



Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Rejón, R

Title:

Automated transcoloration? Skin tone classification algorithms and the prominence of non-epidermic features in the Mexican racial schema

Date:

2025

Citation:

Rejón, R. (2025). Automated transcoloration? Skin tone classification algorithms and the prominence of non-epidermic features in the Mexican racial schema. *Ethnicities*, <https://doi.org/10.1177/14687968251319799>.

Persistent Link:

<https://hdl.handle.net/11343/355604>

License:

[CC BY-NC](#)

Automated transcoloration? Skin tone classification algorithms and the prominence of non-epidermic features in the Mexican racial schema

Ethnicities

2025, Vol. 0(0) 1–20

© The Author(s) 2025



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/14687968251319799

journals.sagepub.com/home/etn**René Rejón** 

Faculty of Arts, University of Melbourne, Parkville, VIC, Australia

Abstract

The scale of anti-racist research employing independent measures for skin tone (i.e., classification algorithms and colorimeters) is increasingly growing. However, recent studies reveal that non-epidermic traits (i.e., hair, lips, gender) contribute to the *transcoloration* –whitening or darkening– of perceived skin tone, signaling that their importance in the racialization process has been underestimated. The magnitude of this effect of non-epidermic features is critical, and the appropriateness of technology-based anti-racist research hinges on it. To what extent are independent measures of skin tone a good approximation to the study of racial discrimination in Mexico? Examining a novel dataset ($n = 3000$), I compare two different skin color variables (one coded by humans and other by an automated algorithm), finding little agreement about exact estimates but substantial—and statistically significant—agreement on the general direction of the variables. These findings corroborate that non-epidermic features slightly alter people's perception of skin tone, without causing a major digression from independent skin color assessments. The results validate and encourage research employing these technological methods, which tend to be more accessible and feasible than their alternatives.

Keywords

Skin color, racial schema, Mexico, structural injustice, racism, colorism

Corresponding author:

René Rejón, Faculty of Arts, The University of Melbourne, Room 148, Old Arts Building, Parkville Campus, VIC 3010, Australia.

Email: rene.rejon@unimelb.edu.au

Data Availability Statement included at the end of the article

Introduction

Racial hierarchies continue to map onto class structures virtually all over the globe (Bonilla-Silva, 2015; Clealand, 2017; De Micheli, 2018; Dixon and Telles, 2017; Keith and Herring, 1991), and Mexico is no exception (Arceo-Gomez and Campos-Vazquez, 2014; Campos-Vazquez and Medina-Cortina, 2019; Moreno, 2008, 2010; Sue, 2020).

Biologically baseless, the social construction of “race” persists pervasively. Through racial schemas, people assign social meaning to certain physical markers –i.e., skin tone, lip shape, hair texture (Bonilla-Silva, 2015; Haslanger, 2012; Sen and Wasow, 2016; Wade, 2012). These racial schemas are complex social phenomena. Like most other invented social categories, race is instable, slippery and mutable (Bonilla-Silva, 2015: 1360; Omi and Winant, 2014: 4). Often, factors such as nationality, culture and birthplace are also invoked to explain the concept (Chandra, 2006; Lancaster, 1991; Loveman, 1997; Marrow, 2003; Nutini, 1997).

Reaching a better understanding of the configuration and dynamics of racial schemas is particularly important in a region with documented “racial fluidity” (De Micheli, 2018: 53, 2021: 6; Irizarry et al., 2023; Marrow, 2003). In Latin America, for decades, race and skin color were treated as overlapping –even interchangeable and indistinct– concepts (Telles, 2004: 79, 2018: 1166). In Brazil, for instance, most people singled out skin color to justify their racial self-identification (Nogueira, 1955) and, to this day, electoral verification commissions use phenotypic criterion to validate the racial self-identification of candidates (Moraes Silva et al., 2024: 7). Moreover, some studies in the region use skin color as a proxy for race (Rejón, 2024; Telles et al., 2015; Trejo and Altamirano, 2016) and researchers in other Latin American countries concur that “race” mostly refers to skin color and that these two terms could be used interchangeably (Htun, 2015: 165; Mitchell-Walthour, 2017: 1; Saldívar, 2014: 90).

However, some scholars question the uncritical “interchangeability” of the terms. According to critics, equating the two is a mistake for at least a couple of reasons. On the one hand, studying race via skin color is to pay attention to only one of the many markers in the phenomenon of race (Monk, 2021: 80). On the other hand, because of the potential interdependence of the elements within the racial schema, “skin color” is difficult to measure and, often, researchers are not investigating what they think they are (Monk, 2016: 415).

Addressing these questions, recent studies have started to test the assumptions and examine whether skin color is indeed the best predictor for racial identification (Monk, 2016; Telles and Paschel, 2014). In Mexico, the most notable example –if not the only one– is the Project on Ethnoracial Discrimination (PRODER), a large-n sample survey that explores the interaction between different racial identifications (i.e., race, ethnic group) and skin color measurements with socioeconomic outcomes. The differentiated findings among various modes to account for racial and skin tone identification, reveal that the Mexican racial schema is more complex than what the “interchangeability” assumption suggests.

These findings raise the question: what are the particular configurations and dynamics of the Mexican racial schema? This issue is especially relevant given the recent increase on the number of studies using “objective” or “independent” skin tone measurements

(i.e., classification algorithms and colorimeters) to analyse various aspects of social life, such as sociological outcomes, electoral success and political representation (Campos-Vazquez and Rivas-Herrera, 2021; Dixon and Telles, 2017; Rejón, 2024; Rejón Piña and Ma, 2023). The extent to which these tools are appropriate and helpful “replacements” for human-coding hinges on their capacity to approximate to the results of a human discrimination process that comprises so much more than a simple assessment of skin tone. If the independent skin tone measurements of these tools diverge significantly from the “skin tone” humans identify influenced by other factors in the racial schema, then the machine-coded variables are less appropriate methods for anti-racist research, as they fail to sufficiently approximate the “skin tone” detected in actual human interaction. On the contrary, if the assessments between algorithms and humans are similar, then the validity of these –still imperfect, but helpful– tools to approach the study of colorism and racism is confirmed.

To shed light into applicability of these tools in relation to the configuration and functioning of the Mexican racial schema, this article examines one sample from two different “skin color” angles: one that independently measures skin tone and one in which humans subjectively assess complexion. Methodologically isolating skin color from other social cues (via an automated algorithm) and comparing this measure with one susceptible to other (admittedly not all) biases (human-coding), helps reveal the extent to which these automated tools are a helpful and appropriate method to facilitate social research in this topic.

To achieve its aims, the article is structured as follows. First, I outline the theoretical framework to which this analysis is circumscribed to; I explain how –through racial schemas– race is actioned for social categorization purposes in everyday life and people are racialized through an undeniably corporeal and ocular dimension. In the following section, I describe the sociohistorical context of the “race issue” in Mexico; reviewing how the current notions of race are rooted in the myth of *mestizaje*, promoted by elites in the 20th Century. In “Methods and data”, I describe the novel dataset of 3,000 portraits I use to compare two different measures of skin color. The first was gathered using CASCo –an automated algorithm that focuses exclusively on the color of the face area of the photo– and the second is a measure of “impressionistic” race, produced by independent human coders. In “Findings and discussion”, I present and analyze the results of the study. I find little inter-coder agreement about exact estimates but substantial –and statistically significant– agreement between the general direction of the two skin color variables. These findings confirm that CASCo –and, arguably, similar tools– provide measurements that do not substantially digress from people’s perception of “skin tone”. I conclude the article with a brief discussion on what these findings entail for the applicability of these tools in social research, arguing that –given their accessibility and satisfactory approximations– they are not only convenient but valid methods to reasonably substitute human coders.

Theoretical framework

Races are invented categories, but they are (socially) real because they are reenacted in the everyday life (Bonilla-Silva, 2015: 1360). They do not exist as biological facts, but are real social forces (Atkin, 2012; Mallon, 2006).

While socially constructed, race has a material foundation (Bonilla-Silva, 2015: 1360) and there is a crucial corporeal dimension to it. According to Omi and Winant (2014: 4), race is “ocular in an irreducible way”. Bodies are “racialized” –read and interpreted– via symbolic meanings and associations; and it continues to identify groups and individuals via somatic markings (i.e., skin tone, lip shape, hair texture) that still signal social status and cause different forms of interaction (Haslanger, 2012: 10). Depending on context, these markers can be cultural, phenotypical, or both. Therefore, some explain race as a “bundle of sticks” or a “racial schema” made of many factors such as societal values, cultural traits and physical attributes (Sen and Wasow, 2016: 506; Wade, 2012).

Note that conceptualizing race as a socially constructed bundle of sticks makes it easy to understand as a category of analysis, but it does not say much about how race is operationalized as category of practice. What is the particular configuration of racial schemas? Are some of their elements more prominent than others? The answers to these questions are not obvious. Perhaps it is for this complexity that the recommended standard to measure and operationalize race is self-identification (CEPAL, 2020; Del Popolo and Schkolnik, 2012).

For sociohistorical reasons, however, in some contexts many people do not identify as members of a racialized groups even if they are racialized (Del Popolo et al., 2009: 63; Sue, 2013; Vaughn, 2013), which complicates the struggle for racial justice. Trying to overcome this obstacle, some scholars resort to the use of skin color as a proxy for race (Rejón, 2024; Telles et al., 2015; Trejo and Altamirano, 2016), sometimes arguing that complexion is the best predictor of racial identification (Banton, 2012; Telles and Paschel, 2014) or even that race and skin color are overlapping –even interchangeable and indistinct– concepts (Htun, 2015: 165; Mitchell-Walthour, 2017: 1; Saldívar, 2014: 90; Telles, 2004: 79, 2012: 1166).

However, this approach has been challenged recently. Critics argue that studying race via skin color is to pay attention to only one of the many markers in the phenomenon of race (Monk, 2021: 80). Furthermore, scholars have criticized the way in which “skin color” is measured, arguing that often researchers are not investigating what they think they are (Monk, 2016: 415). It turns out that measuring skin color is much more complex than social scientists initially thought.

Researchers have measured skin color employing different methods. Using verbal scales that represent a color spectrum (i.e., ranging from “dark” to “light” or “black” to “white”) is probably the most common of all. However, despite its popularity, the approach has been criticized for the way in which these cues skew and influence the answers to the questions asked by imposing rigid categories, even if these are deemed appropriate for prompting racial identity (Harris et al., 1993). Regardless of who is entering the data –interviewer or respondent–, this method also ignores the fact that people are not neutral and detached recorders of physical traits. For instance, research demonstrates that people have limited ability to accurately differentiate physical features of others (Hill, 2002) or even their own (Monk, 2015). Furthermore, skin color self-identification often differs from observers’ classification (Campbell and Troyer, 2007; Pereira and Telles, 2014) and racial self-identifications are volatile (Golash-Boza and Darity, 2008: 904).

Other studies have used exogenous indicators for skin color –color palettes– as benchmarks for different skin tones with the purpose of aiding interviewers classify

respondents more accurately (Dixon and Telles, 2017; Rejón Piña and Ma, 2023: 170). However, this approach is still prone to other distortions of skin color (Harris, 1964). For example, in the United States, men in professional attire are more likely to be categorized as “white”, while people of low socioeconomic levels are more likely to be categorized as “black” (Garcia and Abascal, 2016: 423).

Researchers have also used photo elicitation – presented image(s) to participants and asked them to classify the photos according to their skin tone– to measure skin color (Candelario, 2007; Roth, 2012; Sorokowski et al., 2013). However, research shows that all the elements of a portrait –not just skin tone– affect people’s perception (Hill, 2002). For instance, Garcia and Abascal (2016) demonstrated that skin color ratings were affected by the presence of a racially distinctive name.

Especially relevant to this article, is the phenomenon of “transcoloration” –where perceived color changes– that takes place when people classify themselves (or others) into skin tone categories. As mentioned before, when people “assess” skin color, their judgment is not purely based on complexion but it is influenced by other phenotypical and cultural elements (Garcia and Abascal, 2016; Harris, 1964; Hill, 2002; Omi and Winant, 2014: 2). In the literature, this is sometimes described as “money whitening”, because higher social status, wealth and connections usually lead to whiter classifications (Bonilla-Silva and Glover, 2004: 154; Golash-Boza, 2010; Roth et al., 2022). This transcoloration points to the relevance of other elements of the racial schema. That is, the interplay between the elements is such that the presence/absence of some, shape the way others are perceived, including skin color.

Recent research has started to study the place of skin color in racial schemas, particularly if complexion is in fact the best predictor for racial identification (Monk, 2016; Telles and Paschel, 2014). The most relevant example of this kind of study in Mexico is the Project on Ethnoracial Discrimination (PRODER). Employing data from a large sample survey (n = ~7000), researchers studied the interaction between different racial identifications (i.e., race, ethnic group) and skin color measurements with socioeconomic outcomes. They have found that –beyond skin color– the importance of other phenotypical traits in the racialization process has been underestimated (Krozer and Gómez, 2023; Solís and Güémez, 2020; Solís and Reyes Martínez, 2023). Some of their findings align with those of other studies; for instance, people are strongly sensible to even slight differences in hair texture, stature, the shape of the face, type of nose and gender (Bonilla-Silva, 2009; Bonilla-Silva and Glover, 2004; Feliciano, 2016; Solís et al., 2019, 2023, 2023). Furthermore, some of their results suggest that the effect of skin color is different to that of other features, sometimes just in intensity but others also in direction (Reyes-Martínez et al., 2023; Solís et al., 2023).

Overall, these findings reveal that the Mexican racial schema is more complex than what the race-skin tone “interchangeability” assumption suggests. What are its particular configurations and dynamics? This “racialization” process is still understudied –particularly in contexts where observers categorize others based solely on appearance– and we do not have a good understanding of which phenotypic characteristics are used to assign race (Feliciano, 2016: 391). Scholars have long pointed out that not all physical markers are racialized –at least not in the same way (Wade, 1993); and few studies have examined the role of phenotypic characteristics in shaping how individuals racially

classify others (Feliciano, 2016: 395). This is especially true in Latin America (Golash-Boza and Bonilla-Silva, 2013).

To address this question, I take a measurement (an automated algorithm) that isolates skin color from other social cues and compare it with one susceptible to other (admittedly not all) biases: human-coding. In the next section, I unpack some of the particularities of the Latin American and Mexican contexts, to then develop the hypotheses of the study.

Context

The categories of racial identification in Latin America are complex and porous (Bonilla-Silva and Glover, 2004: 151), to the point that some scholars have described the Latin American racial system as “fluid” (De Micheli, 2018: 53, 2021: 6; Irizarry et al., 2023; Marrow, 2003; Telles and Lim, 1998). In this fluid system, with the right combination of status and behavior, people can “escape” the stigma of certain racial identifications and “whiten” themselves (Degler, 1971; Schwartzman, 2007; Telles and Flores, 2013).

Historically, the ambition for whiteness was so entrenched that it fuelled the development of eugenics in many Latin American countries, including Mexico (Suárez y López Guazo, 2005). The goal to “whiten” was not only personal (Delgado et al., 2017; Hunter, 2011) but national, and it was embedded in public policy (Saldaña-Tejeda, 2022). However, by the 20th Century, this project to phenotypically whiten the population had failed in most Latin American nations; and a cultural whitening project arose in its place (Bonilla-Silva, 2009; Telles, 2014).

In countries like Mexico, progressive elites promoted the idea that society was color-blind and racially unified (Knight, 1990) through the nation-building narrative of *mestizaje*; according to which, miscegenation between the Spanish and the Indigenous had produced a racial group that embodied the best of both peoples: the *mestizo*; and all Mexicans belonged to this group (Martinez Casas et al., 2014; Vasconcelos, 1948).

Scholars argue that, because of this myth (and the public policies that followed), the notions of race and the practices of racism are often disconnected in Mexico; the former are removed from popular consciousness, but the latter persist (Friedlander, 2006; Moreno, 2008: 285). In Mexico, racism is not violent but “sells itself as attractive and appealing, like an obstacle anyone could overcome” (Iturriaga, 2015: 109). With the right combination of education, professional success and wealth, people can reach Mexican “whiteness” (Navarrete, 2016), which is not only a skin color but a form of cultural capital (Richards et al., 2023): physical, social, economic and cultural privilege (Vargas, 2015). The study of whiteness in Mexico is still developing (Ceron-Anaya et al., 2023), but several recent qualitative studies reveal that both phenotype and cultural habits are its constitutive elements (Cerón Anaya, 2023).

Moreover, the public discourse denying the existence of multiple ethno-racial communities along with social stigma, caused these groups to lose legal and political recognition and eventually become “invisible” (Iturralde, 2018). In fact, data on racial self-identification was not collected in Latin America until very recently (CEPAL, 2020).

Amidst these circumstances, and aware of the prominence of skin tone in the racial schema, scholars in the region resorted to the use of skin color as a proxy for race

(Rejón, 2024; Trejo and Altamirano, 2016). In fact, a burgeoning body of literature has found skin color-based discrimination in virtually all spheres of Mexican society –interpersonal relationships (Moreno, 2008, 2010; Sue, 2020), education (Rejón, 2023), the job market (Arceo-Gomez and Campos-Vazquez, 2014), and social mobility (Campos-Vazquez and Medina-Cortina, 2019).

However, as flagged before, recent studies have started to examine this question in more detail, paying attention not only to skin tone but to other elements in the racial schema. These works reveal that the importance of other phenotypical traits in the racialization process has been underestimated (Krozer and Gómez, 2023; Solís and Güémez, 2020; Solís and Reyes Martínez, 2023; Telles and Flores, 2013).

We now know that some *non-epidermic* treats have a transcoloration effect, but its extent and direction remain underexplored. In Latin America, some studies have found no evidence to support the idea that “money whitens” (Golash-Boza, 2010) or, even more, discovered more complex patterns in which money or status can “darken” (De Micheli, 2021; McNamee, 2020; Telles and Paschel, 2014). A few theories attempt to explain these phenomena (De Micheli, 2021; McNamee, 2020), but more empirical research is needed to inform the debate.

This paper aims to contribute such an examination of the racialization process, comparing two different skin color variables for the same subjects, which were gathered through very different methods. A full description of the dataset is presented in the following section. For now, consider that one of the skin color variables is the result of an automated algorithm that focuses exclusively on the color of the face area and the other was produced by independent human coders. These two measurements allow us to observe the impact that other factors (i.e., hair color and texture, eyes, lip shape, gender, jewelry, attire) have on skin color assessment and whether they significantly contribute to “transcoloration” (whitening or darkening) of the measurement. Call these other elements, *non-epidermic* factors.

The extent to which these *non-epidermic* factors influence people’s “skin color” assessment impacts the validity of machine-coded skin tone variables, especially when their effect causes the assessment to run in different directions.¹ If independent measures for skin tone (i.e., classification algorithms and colorimeters) are to be appropriate and helpful “replacements” for human-coding, they need to satisfactorily approximate human measurements, which comprise much more than a simple assessment of skin tone. If *non-epidermic* factors (which these technologies cannot detect yet) cause the variables to significantly differ from each other, then the machine-coded variables are less appropriate methods for anti-racist research. If, however, the two variables are substantially similar, the validity of these tools to approach the study of colorism and racism is reaffirmed.

Methods and data

Data

Analyzing the dynamics of the interaction among all the different elements of the Mexican racial schema requires to look at an existing data set from different angles or to collect a data set for this specific purpose. Here, I examine two different skin color variables produced for the same subjects. Both variables are processed from 3000 studio portraits of

legislators in the Mexican Chamber of Deputies. The portraits were taken against a white background with similar light, contrast and zoom conditions. These portraits, published in the official website of the Chamber, showcase various characteristics and markers that could affect skin color classification: complexion, hair color and texture, lip shape, and gender. To a lesser extent, the photos also display some class markers, such as clothing and jewelry. [Figure 1](#) presents a panel of sample photos for reference, but the full albums are available as supplementary material ([Rejon, 2025](#)).

Some might challenge the representativeness of the sample on the grounds that some studies suggest the pool of legislators are “whiter” than the average population ([Campos-Vazquez and Rivas-Herrera, 2021](#)), but other research argues the opposite ([Lawson et al., 2010](#); [Rejón, 2024](#)). At worst, the question of representativeness is empirical, and it remains unanswered; by no means does it affect the validity of the sample.

Admittedly, however, there are some limitations in using these portraits to conduct the skin tone measurements. First, photos “flatten” some facial features and do not include all elements of the racial schema (i.e., stature, accent, size). This is undeniable true. Like any other method in the social sciences, this one has disadvantages, including the fact that it cannot totally recreate the conditions of actual human interaction. Still, its advantages outweigh the limitations: it exposes participants to enough elements of the racial schema to allow a comparison with independent and automated measurements. Secondly, the skin tone of the face is susceptible to modifications (i.e., make up). While true, I find this limitation to be less relevant. In fact, the creators of classification algorithms for skin color have anticipated and responded to “the make-up objection”. They conceded that make-up could distort the measurements but argued that the same is true for other methods too (i.e., interviewers) and that, ultimately, if subjects usually present using make up, then measuring that –rather than actual skin tone– would better explain their experiences ([Rejón and Ma, 2023](#)). Afterall, skin color “as seen by others” is the important factor when examining racial inequality and discrimination” ([Telles 2012: 1167](#)).

Methods

The first skin tone measurement was obtained using the Classification Algorithm for Skin Color (CASCo), introduced by [Rejón and Ma \(2023\)](#). CASCo is a Python library that



Figure 1. Panel of the portraits in the sample.

detects the skin color of the face area of a given portrait and classifies it to one of the categories in the PERLA color palette (Telles, 2014). CASCo maps the face area on the HSV color space rather than RGB, which makes its measurements more accurate. In simpler terms, CASCo is a modern technological tool that identifies skin color alone, ignoring all other factors. It is an independent and automatic measure that focuses on skin color *exclusively*.

CASCo's measurements are innovative and valuable precisely because they overcome the shortcomings of other methods to measure skin color such as interviewers' bias and money whitening, as reviewed in previous sections of the paper. At the same time, however, this independent measure effectively isolates skin color from *non-epidermic* features –such as hair color and texture, eyes, lip shape, gender, jewelry, and attire– allowing us to compare its results with those gathered through human-based methods.

The second skin tone measurement was coded by humans, and conducted via an online (Qualtrics) survey that was distributed through social media by RacismoMX, an anti-racist organization in Mexico. This sampling process via activist organizations is common feature in academic research (Brinkerhoff, 2014; Brinkerhoff et al., 2019). The survey included one eligibility question: “In total, how many years have you lived in Mexico?” and respondents with less than 10 years of lived experience in Mexico were excluded from the sample. To eligible participants, the survey asked: “From the following color palette, select the skin color of the face of the person in the portrait”; then, it showed them one portrait per page along with the colors in the PERLA palette as options. Each participant was given 50 photos to classify (10 for each legislature in the sample). Asking explicitly about color (rather than “race”) and giving respondents the PERLA palette as options entails that all variation between their scores and CASCo's are due to non-epidermic factors.

In total, the survey had 367 responses but not all participants classified the 50 photos (i.e. they left the survey before completing it). Therefore, the first legislatures asked (the oldest ones) had more classifiers/coders than the recent ones. All classifications were preserved, to permit older photos (with less quality) to have a larger number of coders. Overall, each portrait was classified by 3-7 human coders. For each photo, the exact number of coders and their scores are available as supplementary material. The assessments of these human coders were averaged and rounded to the closest integer, to produce a categorical variable comparable with that of CASCo.

I anticipate some objections regarding the suitability of using the PERLA color palette, particularly in the Mexican context. The tool has been subjected to criticism on the grounds that there it presents significant “technical issues” that prevents it to appropriately capture the main skin tone variations of people in Mexico (COLMEX, 2023). However, while these criticisms might be substantiated, they have not been published or significantly outlined yet. For these reasons, I have not been able to corroborate them. In turn, the PERLA palette has been widely used in social research in different settings, including Mexico (Campos-Vazquez and Medina-Cortina, 2019; Rejón, 2023; Telles et al., 2015); therefore, employing it logically allows for comparative studies and facilitates the communication of the results. For these reasons, here I stick to the PERLA color palette.

In sum, I examined skin color in this sizeable sample from different angles: one that independently measures skin tone exclusively and one in which humans subjectively assess “skin color”. Comparing the results of these two methods gives us an indication of the extent to which *actual* skin color correlates to human “skin tone” classification, and sheds light on how much people are influenced by non-epidermic features when they assess “skin color”.

Now, there are several different statistical measures for assessing the extent of agreement among two or more raters. These computations are usually referred to as inter-rater reliability (IRR). I use several approaches to assess the degree of concordance between CASCo and human coders: percent agreement, Cohen’s κ , Spearman’s ρ , Krippendorff’s α , Lin’s Concordance Correlation Coefficient, Kendall’s W, Fleiss’ κ , and Gwet’s AC2. For the benefit of the reader unfamiliar with these computations, [Table 1](#) summarizes the characteristics of each one of them.

Most of these measurements vary from 0 to 1 (with some ranging from -1 to 1 to account for perfect disagreement), but the “accepted” levels of agreement are conventional and vary per discipline. For instance, clinical diagnoses usually require high kappa values but the social sciences accept lower rates.

In this paper, I follow the scale introduced by [Landis and Koch \(1977\)](#), widely used in social research. According to this scale, the strength of agreement can be interpreted as follows: values < 0 poor, 0.00–0.20 slight, 0.21–0.40 fair, 0.41–0.60 moderate, 0.61–0.80 substantial, and 0.81–1.00 almost perfect. Following [Krippendorff \(2004: 241\)](#), I am cautious when interpreting the findings: discarding data with agreement measures < 0.667 , drawing only tentative conclusions from data ≥ 0.667 but < 0.800 , and relying on data ≥ 0.800 .

All replication data and codes are included in a R Markdown file as supplementary material to this paper.

Findings and discussion

[Table 2](#) presents part of the inter-rater reliability computations. The level of agreements in this first set of IRR varies from slight to moderate. For instance, the percent agreement –where raters either agree or disagree, with no other factor considered– is 5.2% between CASCo and human coders, which indicate a low level of agreement.

The Cohen’s κ is 0.015, with a significant p -value of $1.43e^{-6}$; which suggests very limited agreement between CASCo and human coders. Given that PERLA can be considered an ordinal variable, it makes sense to use quadratic weights too ([Cohen, 1968](#); [Janusz et al., 2023](#)). Doing so, the κ is 0.245, which indicates that accounting for squared weights better captured the agreement structure in the data and increased the observed agreement. However, the strength of agreement is slight; that is to say, the two skin color measurements substantially differ from each other, which would indicate that *non-epidermic* features have a strong influence in the Mexican racial schema. The same is true for the Krippendorff’s α , which is 0.227, suggesting only a fair –relatively low– level of agreement among raters.

Table 1. Inter-rater reliability computations.

<i>IRR</i>	<i>Characteristics</i>
Percent agreement	Percent agreement is perhaps the simplest test available. It is the average amount of agreement expressed as a percentage.
Cohen's κ	Cohen's kappa (κ), which also measures absolute agreement and treat all disagreements equally. The advantage of this measure is that it factors in the role of chance when evaluating agreement. Conventionally, a κ lower than 0.61 is not considered substantial.
Spearman's ρ	Spearman's rank correlation coefficient, also known as Spearman's rho (ρ). This is a nonparametric measure of rank correlation (statistical dependence between the rankings of two variables) that assesses how well the relationship between two variables can be described using a monotonic function, varying from -1 to 1 . Spearman's ρ is appropriate for both continuous and discrete ordinal variables, such as the skin tone variables in this analysis.
Krippendorff's α	is Krippendorff's alpha (α), which is primarily used to measure the agreement achieved when coding a set of units of analysis. Krippendorff's α is applicable to any number of coders and categories, different types of variables (i.e., binary, nominal, ordinal, interval, ratio), and it adjusts itself to small sample sizes of the reliability data.
Lin's Concordance Correlation Coefficient	Lin's (1989) Concordance Correlation Coefficient (CCC), widely used in fields such as medical research, biology and psychology. It provides a single numerical value (-1 to 1), where: 1 indicates perfect positive agreement. -1 indicates perfect negative agreement and 0 indicates no agreement. Lin's CCC is better suited to quantify agreement between two continuous variables, a category both variables considered here could fit.
Kendall's W	Kendall's W is a measure of inter-rater reliability that accounts for agreement not only in Manichean terms, but for the strength of the relationship between multiple ratings. In other words, it measures the extent of (dis)agreement rather than only absolute agreement, differentiating between near misses versus not close at all. For these reasons, Kendall's W is very helpful with ordinal ratings, such as PERLA. It ranges from 0 to 1 , where higher values indicate stronger inter-rater reliability.
Fleiss' κ	Fleiss' kappa (κ) is a statistical measure for assessing the reliability of agreement between a fixed number of raters when assigning categorical ratings to a number of items. This contrasts with other kappas, which only work when assessing the agreement between not more than two raters. Furthermore, the measure calculates the degree of agreement in classification over that which would be expected by chance.
Gwet's AC2	Gwet's AC2 also compares the observed agreement rate with a comparator, it does not depend upon the assumption of independence between raters (Gwet, 2008).

Lin's Concordance Correlation Coefficient is 0.375, with confidence intervals (0.3520, 0.397) which, again, suggest a fair level of agreement between CASCo and human coders.

The only slightly different result amongst this first set of IRR, is the Spearman's ρ ; which, at 0.491, suggests a significant and moderate positive correlation between the two skin color variables. This result, however, should be interpreted with caution because ties in the data may affect the accuracy of the rank correlation coefficient.

In general, these set of IRR support the idea that *non-epidermic* features cause a significant diversion between the independent and subjective measurements of skin color. These outcomes would imply that other factors such as hair, eyes, or lips, are very salient in the Mexican racial schema, rendering measurements of actual skin color (i.e., skin tone classification algorithms or colorimeters) less relevant. However, more sophisticated measures are available to measure inter-rater reliability. In particular, some that are more appropriate to data with the characteristics of our skin color variables.

Table 3 shows the rest (and more sophisticated) of the measures for inter-rater reliability. Kendall's W (0.7143 and 0.7455) suggest a substantial degree of agreement among raters in their rankings. The corrected value takes into account the presence of ties, providing a more accurate estimate and a slightly higher estimate of agreement.

The Fleiss' κ value of 0.934 associated with ordinal weights, also suggests a high and statistically significant level of agreement among raters. When considering quadratic weights, κ is a bit higher (0.959) also indicating a high level of agreement among raters. Similarly, using ordinal weights, the Gwet's AC2 value of 0.727 with $p = 0$ suggests a statistically significant high level of agreement among raters beyond what would be

Table 2. Inter-rater reliability between CASCo and human coders.

Test	Estimate	Agreement
Percent Agreement	5.2	Slight
Cohen's κ	0.015	Slight
Cohen's κ (Quadratic)	0.245	Fair
Krippendorff's α	0.227	Fair
Lin's CCC	0.375	Fair
Spearman's ρ	0.491	Moderate

Table 3. Fleiss' Kappa and Gwet's AC2 between CASCo and human coders.

Test	p_a	p_e	coeff.	se	conf.int	p	A
Kendall's W	Uncorrected (ties)	0.7143					Substantial
	Corrected (ties)	0.7455					Substantial
Fleiss' κ	Ordinal	0.934	0.912	0.252	0.018	(0.217,0.287)	0 High
	Quadratic	0.959	0.943	0.270	0.018	(0.235, 0.306)	0 High
Gwet's AC2	Ordinal	0.934	0.758	0.727	0.006	(0.715,0.74)	0 High
	Quadratic	0.959	0.801	0.791	0.005	(0.781,0.801)	0 High

expected by chance. The observed agreement (pa) of 0.934 is much higher than the expected agreement by chance (pe) of 0.758. Using quadratic weights confirms the high level of agreement, with an AC2 value of 0.791, the observed agreement (0.959) is much higher than the expected agreement by chance (0.801).

In sum, while simplistic tests suggest a lack of agreement between CASCo and human coders, a more detailed analysis—one that takes into account the characteristics of the data—indicates that there is a high level of agreement between both variables. In other words, there is little inter-coder agreement about *exact* estimates but substantial—and statistically significant—agreement between the general direction of the two skin color variables. These findings directly engage with previous research in at least two ways: First, they confirm that non-epidermic features indeed alter people's perception of skin tone, even if slightly. Secondly, they validate the findings of studies using independent and automated measures of skin color, as they reveal that the elements of the racial schema that these tools cannot account for cause no major digression from independent skin color assessments. That is, they confirm the reliability of CASCo as an accurate substitute for human coding, at least in Mexico.

As reviewed in previous sections, measurements of skin color in the United States are subject to distortions triggered by somatic and class markers such as hair color and texture, lip shape, make-up style, jewelry and clothing. These markers impact people's perception of affluence and professional success which, in turn, makes subjects be observed to be whiter (Harris, 1964). Similarly, people perceived to be of low socioeconomic levels based on these markers, are more likely to be categorized as “black” (Garcia and Abascal, 2016: 423).

In Mexico, recent studies demonstrate that the importance of phenotypical traits—other than skin color—has been underestimated (Krozer and Gómez, 2023; Solís and Güémez, 2020; Solís and Reyes Martínez, 2023; Telles and Flores, 2013); revealing that people are sensible to even slight differences in hair texture, stature, the shape of the face, type of nose and gender (Solís et al., 2019, 2023, 2023). The magnitude of the effect of non-epidermic features is critical, as it has the potential to confirm or decimate the significance of studies examining the correlations between independent measures of skin color and variables such as electoral success, political representation or socioeconomic outcomes.

This empirical data suggests that in Mexico, subjectively perceived skin color (while not identical) is not very different to actual skin color. These findings are important because they confirm qualitative research arguing that skin tone is a good proxy for race in Mexico (Rejón, 2024; Saldívar, 2014: 90) and possibly Latin America (Telles et al., 2015). If assumptions and prejudices are attached to racial schemas (Sen and Wasow, 2016: 506; Wade, 2012) which are triggered by physical markers (Haslanger, 2012: 307), the fact that independent and subjective measures of skin color are not significantly different in Mexico indicates that independently measuring skin tone is, in fact, a legitimate approach to study racism.

Conclusion

Racism continues to be pervasive around the world. People still deploy notions of race to classify and assign value to others. They do so employing racial schemas to assign social

meaning to certain physical markers. A growing body of literature is devoted to examining the configuration and dynamics of these racial schemas. In Mexico, these investigations reveal that the importance of non-epidermic traits has been underestimated. Given that studies employing independent measures for skin tone (i.e., classification algorithms and colorimeters) are increasingly frequent, the magnitude of this effect of non-epidermic features is critical; it has the potential to challenge the appropriateness of these tools and, consequently, debunk the significance of said research corpus.

Using a novel dataset of 3000 portraits, here I compared two different measures of skin color. One was processed with an automated and independent algorithm that focuses exclusively on the color of the face area of the photo. The other one is a measure of “impressionistic” race, produced by human coders. In other words, the first variable effectively isolates skin color from *non-epidermic* features—such as hair color and texture, eyes, lip shape, gender, jewelry, and attire—while the human-coded variable is influenced by all those elements. This methodological approach isolates skin color from other social cues (via an automated algorithm) and compares it with a measure susceptible to other biases (human-coding), which allows us to observe the extent to which these non-epidermic features cause people’s perceptions to digress from actual skin color.

The findings reveal little inter-coder agreement about exact estimates but substantial—and statistically significant—agreement between the general direction of the two skin color variables; which suggests that while non-epidermic features slightly alter people’s perception of skin tone they do not cause major digression from independent skin color assessments.

These findings have significant implications for anti-racist research in Mexico. If, as this empirical data suggest, subjectively perceived and actual skin color are not very different, then quantitative research employing independent skin tone measures (i.e., classification algorithms) are likely to find valid correlations. Given that these methods usually facilitate research that is otherwise not feasible due to limitations on human or financial resources, the findings of this article are encouraging as they entail that these accessible research tools can indeed contribute valuable scholarship.

At the same time, these results confirm that differences between the measurements do exist. This entails that, while these technology-based independent skin color measures are good approximations to human-coding, they should be accompanied by human-coded variables whenever possible. The combined use of these two types of variables would buttress the analyses and account for any divergence with human appreciations. In any case, these findings confirm the benefits of these tools, including the fact that they showcase how human codification diverges from pigment tracing.

Note that these findings say nothing about the influence of *non-epidermic* features in racial classification and self-identification. Given that this dataset does not include a variable of the legislators’ racial classification (self-ascribed or otherwise), it should not inform studies which find interdependence between these features and racial classification. In that sense, it is still the case that skin color is only one of the many markers in the phenomenon of race (Monk, 2021: 80). However, we now know a little bit more about the Mexican racial schema.

Another interesting potential implication of these findings has to do with the way in which “racial” discrimination is framed in the Mexican case. The findings suggest that

skin color assessment is not significantly distorted by other *non-epidermic* features, so one might argue that colorism is the best way to describe the skin color discrimination that recent research has unveiled. I think, however, that such a judgement would depend on each of the relevant studies. Afterall, the data I analyzed here comprehends *non-epidermic* features, much of which are still phenotypical (i.e., hair, eyes, lips) even if gender, jewelry, and attire are also marginally included. The data does not include cultural traits or norms, such as the language one speaks, one's accent, or one's demeanor. These elements might as well prove relevant in the Mexican racial schema and in the racialization process through which some people are discriminated against. Therefore, I conclude this article calling for more empirical research on the topic.

Declaration of conflicting interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

René Rejón  <https://orcid.org/0000-0001-5516-3278>

Data availability statement

The data that support the findings of this study are openly available at <https://doi.org/10.26188/24873330> (Rejon, 2025).

Note

1. In some cases, non-epidermic features will influence human “skin tone” assessment to align with machine-coded variables which focus solely on skin color. For example, a human coder would be more likely to “correctly” (according to the algorithmic measure) identify pigment level of someone with dark skin and curly hair, than if that person's hair was straight. In this case, hair texture matters, but it would not cause divergence between the two measures. I thank the anonymous reviewer which pointed out this possibility.

References

- Arceo-Gomez EO and Campos-Vazquez R (2014) Race and marriage in the labor market: a discrimination correspondence study in a developing country. *American Economic Review* 104(5): 376–380.
- Atkin A (2012) *The Philosophy of Race*. New York: Acumen Publishing.
- Banton M (2012) The colour line and the colour scale in the twentieth century. *Ethnic and Racial Studies* 35(7): 1109–1131.

- Bonilla-Silva E (2009) Are the Americas ‘sick with racism’ or is it a problem at the poles? A reply to Christina A. Sue. *Ethnic and Racial Studies* 32(6): 1071–1082.
- Bonilla-Silva E (2015) The structure of racism in color-blind, “post-racial” America. *American Behavioral Scientist* 59(11): 1358–1376.
- Bonilla-Silva E and Glover K (2004) “We are all Americans”: the Latin Americanization of race relations in the United States. In: Krysan M and Lewis A (eds). *The Changing Terrain of Race and Ethnicity*. New York: Russell Sage Foundation.
- Brinkerhoff JM (2014) Diaspora philanthropy: lessons from a demographic analysis of the Coptic diaspora. *Nonprofit and Voluntary Sector Quarterly* 43(6): 969–992.
- Brinkerhoff JM, McGinnis Johnson J and Gudelis D (2019) Are our assumptions about diaspora and immigrant philanthropy generalizable? Exploring the relevance to high-income countries of origin. *Nonprofit and Voluntary Sector Quarterly* 48(5): 1094–1109.
- Campbell ME and Troyer L (2007) The implications of racial misclassification by observers. *American Sociological Review* 72: 750–765.
- Campos-Vazquez R and Medina-Cortina EM (2019) Skin color and social mobility: evidence from Mexico. *Demography* 56(1): 321–343.
- Campos-Vazquez R and Rivas-Herrera C (2021) The color of electoral success: estimating the effect of skin tone on winning elections in Mexico. *Social Science Quarterly* 102(2): 844–864.
- Candelario G (2007) *Black behind the Ears: Dominican Racial Identity from Museums to Beauty Shops*. Durham: Duke University Press.
- CEPAL (2020) Aspectos conceptuales de los censos de población y vivienda: desafíos para la definición de contenidos incluyentes en la ronda 2020. In: *Serie Seminarios Y Conferencias 94*. Santiago. Available at: <https://www.cepal.org/apps>
- Cerón Anaya H (2023) Color de piel humilde, color de piel privilegiado. Elites y Blancura en América Latina. *Nueva Sociedad* (303): 50–63.
- Ceron-Anaya H, de Santana Pinho P and Ramos-Zayas A (2023) A conceptual roadmap for the study of whiteness in Latin America. *Latin American and Caribbean Ethnic Studies* 18(2): 177–199.
- Chandra K (2006) What is ethnic identity and does it matter? *Annual Review of Political Science* 9(1): 397–424.
- Clealand D (2017) *The Power of Race in Cuba*. Oxford: Oxford University Press. Available at: <https://academic.oup.com/book/5537>
- Cohen J (1968) Weighted kappa: nominal scale agreement provision for scaled disagreement or partial credit. *Psychological Bulletin* 70(4): 213–220.
- COLMEX (2023) Presentación de encuesta del proyecto sobre discriminación étnico-racial en México (PRODER). Available at: https://www.youtube.com/live/YSjV_FaOrRY (accessed 29 November 2024).
- De Micheli D (2018) The racialized effects of social programs in Brazil. *Latin American Politics and Society* 60(1): 52–75.
- De Micheli D (2021) Racial reclassification and political identity formation. *World Politics* 73(1): 1–51.
- Degler C (1971) *Neither Black Nor White: Slavery and Race Relations in Brazil & the United States*. New York: The Macmillan Company.

- Del Popolo F and Schkolnik S (2012) Indigenous Peoples and Afro-descendants: the difficult art of counting. In: Angosto Ferrández LF and Kradolfer S (eds) *Everlasting Countdowns: Race, Ethnicity and National Censuses in Latin American States*. Cambridge: Cambridge Scholars Publishing, 304–334.
- Del Popolo F, Oyarce AM, Schkolnik S, et al. (2009) *Censos 2010 y la inclusión del enfoque étnico : hacia una construcción participativa con pueblos indígenas y afrodescendientes de América Latina ; Serie Seminarios y conferencias 57, September*. Santiago: Naciones Unidas. CEPAL.
- Delgado AN, Deister VG and Beltrán CL (2017) ¿De qué me ves cara? Narrativas de herencia, genética e identidad inscritas en la apariencia. *AIBR Revista de Antropología Iberoamericana* 12(3): 313–337.
- Dixon A and Telles E (2017) Skin color and colorism: global research, concepts, and measurement. *Annual Review of Sociology* 43: 405–424.
- Feliciano C (2016) Shades of race. *American Behavioral Scientist* 60(4): 390–419.
- Friedlander J (2006) Being Indian revisited. In: *Being Indian in Hueyapan*. New York: Palgrave Macmillan US, 181–223. Available at: https://link.springer.com/10.1057/9780230601659_10
- García D and Abascal M (2016) Colored perceptions: racially distinctive names and assessments of skin color. *American Behavioral Scientist* 60(4): 420–441.
- Golash-Boza T (2010) Does whitening happen? Distinguishing between race and color labels in an African-descended community in Peru. *Social Problems* 57(1): 138–156.
- Golash-Boza T and Bonilla-Silva E (2013) Rethinking race, racism, identity and ideology in Latin America. *Ethnic and Racial Studies* 36(10): 1485–1489.
- Golash-Boza T and Darity W (2008) Latino racial choices: the effects of skin colour and discrimination on Latinos' and Latinas' racial self-identifications. *Ethnic and Racial Studies* 31(5): 899–934.
- Gwet KL (2008) Computing inter-rater reliability and its variance in the presence of high agreement. *British Journal of Mathematical and Statistical Psychology* 61(1): 29–48.
- Harris M (1964) Racial identity in Brazil. *Luso-Brazilian Review* 1: 21–28.
- Harris M, Gomes Consorte J, Lang J, et al. (1993) Who are the whites? imposed census categories and the racial demography of Brazil. *Social Forces*, Available at: <https://www.jstor.org/stable/2579856>
- Haslanger S (2012) *Resisting Reality: Social Construction and Social Critique*. Oxford: Oxford University Press.
- Hill ME (2002) Skin color and the perception of attractiveness among African Americans: does gender make a difference? *Social Psychology Quarterly* 65(1): 77.
- Htun M (2015) *Inclusion without Representation in Latin America*. Cambridge: Cambridge University Press.
- Hunter M (2011) Buying racial capital: skin-bleaching and cosmetic surgery in a globalized world. *Journal of Pan African Studies* 4(4): 142–165.
- Irizarry Y, Monk EP and Cobb RJ (2023) Race-shifting in the United States: Latinxs, skin tone, and ethnoracial alignments. *Sociology of Race and Ethnicity* 9(1): 37–55.
- Iturralde G (2018) Invisibilidad. Las personas afrodescendientes y el racismo. In: Nieto GI (ed) *Caja de Herramientas Para Identificar El Racismo En México*. Mexico City: Afrodescendencias en México Investigación e Incidencia, A.C.

- Iturriaga E (2015) La ciudad blanca de noche: las discotecas como espacios de segregación. *Alteridades* 25(50): 105–115.
- Janusz A, Jarrin A, Faerman T, et al. (2023) Electoral colorism: candidate skin color and list placement in Ecuador. *Electoral Studies* 81: 102555.
- Keith VM and Herring C (1991) Skin tone and stratification in the black community. *American Journal of Sociology* 97(3): 760–778.
- Knight A (1990) The idea of race in Latin America, 1870–1940. In: Graham R (ed.) *Racism, Revolution, and Indigenismo: Mexico, 1910–1940*. Austin: University of Texas Press, pp. 71–114.
- Krippendorff K (2004) *Content Analysis: An Introduction to its Methodology*. 2nd edition. Thousand Oaks, CA: Sage.
- Krozer A and Gómez A (2023) Not in the eye of the beholder: racialization, whiteness, and beauty standards in Mexico. *Latin American Research Review* 58(2): 422–439.
- Lancaster RN (1991) Skin color, race, and racism in Nicaragua. *Ethnology* 30(4): 339–353.
- Landis JR and Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* 33(1): 159.
- Lawson C, Lenz GS, Baker A, et al. (2010) Looking like a winner: candidate appearance and electoral success in new democracies. *World Politics* 62(4): 561–593.
- Lin LI-K (1989) A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 45(1): 255.
- Loveman M (1997) Is ‘race’ essential? *American Sociological Review* 64(6): 891–898.
- Mallon R (2006) ‘Race’: normative, not metaphysical or semantic. *Ethics* 116(3): 525–551.
- Marrow H (2003) To be or not to be (Hispanic or Latino). *Ethnicities* 3(4): 427–464.
- Martinez Casas R, Saldivar E, Flores R, et al. (2014) The different faces of mestizaje: ethnicity and race in Mexico. In: Telles E (ed) *Pigmentoocracies: Ethnicity, Race, and Color in Latin America*. Chapel Hill: University of North Carolina Press, 36–80.
- McNamee L (2020) Colonial legacies and comparative racial identification in the Americas. *American Journal of Sociology* 126(2): 318–353.
- Mitchell-Walthour G (2017) *The Politics of Blackness*. Cambridge: Cambridge University Press. Available at: <https://www.cambridge.org/core/product/identifier/9781316888742/type/book>
- Monk EP Jr (2015) The Cost of Color: Skin Color, Discrimination, and Health among African-Americans. *American Journal of Sociology* 121(2): 396–444. doi: 10.1086/682162.
- Monk EP (2016) The consequences of “race and color” in Brazil. *Social Problems* 63(3): 413–430.
- Monk EP (2021) The unceasing significance of colorism: skin tone stratification in the United States. *Daedalus* 150(2): 76–90.
- Moraes Silva G, Toste Daflon V and Giraut C (2024) Seeing race like a state: higher education affirmative action verification commissions in Brazil. *Latin American Politics and Society* 66(1): 1–26.
- Moreno M (2008) Historically rooted transnationalism: slightedness and the experience of racism in Mexican families. *Journal of Intercultural Studies* 29(3): 283–297.
- Moreno M (2010) Distributed intensities: whiteness, mestizaje and the logics of Mexican racism. *Ethnicities* 10(3): 387–401.
- Navarrete F (2016) *México Racista. Una Denuncia*. Mexico City: Penguin Random House.

- Nogueira O (1955) Preconceito racial de marca e preconceito racial de origem: sugestão de um quadro de referência para a interpretação do material sobre relações raciais no Brasil. *Revista Anhembi* 60: 554–557.
- Nutini H (1997) Class and ethnicity in Mexico: somatic and racial considerations. *Ethnology* 36(3): 227–238.
- Omi M and Winant H (2014) *Racial Formation in the United States*. 3rd edition. New York, NY: Routledge.
- Perreira KM and Telles E (2014) The color of health: skin color, ethnoracial classification, and discrimination in the health of Latin Americans. *Social Science and Medicine* 116: 241–250.
- Rejón R (2023) Avoiding backlash: narratives and strategies for anti-racist activism in Mexico. *Ethnicities* 23(6): 843–866.
- Rejón R (2024) The colours of lawmaking: ethno-racial descriptive representation in Mexico. *Ethnic and Racial Studies* 47(4): 785–808.
- Rejon R (2025) *Automated transcoloration? Skin tone classification algorithms and the prominence of non-epidermic features in the Mexican racial schema*. The University of Melbourne. Dataset. DOI: [10.26188/24873330.v1](https://doi.org/10.26188/24873330.v1).
- Rejón Piña RA and Ma C (2023) Classification Algorithm for Skin Color (CASCo): a new tool to measure skin color in social science research. *Social Science Quarterly* 104(2): 168–179.
- Reyes-Martínez J, Alcántara I and Solís P (2023) *Ethnoracial Disparities in Self-Rated Health: Exploring the Impact of Skin Color and other Ethnoracial Characteristics in Mexico*. Mexico City: Proyecto sobre Discriminación Étnico-Racial en México (PRODER) 9.
- Richards BN, Ceron-Anaya H, Dumais SA, et al. (2023) What’s race got to do with it? Disrupting whiteness in cultural capital research. *Sociology of Race and Ethnicity* 9(3): 279–294.
- Roth W (2012) *Race Migrations: Latinos and the Cultural Transformation of Race*. Stanford, CA: Stanford University Press.
- Roth W, Solís P and Sue C (2022) Beyond money whitening: racialized hierarchies and socio-economic escalators in Mexico. *American Sociological Review* 87(5): 827–859.
- Saldaña-Tejeda A (2022) Genómica y eugenesia: nuevos imaginarios reproductivos y viejos racismos de estado. *Estudios Sociológicos de El Colegio de México* 40: 151–178.
- Saldívar E (2014) ‘It’s not race, it’s culture’: untangling racial politics in Mexico. *Latin American and Caribbean Ethnic Studies* 9(1): 89–108.
- Schwartzman LF (2007) Does money whiten? Intergenerational changes in racial classification in Brazil. *American Sociological Review* 72(6): 940–963.
- Sen M and Wasow O (2016) Race as a bundle of sticks: designs that estimate effects of seemingly immutable characteristics. *Annual Review of Political Science* 19: 499–522. doi: [10.1146/annurev-polisci-032015-010015](https://doi.org/10.1146/annurev-polisci-032015-010015).
- Solís P and Güémez B (2020) Características étnico-raciales y desigualdad de oportunidades económicas en México. *Estudios Demográficos y Urbanos* 36(1): 255–289.
- Solís P and Reyes Martínez J (2023) Discriminación percibida, características etnoraciales y género. In: Rodríguez Zepeda J and González Luna T (eds) *Interseccionalidad. Teoría Antidiscriminatoria y Análisis de Casos*. Mexico City: UAM, 139–178.
- Solís P, Krozer A, Arroyo BC, et al. (2019) Discriminación étnico-racial en México: una taxonomía de las prácticas. In: Rodríguez Zepeda J and González Luna Corvera T (eds) *La Métrica de Lo*

- Intangible: Del Concepto a La Medición de La Discriminación*. Mexico City: CONAPRED, 55–94.
- Solis P, Güémez B and Campos-Vázquez RM (2023) Skin tone and inequality of socioeconomic outcomes in Mexico: a comparative analysis using optical colorimeters and color palettes. *Sociology of Race and Ethnicity*, Epub ahead of print. DOI: [10.1177/23326492231217232](https://doi.org/10.1177/23326492231217232).
- Sorokowski P, Sorokowska A and Kras D (2013) Face color and sexual attractiveness: preferences of yali people of Papua. *Cross-Cultural Research* 47(4): 415–427.
- Suárez y López Guazo L (2005) *Eugenesis Y Racismo En México*. Mexico City: UNAM.
- Sue C (2013) *Land of the Cosmic Race: Race, Mixture, Racism and Blackness in Mexico*. Oxford: Oxford University Press.
- Sue C (2020) The dynamics of color. In: Nakano Glenn E (ed) *Shades of Difference*. Stanford: Stanford University Press, 114–128. Available at: <https://www.degruyter.com/document/doi/10.1515/9780804770996-009/html>
- Telles E (2004) *Race in Another America: The Significance of Skin Color in Brazil*. Princeton: Princeton University Press.
- Telles E (2012) The overlapping concepts of race and colour in Latin America. *Ethnic and Racial Studies* 35(7): 1163–1168.
- Telles E (ed) (2014) *Pigmentocracies: Ethnicity, Race, and Color in Latin America*. Chapel Hill: University of North Carolina Press.
- Telles E and Flores R (2013) Not just color: whiteness, nation, and status in Latin America. *Hispanic American Historical Review* 93(3): 411–449.
- Telles E and Lim N (1998) Does it matter who answers the race question? Racial classification and income inequality in Brazil. *Demography* 35(4): 465–474.
- Telles E and Paschel T (2014) Who is black, white, or mixed race? How skin color, status, and nation shape racial classification in Latin America. *American Journal of Sociology* 120(3): 864–907.
- Telles E, Flores RD and Urrea-Giraldo F (2015) Pigmentocracies: educational inequality, skin color and census ethnoracial identification in eight Latin American countries. *Research in Social Stratification and Mobility* 40: 39–58.
- Trejo G and Altamirano M (2016) The Mexican color hierarchy. In: Hooker J and Tillery ABJ (eds) *The Double Bind: The Politics of Racial and Class Inequalities in the Americas*. Washington, DC: American Political Science Association, 1–14.
- Vargas S (2015) México : la pigmentocracia perfecta. Available at: <https://horizontal.mx/mexico-la-pigmentocracia-perfecta/> (accessed 11 June 2019).
- Vasconcelos J (1948) *La Raza Cósmica*. BuenosAires: Espasa-Calpe.
- Vaughn B (2013) México Negro: from the shadows of nationalist mestizaje to new possibilities in Afro-Mexican identity. *The Journal of Pan African Studies* 6(1): 227–240.
- Wade P (1993) *Blackness and Race Mixture: The Dynamics of Racial Identity in Colombia*. Baltimore, MD: John Hopkins University Press.
- Wade P (2012) Skin colour and race as analytic concepts. *Ethnic and Racial Studies* 35(7): 1169–1173.