RGB Data Analysis from Massey Martin Scale for Quantitative Imaging and Measurement Standardization

Dr. Klaus Reinhardt[1]*, Dr. Anna Fischer[1], Dr. Markus Weber[2], Dr. Lena Hoffmann[2]

[1]University of Heidelberg, Department of Medical Imaging and Digital Measurement Systems, Heidelberg, Germany

Abstract

Current events have placed the discussion of racism in the police force at the forefront. Until now, there has not been a systematic and accurate way to analyze racism. In the past, either individual judgment or demographically imperfect bifurcation was used. In prior research, either observational guesses, demographic data, or the Massey-Martin Scale of Skin Color Darkness were used to assess race. With RGB colorimeter or eyedropper technology, a total picture of racism can be seen through the entire spectrum of skin tone. Nuances to racism were challenging to flesh out with prior methods. Without exact, consistent, and unbiased data, no research can be trusted. In the past, research on racism has not used exact, consistent, and unbiased data.

Keywords: Massey-Martin Scale, racism, RGB, colorimeter, bimodal means, skin tone discrimination, colorism, chromaticity diagram, PANTONE, PERLA

Evidence-Based Research for Policing

Operational orders and directives can be shaped by both public opinion and evidence of disproportionate minority contact. The data cannot be accurate if there is no systematic way of describing minorities. Racism is not based on demographics. Racism is based on the perception of skin color. Large discrepancies exist when research is judging race based on skin color. The Massey-Martin Scale of Skin Color Darkness is a systematic way of describing race, but it also suffers from a reliance on an individual's inaccurate judgment. With the increased need to answer questions of racism in the police force (Krieger et al., 2015), a much more accurate and systematic process for analysis is needed (Meissner & Brigham, 2001). Equally important is the ability to define fully and contextually explain neighborhood conflict. A poignant example being that under the contested boundaries hypothesis, it requires a precise and bias-free measure of skin tone (Legewie & Schaeffer, 2016).

Evidence of Use of Force on Skin Tone

A leading cause of death among young men is from the police. Black men, compared to White men, have an exceptionally high risk of death by Police Officer, or death by cop (DCP). Because death by Police is a leading cause of death in young men, it makes it more than a public health issue, but a racial inequality factor as well (Edwards et al., 2019). No matter what it is called, police homicides, police-related deaths, or legal intervention deaths (Finch et al., 2019), the number of those that die by Police is not small. The disproportionate amount of Black men that die by Police makes the analysis of skin tone imperative in our understanding of not only the question as to why but the answer of how to fix it. The bifurcation process of demographics will not answer these questions or find solutions.

Although there is no governmental requirement in the United States to track those killed by the police (Krieger et al., 2015), initial research suggests that the risk of DCP is going down, but future research is needed to determine if this is correct. United States government data is unreliable, inconsistent (Finch et al., 2019) and underreported (Currie et al., 2015). The data on DCP rely on Fatal Encounters (FE), which is documented by public records and journalists. The FE data used photographs to classify deaths by race, and admittedly, can not take in variables of skin tone, of which the authors suggest should be closely considered in the future (Edwards et al., 2019). Without useful data, the extent of the issue is unknown (Alang et al., 2017).

Black men face approximately a 1 in 1,000 chance of being killed by the police, and White men face an approximately 1 in 2,500 chance (Edwards et al., 2019). Neighborhood resources and race conflict stymie scholars, policymakers, and pundits as a need for exact and specific measures of fuzzy boundaries and polarized borders need concise measures (Legewie & Schaeffer, 2016).

Skin tone is a vital point to be considered in providing police officers with evidence-based research. It will help them in understanding disproportionate minority contact and the ramifications of racism in their contact with minorities. Discrimination leads to anger, and anger leads to adverse outcomes with police contact. The darker the skin tone, the more discrimination is someone will encounter, the more someone will then internalize oppression, the more someone will be angry, and that ultimately leads to more negative outcomes with police contact (Krieger et al., 1998).

Black men report more discrimination than Black women. Black men experienced discrimination across the board in the educational system (Campos-Castillo, 2019), but only darker-skinned Black men reported racism from the police if they are in the working-class.

Darker-skinned racism was reported less for Black professional men regarding police contact (Krieger et al., 1998). The implications are thought-provoking, but the concluding remarks of the researchers were that there was insufficient data to provide a definitive answer to the effects of racism on emotional health. With a Panchromatic Vicinage Chart (PVC), new insights could emerge from this research.

Skin Tone Measures, Colorism vs. Racism

Color has long been a way to identify people. Demographically separating people is not enough. Looking at the spectrum of skin color is becoming more important for looking at inequality and combatting disparities more accurately (Campos-Castillo, 2019). Color has been used in police work as long as there has been police work. The use of original points and evolved points in a bimodal histogram of a fingerprint has been used to prove identification accurately (Tobias & Seara, 2002) for over 100 years (Uenuma, 2018). Skin tone data is rarely collected, although the Equal Employment Opportunity Commission has seen a significant increase in discrimination based on skin tone. Increasingly popular skin tone scales are fraught with bias (Hill, 2002). White coders are three times more likely to code someone darker than a Black coder would (Lance Hannon et al., 2014). In a study of 39 research articles with 91 independent samples of nearly 5,000 participants, it was shown that there is an own-race bias phenomenon (Meissner & Brigham, 2001), also known as the cross-race effect (Bornstein et al., 2013). Besides ease, there is no reason to categorize people by race in evidence-based police decision making. Skin color is a first-order abstraction, while race is a second-order abstraction. Skin tone is measurable and can be visually verified. Visually, race can only be guessed at based on skin tone and other stereotypical characteristics (Banton, 2012). Until a scientific way of removing all bias from data collection, it will remain challenging for the Equal Employment Opportunity

Commission to work successfully. Using eyedropper technology, exact and repeatable data about skin tone can be extracted.

Massey-Martin Scale of Skin Color Darkness

The Massey-Martin Scale is considered one of the most important skin tone measurement instruments (Lance Hannon & Defina, 2016). The scale, as seen in Figure 1, has widespread acceptance, but the main problem with using the Massey-Martin is that it suffers from low intercoder reliability. It is also only on a scale of 1 (very light skin tone) to10 (very dark skin tone) and does not measure the subtle difference of skin tone. The scale an evenly scaled, as can be seen in Figure 2, so with few data points, it can not distinguish differences well across concentrated skin tones. As can be seen by comparing Figure 2 and Figure 3, the majority of darker skin tones are unlikely to be distinguishable by a human coder. In addition, the considerable reduction in the color blue in the number 9 hand, makes it a lighter skin tone overall.

Some researchers use the chart as an 11-point scale adding 0 to represent albinism (Hersch, 2008), but there are much larger consistency issues in using this method. In a study of trustworthiness, the researchers classified five and below as light skin, and classified six and above as dark skin (Birdsong, 2017). Just a cursory glance at Figure 1 reveals the problem with arbitrarily assigning this value. It would appear that this is a good indication of separation, but when looking at the amount of the sample falling between four and five and five and six, the problem becomes evident. There is too much of the population that falls between the Massey-Martin Scale from four to six. Further research would be warranted to determine how this point changes between communities and regions.

The Massey-Martin Skin Color Scale is considered the standard for coding skin tone, but there is a lack of consistency even beyond what is created by human error or judgment. Most research only takes one measurement of the data point, and research is not consistent across raters (Lance Hannon & Defina, 2016). This inconsistency should be a cause for great concern since it has been shown that interviewer effects can materially change survey data and results (Davis et al., 2010). The guide is used as a comparative judgment made by others about skin color and how it coordinates with the 10 point scale (Campos-Castillo, 2019). Some researchers will count 1, 2, and 3 as light skin tone while others will only count 1 and 2 on the scale as light skin tone (Louie & Wilkes, 2018). On a ten-point scale, that can lead to material differences and change findings. If there existed a standard in research, it would still not be unusual to see a 10 percent variation between coders (Louie & Wilkes, 2018).

As the world becomes more ethnically diverse over time, the ability to look at skin tone over race is warranted (L Hannon & DeFina, 2013). As skin tone curves flatten out, the nuances of pigmentocracy need more exact measures to understand the prevalence of color relating to wealth and social status (Christensen, 2016). Many new techniques can also help those at high risk. The darker women are, the more likely it leads to low self-esteem, which means more likely to be involved in domestic violence, which leads to a higher risk for high-risk sexual behavior, which can lead to Human Immunodeficiency Virus (HIV), of which Black women in the south represent 71 percent of the new infections.

Although the Massey-Martin Scale proved to be the leading way of studying skin tone, another more scientific method was highlighted in an earlier study. In this study, they took light values off an area of the upper arm not typically exposed to the sun. They used a Photovolt 577 reflectance meter to take the reading using amber, blue and green filters. The research used the

amber filter, although there appeared to be no statistical reason to say one was better than the other (Krieger et al., 1998). The Massey-Martin Scale of Skin Color Darkness was a good start, but the PERLA Color Palette is easier to implement and also utilizes a 0-10 scale (Holland, 2020). The PERLA Color Palette can be seen in Figure 4. The Project on Ethnicity and Race in Latin America (PERLA) has shown that darker skin color correlates to lower education and socioeconomic status (Horta, 2012). As part of a survey that included Acculturation Rating Scale for Mexican Americans-II (ARSMA-II), Multigroup Ethnic Identity Measure-Revised (MEIM-R), University Environment Scale (UES), Cultural Congruity Scale (CCS), Psychological Well-Being Short Scale (PWBSS), and Satisfaction with Life Scale (SWLS) looked to exam the effects of darker skin tones. Using one-way MANOVA, the study showed that darker-skinned students had substantially lower grade point averages (Lamas, 2018). Darker-skinned people have lower cognitive ability, lower crystallized intelligence, and lower intelligence quotient (Fuerst et al., 2019).

PANTONE SkinTone™ Guide

In the same venue as the PERLA Color Palette is the PANTONE SkinTone™ Guide, except it consists of 110 colors that try to duplicate all possible human skin tones (PANTONE, 2020). In many ways, the PANTONE system, which can be seen in Figure 5, is perfect. Its main drawback is the time it would take for someone to use 110 colors in a fandeck to assess someone's color. Although the results could be consistent and accurate, the time commitment would make this method impossible to implement with a reasonably large sample. The use of electronic means would make this process jump that hurdle. The company does indeed have electronic solutions, but this turns into an expensive rabbit hole when an actual perfect alternative solution is already available for use now.

RGB Eyedropper

The best way to collect evidence-based data on discrimination and racism is through skin tone and not by judgmental categories or lineage. Racism lies in the judgment of skin tone. There is currently no accurate way to categorize skin tone (Villarreal, 2010). Utilizing an RGB-eyedropper now makes this statement untrue. Analyzing skin tone with an RGB eyedropper is accurate, consistent, and unbiased.

Often researchers will code observations based on perceived skin tone, hair: texture, eye color, and facial features (Louie & Wilkes, 2018). Observing data in this manner is half the definition of racism itself. A systematic and consistent way of analyzing skin tone is with a colorimeter or an RGB eyedropper. With an RGB eyedropper or the colorimeter, an exact measurement can be made of a photo, a website, or an actual person. These products are incredibly accurate and providing scientific data as opposed to human interpretation. RGB is the best method to distinguish the skin tone of different races (Kim et al., 2012).

Table 1 gives an example as to the accuracy of the Massey-Martin Scale of Skin Color Darkness. With this experiment, a Google Chrome eyedropper plug-in was used to take repeated readings of the hands in the Massey-Martin Scale. As can be seen in Table 1, the readings are more than significant to produce reliable and trustworthy research. The eyedropper discovered issues with the scale in reference to being nonlinear, and an alarming number 9 image consistency.

Future Implications

Using a normalized RGB chromaticity diagram, not only could more accurate skin tone be recorded, but identity verification advances as well (Soetedjo et al., 2010). Using RGB chromaticity diagram lips of a face could increase the discrimination of lips from the skin. For

police, this would mean the ability to read facial expressions and understand speech from far away or under challenging conditions using the Automatic Visual Speech Recognition System. This technology could possibly be used as a field lie detector test or warning of increasing aggression.

Facial recognition is a difficult task and requires complex processes. RGB data and a depth map, it is simple to make a 3D image of a person (Goswami et al., 2013). Knowing what or where people are looking can provide valuable clues into the future and past behavior of someone coming into contact with a police officer. Suppose someone is looking at an officer's gun during a field interview, better facial tracking could provide possible life-saving information. If someone continues to look at an exact spot during questioning, it may imply possible evidence or flight path.

Light-skinned people are disproportionately thought of as being more trustworthy and innocent (Birdsong, 2017). Being able to read the person being interviewed, and the police officer, could provide valuable feedback as to if the officer is effectively reading the trustworthiness and innocence of the person. This is especially important in a judicial scenario when sentencing could be adjusted to eliminate racial bias. The color of someone's skin has been shown to influence sentencing outcomes (Blair et al., 2004).

Black women are disproportionately infected with the Human Immunodeficiency Virus (HIV). The root cause of this is low self-esteem caused by darker than average skin tone (Christensen, 2016). With more exact measurements, more studies can be done to more specifically mitigate these risk factors early on.

Community edge intensity is a sociospatial feature that shows the composition of space of communities and their borders. The ability to distinguish homogeneous boundaries, no

boundaries, clear-cut boundaries, and edge intensity (Legewie & Schaeffer, 2016) can all be strengthened vis-à-vis a robust way to measure skin color. RGB data from a computer algorithm can significantly enhance our understanding of ethnoracial composition and neighborhood conflict.

Many countries ask their citizens to self report their race and ethnicity. This gives researchers a great deal of data to study. However, ethnoracial categories are poorly captured, and classifications are ambiguous as to category and color heterogeneity (Telles et al., 2015). A substantial improvements could be made if instead of self-reporting, governments took RGB data off passports, driver's licenses, or other forms of government identification.

Conclusion

Not much understanding of race relations is created by separating people as either Black or White. The world is not that simple, and the solutions to everyone being equal is not simple as well. Racism has also evolved to a point where research is unable to make further progress without a system for understanding the minute details of racism on skin tone. In the past, research has focused on using the Massey-Martin Scale of Skin Color Darkness. It has been used because of its simplicity of design, but due to its bias and inaccuracy, it should no longer be used in modern research on race relations. The use of an eyedropper like that found in many software like Adobe Photoshop is a much more accurate and concise measure. The eyedropper is the preferred tool for further research in understanding race relations and the negative implications society can place on those with darker skin tones.

References

Alang, S., McAlpine, D., McCreedy, E., & Hardeman, R. (2017). Police brutality and black health: Setting the agenda for public health scholars. *American Journal of Public Health*, *107*(5), 662–665. https://doi.org/10.2105/AJPH.2017.303691

Banton, M. (2012). The colour line and the colour scale in the twentieth century. *Ethnic and Racial Studies*, *35*(7), 1109–1131.

Birdsong, C. K. (2017). Does Appearance Matter?: The Effect of Skin Tones on Trustworthy and Innocent Appearances. In *ProQuest Dissertations and Theses*. http://ir.ua.edu/handle/123456789/3276

Blair, I. Y., Judd, C. M., & Chaplean, K. M. (2004). The influence of Afrocentric facial features in criminal sentencing. *Psychological Science*, *15*(10), 674–679. https://doi.org/10.1111/j.0956-7976.2004.00739.x

Bornstein, B. H., Laub, C. E., Meissner, C. A., & Susa, K. J. (2013). The Cross-Race Effect: Resistant to Instructions. *Journal of Criminology*, *2013*, 1–6. https://doi.org/10.1155/2013/745836

Campos-Castillo, C. (2019). Exploring skin color and black-white differences in trust in physicians in a cross-sectional study of US adults. *National Medical Association*. https://www.sciencedirect.com/science/article/pii/S0027968418302098

Christensen, P. R. (2016). Interrelationships of colorism, violence, and sexual behaviors among southern African-American women [Walden]. In *ProQuest Dissertations and Theses*. https://scholarworks.waldenu.edu/dissertations

Currie, M., Paris, B., Pasquetto, I. V, Pierre, J., Sands, A. E., & Lievrouw, L. A. (2015). Police officer-involved homicide database project. In *University of California*.

http://www.bjs.gov/index.cf

Davis, R. E., Couper, M. P., Janz, N. K., Caldwell, C. H., & Resnicow, K. (2010). Interviewer effects in public health surveys. *Health Education Research*, *25*(1), 14–26. https://doi.org/10.1093/her/cyp046

Edwards, F., Lee, H., & Esposito, M. (2019). Risk of being killed by police use of force in the United States by age, race–ethnicity, and sex. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(34), 16793–16798. https://doi.org/10.1073/pnas.1821204116

Finch, B. K., Beck, A., Burghart, D. B., Johnson, R., Klinger, D., & Thomas, K. (2019). Using crowd-sourced data to explore police-related-deaths in the United States (2000–2017): the case of fatal encounters. *Open Health Data*, *6*(1), 1. https://doi.org/10.5334/ohd.30

Fuerst, J. G. R., Lynn, R., & Kirkegaard, E. O. W. (2019). The effect of biracial status and color on crystallized intelligence in the us-born African–European American population. *Psych*, *1*(1), 44–54.

Goswami, G., Bharadwaj, S., Vatsa, M., & Singh, R. (2013). On RGB-D face recognition using Kinect. *IEEE 6th International Conference on Biometrics: Theory, Applications and Systems, BTAS 2013*. https://doi.org/10.1109/BTAS.2013.6712717

Hannon, L, & DeFina, R. (2013). *Just Skin Deep? Interviewer Assessment of African American Respondent Skin Tone*.

Hannon, Lance, & Defina, R. (2016). Reliability concerns in measuring respondent skin tone by interviewer observation. *Public Opinion Quarterly*, *80*(2), 534–541. https://doi.org/10.1093/poq/nfw015

Hannon, Lance, Problems, R. D.-R. and S., 2014, undefined, Defina, R., Problems, R. D.-R. and

S., & 2014, undefined. (2014). Just skin deep? The impact of interviewer race on the assessment of African American respondent skin tone. *Springer*, *6*(4), 356–364. https://doi.org/10.1007/s12552-014-9128-z

Hersch, J. (2008). Scholarship@Vanderbilt Law Profiling the New Immigrant Worker: The Effects of Skin Color and Height Recommended Citation Profiling the New Immigrant Worker: The Effects of Skin Color Profiling the New Immigrant Worker: The Effects of Skin Color and Height. *Source: Journal of Labor Economics*, *26*(2), 345–386. https://doi.org/10.1086/587428

Hill, M. E. (2002). Race of the interviewer and perception of skin color: Evidence from the multi-city study of urban inequality. *American Sociological Review*, 99–108.

Holland, S. (2020). *Examining Diversity in YouTube Music Videos with Queer Women Couples from 2006–2019*. http://search.proquest.com/openview/59d13b97e58a53ab76e69b7937726415/1?pq-origsite=gscholar&cbl=18750&diss=y

Horta, M. (2012). *PERLA*. PERLA. https://perla.princeton.edu/category/blog/

Kim, D., Kyung, W., … H. H.-2012 I. 16th, & 2012, U. (2012). Selective skin tone reproduction using preferred skin colors. *Ieeexplore.Ieee.Org*. https://ieeexplore.ieee.org/abstract/document/6241721/

King James Version. (2016). Holy Bible. In *Thomas Nelson* (Issue 215). https://doi.org/10.1093/nq/s4-IX.215.118-c

Krieger, N., Chen, J. T., Waterman, P. D., Kiang, M. V., & Feldman, J. (2015). Police Killings and Police Deaths Are Public Health Data and Can Be Counted. *PLoS Medicine*, *12*(12). https://doi.org/10.1371/journal.pmed.1001915

Krieger, N., Sidney, S., & Coakley, E. (1998). Racial discrimination and skin color in the CARDIA study: Implications for public health research. *American Journal of Public Health*, *88*(9), 1308–1313. https://doi.org/10.2105/AJPH.88.9.1308

Lamas, A. (2018). *The salient effects of colorism among Mexican American college males*. http://search.proquest.com/openview/3ba9c4494af34add446becdfa290ca19/1?pq-origsite=gscholar&cbl=18750&diss=y

Legewie, J., & Schaeffer, M. (2016). Contested boundaries: Explaining where ethnoracial diversity provokes neighborhood conflict. *American Journal of Sociology*, *122*(1), 125–161. https://doi.org/10.1086/686942

Louie, P., & Wilkes, R. (2018). Representations of race and skin tone in medical textbook imagery. *Social Science and Medicine*, *202*, 38–42. https://doi.org/10.1016/j.socscimed.2018.02.023

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: a meta-analytic review. *Psychology Public Policy and Law*, *7*(1), 3–35. https://doi.org/10.1037/1076-8971.7.1.3

New International Version. (1994). *Quest Study Bible* (M. Shelley (ed.)). Zondervan.

PANTONE. (2020). *PANTONE SkinTone Guide*. PANTONE. https://www.pantone.com/products/fashion-home-interiors/pantone-skintone-guide

Soetedjo, A., Yamada, K., & Limpraptono, F. Y. (2010). Lip detection based-on normalized Rgb chromaticity. *The 6th International Conference on Information & Communication Technology and Systems*, 63–67. http://eprints.itn.ac.id/3707/

Telles, E., Flores, R. D., & 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. https://doi.org/10.1016/j.rssm.2015.02.002

Uenuma, F. (2018). *The first criminal trial that used fingerprints as evidence*. Smithsonian Magazine. https://www.smithsonianmag.com/history/first-case-where-fingerprints-were-used-evidence-180970883/

Villarreal, A. (2010). Stratification by skin color in contemporary Mexico. *American Sociological Review*, *75*(5), 652–678.

Tables

Descriptive Statistics--Massey-Martin Color Guide												
	N	Range	Minimum	Maximum	Mean		Std. Deviation	Variance	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
1	48	6	212	218	214.52	.217	1.502	2.255	-.061	.343	-.584	.674
2	48	6	168	174	171.21	.166	1.148	1.317	.013	.343	.296	.674
3	48	8	148	156	150.17	.223	1.548	2.397	1.038	.343	2.922	.674
4	48	4	124	128	125.85	.160	1.111	1.234	.300	.343	-.685	.674
5	48	5	107	112	109.50	.158	1.092	1.191	.000	.343	.358	.674
6	48	5	90	95	92.38	.185	1.282	1.644	-.057	.343	-.499	.674
7	48	4	78	82	80.10	.161	1.115	1.244	-.213	.343	-.754	.674
8	48	4	71	75	73.23	.155	1.077	1.159	-.161	.343	-.827	.674
9	48	3	76	79	77.40	.129	.893	.797	-.140	.343	-.780	.674
10	48	3	61	64	62.62	.097	.672	.452	.178	.343	-.271	.674
Valid N (listwise)	48											

Table 1

Descriptive Statistics for eyedropper of Massey-Martin Color Guide. IBM SPSS Statistics Version 26.

Figures

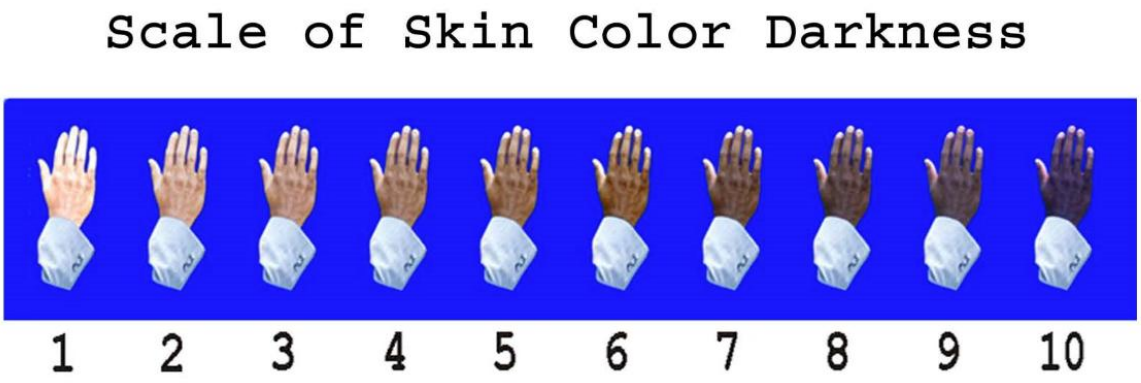

Figure 1. Massey-Martin skin color guide for coders.

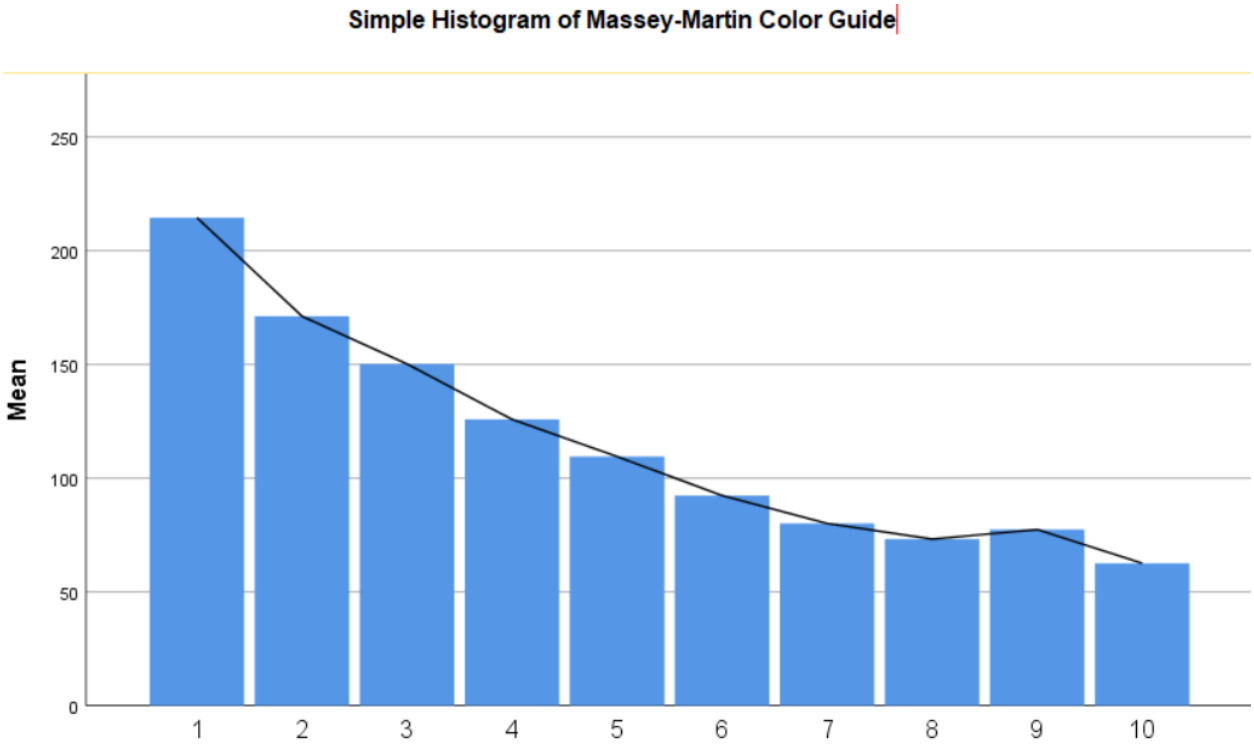

Figure 2. Massey-Martin skin color guide for coders. IBM SPSS Statistics Version 26

Figure 3. PANTONE SkinTone™ Guide (PANTONE, 2020)