.36, $p = .001$
- *All coefficients standardized*

MIMICS: AFFORDANCES

Within the broader CQ perspective, MIMICs represent one approach to identify and model how racism (and other forms of inequity) contribute to measurement

As demonstrated here, even in a dataset designed to capture the experiences of Black students, items “undermeasured” anti-Black racism in schools

MIMIC models simple, yet effective, strategy to detect and control for items that may be biased

Note: This does not establish Measurement Invariance (MI), a more intensive & stepwise process

Sample size efficient: Because sample is not split into multiple groups (a la measurement invariance), MIMICs are feasible with smaller samples

This is particularly important in studying minoritized populations, who by definition are numerical minorities -> more difficult to recruit & retain (in longitudinal samples)

MIMICS: LIMITATIONS

- Group membership is modeled as a dichotomy -> oversimplifies racial/ethnic identification, intersectional identities, etc into a 0 or 1
 - Tradition to specify dichotomous exogeneous covariates in MIMICs, yet continuous covariates are mathematically possible & better reflect the nuance of contemporary theory
- “The multifaceted nature of race and ethnicity suggests that when race is operationalized as a stable, homogenous entity (e.g., a simple dummy or categorical variable like “1” if white, “0” if nonwhite), any statistical association will typically offer little or no insight as to which elements are the key mechanisms of action—be it fear of an out-group, neighborhood effects, or some other factor.” (p. 517, Sen & Wasow, 2016).
 - MIMICs do not capture the specific mechanisms (e.g., stereotype threat, racial identity status) that contribute bias (i.e., construct-irrelevant variance) to scores
- MIMICs measure racism by proxy (racial/ethnic category)
 - ‘Race as category’ measures structural racism, inequitable policing, anti-Blackness indirectly via their manifestation in each category

CONCLUSION

- CQ, QuantCRiT increasing attention lately (e.g., SREE webinar series, journal special issues)
- Both hold quant methodology to account for its history -> but also repurposing these approaches to “do better” (e.g., identify and remediate racism in measures)
- MIMICs hold promise because of their (relative) simplicity, and because of their potential to help identify and (as illustrated here) challenge racism in measurement
- Working paper on Crit Quant & MIMIC modeling (w/ lavaan/Mplus/Stata code) can be shared by contacting first author
- Thank you!