"SIS, ARE YOU OKAY?": A FACTOR ANALYSIS TO IDENTIFY THE RAMIFICATIONS OF SYSTEMIC RACISM ON THE PSYCHOLOGICAL WELL-BEING OF BLACK WOMEN IN U.S. AMERICAN ACADEMIA

Authored By

BENNETT McAULEY
KATELYN SETTLEMYRE
&
DIDIER TURCIOS

North Carolina State University
Department of Statistics

in partial fulfillment of the requirements for
the degree of Master of Statistics;
and in collaboration with our client et al.:

JOCELYN TALIAFERRO, MSW, PhD

North Carolina State University
School of Social Work

&

DIAMOND MEADOWS, MSW

East Tennessee State University
Department of Psychology

May 2023

1 Introduction

The United States has a persistent history of encouraging and enforcing discrimination, oppression, and other forms of prejudice against African Americans, who are often simply characterized as "Black". The turbulent relationship between the Black community and the institutions of government, education, healthcare, law enforcement, social services, and even housing, is a compounded byproduct of historical bylaws–most infamously, the era of Jim Crow laws (c. 1877-1964)–that have yet to be unraveled from the fabric of U.S. American culture and politics.

Racially-motivated violence has been the cornerstone of the Black experience for the entirety of the community's documented history. Paradoxically, Black people are more susceptible to bouts of national trepidation but are expected to show the most resilience. With the additional intersection of gender, Black *women* are especially vulnerable but are often neglected in policies and actions designed to alleviate the stressors of U.S. residents.

In the spring of 2020, the deficit of adequate resources for Black women was proven evident with the surge of COVID-19 and the subsequent deaths of George Floyd and Breonna Taylor. Even with nationwide attention on these issues, Black women continued to suffer in silence under the pretense that they are "strong": perpetually shouldering the burdens of others with no space for their own (Pappas, 2021). Consequently, Black women as a population experience disparities in quality of life that no other group in the U.S. do, most of which can be attributed to social-environmental factors. They are at higher risk for breast cancer, stroke, diabetes, and hypertension; the Black community overall experienced two to three times more complications or death from COVID-19 than the general U.S. population (Menifield & Clark, 2021). Concerning mental health, Black women are half as likely to seek services compared to their white counterparts, and documented studies suggest one reason for this is the lack of cultural competency among providers (Abrams et al., 2018).

In the context of this study, the population of interest is Black women in academia. They make up 4% of college and university faculty, among whom they must navigate unwelcoming–sometimes hostile–environments (NCES, 2020). The barriers previously mentioned have contributed to the marginalization and isolation of Black women in academia. By examining potential associations extracted from respondent data, methods for healthy and impactful coping mechanisms can be identified. Data is derived from a survey conducted in April 2022, to learn how Black women were coping post-pandemic. Approximately 2680 participants responded to 70 questions. In addition to demographic information, the topics covered in the survey included COVID-19, anti-Black racism, depression, anxiety, and coping. The analysis presented in the following sections is driven by structural modeling methods. In lieu of significance testing, the objective of the study is to determine prominent factors that influence the extent to which Black women suffer from depression, anxiety, and other mental health issues.

2 Data Analysis Methodology

2.1 The Data

Main points:

- Collected using paid panel

- 70 question survey to find out how Black women were coping

- Asked about Covid, Racism, Depression, Anxiety, Coping, and Demographics

- Screened out participants to only include consenting Black adult (18+) females

- Raw data contained 157 variables, many removed for analysis such as personal information and variables that were recoded

- ~79 variables left after subsetting to remove and removing irrelevant variables

- Went from 2631 observations to 2260 after removing those who were screened out

- Data cleaning (such as updating character entries found in a numeric variable), handle missing values (such as age) and repeats (survey refresh/kicked out and restarted)

- *what else?*

2.2 Structural Equation Modeling (SEM)

Structural Equation Modeling (SEM) is a technique for investigating relationships by specifying a model that represents predictions of that theory among plausible constructs measured with appropriate observed variables (Hayduk, Cummings, Boadu, Pazderka-Robinson, & Boulianne, 2007). SEM can be thought of as a combination of regression analysis and factor analysis; it uses *latent* variables or constructs that are measured by multiple observed variables or indicators. For example, a SEM model can be used to estimate the association between outcomes, such as levels of depression or anxiety, and underlying attitudes that influence those, such as coping mechanisms and the effects of Covid-19 or anti-Black racism. SEM allows us to include multiple influences, to allow unobserved concepts that control for the observed indicators, to specify how those concepts influence one another, to assess the model's overall compatibility to the data, and to determine whether how well the model fits the data (Chapman & Feit, 2019).

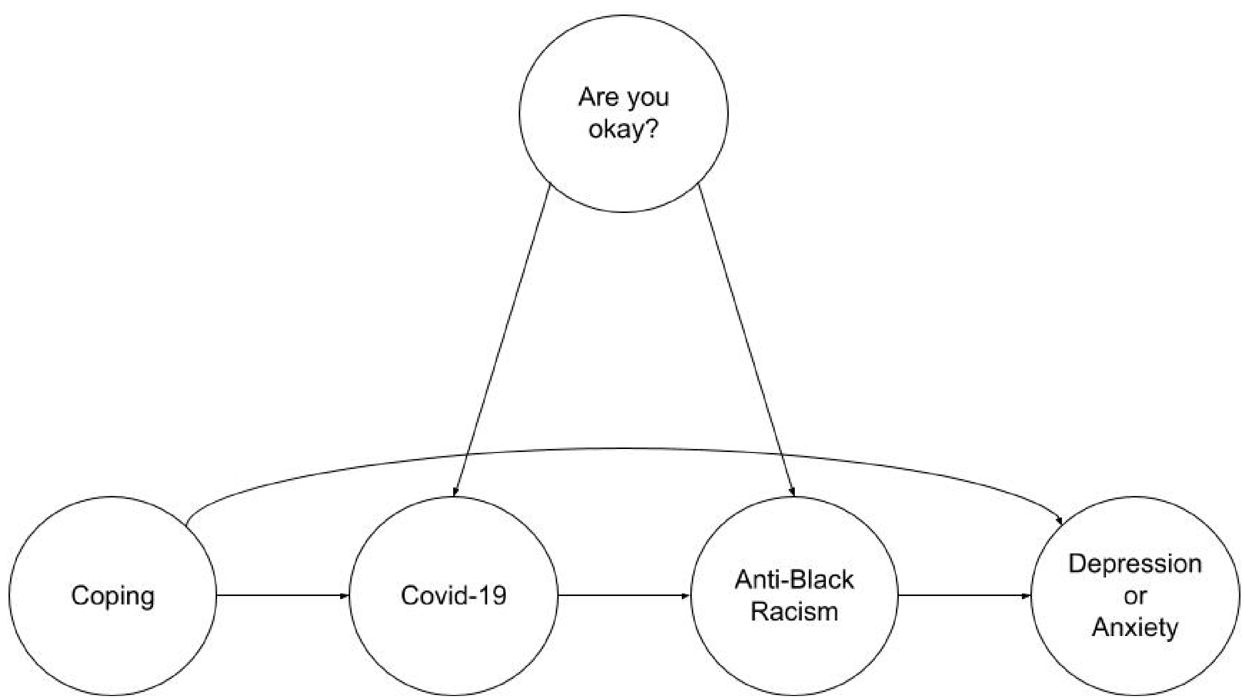

Figure 1: A model measuring levels of depression and anxiety. In this model, coping ability and strategies is associated with Covid-19 and anti-Black racism factors, while the overarching response to *Are you okay?* closely relates to both Covid-19 and anti-Black racism influences, which is then associated to levels of depression and anxiety.

To use SEM as an analysis tool, a graphical path diagram of influences is first created. Then, the strengths of the relationships for each path in the model are estimated. The paths in the model can be categorized as either observed variables, i.e., have data points, or latent variables that may underlie the observed data. A model is then fitted to the data, using the structural model paths created. Figure 1 demonstrates how an SEM model can be constructed to investigate the objectives of this study. For example, the Covid-19 latent variable can be observed as survey items Q1, Q2, and Q3. More complexly, the Coping latent variable can be observed by both survey items *and* other latent variables (i.e., Covid-19, anti-Black racism, etc). Upon fitting the SEM model, the strength of the relationships between the latent variables are compared, as well as the degree to which the model fits the observed data. Alternative models may then be fitted and compared to the original (full) model. How well a model fits the data is not the goal of SEM, however.

2.3 Exploratory Factor Analysis (EFA)

In Chapter 8.3 of *R For Marketing Research and Analysis*, Chapman and Feit (2019) describe Exploratory Factor Analysis (EFA) as a family of techniques used to assess the relationship of *constructs* (concepts) in surveys and psychological assessments. Factors are regarded as *latent variables* that cannot be observed directly–as one might in a simple linear model–but are associated through their relationship to other variables. For example, abstract concepts such as intelligence, personal preference, and emotions are not observable independently or empirically. Instead, they are measured with a number of indicators known as *manifest variables*, which are empirical factors and behaviors(e.g. GPA, survey responses, and so on) that, in tandem, provide an imperfect explanation to the underlying latent variables.

The goal of EFA is to find the degree to which the latent variables account for observed the variance of the manifest variables. The result is similar to Principle Components Analysis (PCA), where linear combinations or mixtures of the initial variables reduce the dimensionality of the data while retaining as much information as possible. What separates EFA from PCA, however, is the interpretability of the loadings on the variables. In other words, EFA optimizes solutions that favor intuition in terms of the original variables so the user can test whether the data is consistent with expectations.

To find a solution using EFA the following steps are generally implemented with a variety of tools and techniques:

1. **Determine the number of factors to estimate and test**. In R, this can be done by constructing a *scree plot*, a line plot of the eigenvalues of factors or principal components and retaining the ones that are greater than 1.0. An eigenvalue of 1.0 corresponds to the amount of variance that can be attributed to a single independent variable.

2. **Extract factors for modeling**. Determine the variable loadings using PCA, maximum likelihood, or other methods to select the k (limited to the number suggested in Step 1) factors to include and interpret. Trying more than one value of k–none beyond $k \pm 1$–is recommended to determine which model yields stronger interpretability.

3. **Rotate the loadings**. The purpose of rotation is to determine whether correlation between factors should be allowed or not. Software cannot interpret the context of the *latent variables*, but rotation can assist in deciding if, conceptually, factors should be considered independent or if it makes more sense for them to be related. One of two types of rotations can be applied: *orthogonal* and *oblique*. Orthogonal rotation, as its name implies, assumes that factors *do not* correlate. While this method is the default among most software and languages, it is not always the more plausible. Oblique rotation, on the other hand, assumes that factors can and *do* correlate, and considers how.

2.4 Why these methods?

Main points

- High dimensionality in our data

- These are dimension reduction methods

- Interest in predicting anxiety and depression, which have many influencing (manifest) variables and thus cannot be observed directly (anxiety and depression $\Rightarrow$ latent variables)

- Why not regression? regression aims to predict, while EFA and SEM aim to explain relationships between variables

- PCA is also a dimension reduction method, but we lose interpretability and new factors are uncorrelated

- EFA takes correlation into account – we can choose uncorrelated (orthogonal) or allow correlation (oblique) [we might need to allow correlation]

- part of PCA vs EFA from Wikipedia: "Factor analysis is clearly designed with the objective to identify certain unobservable factors from the observed variables, whereas PCA does not directly address this objective; at best, PCA provides an approximation to the required factors. From the point of view of exploratory analysis, the eigenvalues of PCA are inflated component loadings, i.e., contaminated with error variance."

- From ST 537 notes: PCA makes no model assumption, attempts to capture most of the total variance, doesn't account for measurement errors in observed vars, orthogonal by construction, and PCs are linear combos of vars. EFA assumes a specific model and estimates params, tries to max the variance due to common factors, can accommodate measurement errors in vars through specific factors, is able to have oblique models that allow for correlation, and the vars are linear combos of latent factors.

- *need more on SEM*

- *EFA and SEM are similar. How do they differ and why do we want this?*

References

Abrams, J. et al. 2018. *Underneath the Mask of the Strong Black Woman Schema: Disentangling Influences of Strength and Self-Silencing on Depressive Symptoms Among u.s. Black Women.* Sex Roles, 80(9-10), 517–526. https://doi.org/10.1007/s11199-018-0956-y.

Chapman, E., C. & Feit. 2019. *R for Marketing Research and Analysis, Second Edition.* Springer Nature Switzerland AG.

Menifield, & Clark, C. E. 2021. *Pandemic Planning in the United States: An Examination of COVID-19 Data.* Public Administration Review, 81(6), 1102–1109. https://doi.org/10.1111/puar.13326.

National Center for Education Statistics. 2022. *Characteristics of Postsecondary Faculty. Condition of Education.* U.S. Department of Education, Institute of Education Sciences. https://nces.ed.gov/programs/coe/indicator/csc.

Pappas, S. 2021. *Effective Therapy with Black Women.* American Psychological Association. https://www.apa.org/monitor/2021/11/ce-therapy-black-women.