Race and the beauty premium: Mechanical Turk workers' evaluations of Twitter accounts

Anne Groggel
PhD candidate at the Department of Sociology, Indiana University Bloomington

Shirin Nilizadeh
Assistant Professor at the Department of Computer Science and Engineering, University of Texas at Arlington

Yong-Yeol Ahn
Associate professor at the School of Informatics, Computing, and Engineering, Indiana University Bloomington

Apu Kapadia
Associate Professor of Computer Science at the School of Informatics, Computing, and Engineering, Indiana University Bloomington

Fabio Rojas
Professor at the Department of Sociology, Indiana University Bloomington

The Version of Record of this manuscript has been published and is available in Information, Communication & Society, 15 Nov 2018, http://www.tandfonline.com/, DOI: 10.1080/1369118X.2018.1543443.

Abstract

Race, gender, and physical attractiveness strongly affect perceptions of trustworthiness and subsequent face-to-face interactions. This study examines how social media users' perceived gender, race, and physical attractiveness can impact their standing online. We test these broad hypotheses by having Amazon Mechanical Turk (MTurk) workers evaluate a sample of 816 Twitter accounts. Our results show a 'beauty premium,' where MTurk workers say they are more likely to follow Twitter accounts with attractive profile photos, and attractive photos are positively associated with evaluations of trust. However, very attractive Black male and female Twitter accounts are associated with lower evaluations of trust compared to their White counterparts. These findings suggest that social media users' social characteristics, perceived from their username or profile image, can replicate offline inequality online.

Introduction

Research on face-to face interactions, such as job interviews, has found that race, gender, and attractiveness strongly affect perceptions of competence, trustworthiness, and personality traits (Segal-Caspi, Roccas, & Sagiv, 2012; Zarghooni, 2007). It is important to understand the role these social characteristics play online as social media becomes a dominant information source (Silverman & Singer-Vine, 2016). To understand how users evaluate Twitter accounts, we investigate the significance of their username or profile image as social cues. Individuals are dependent on these cues to classify whether someone is similar to himself or herself and perceive credibility online (Lin, Spence, & Lachlan, 2016; Walther & Jang, 2012). These cues take on greater importance because Internet users often do not personally know the source of the information (Metzger, Flanagin, & Medders, 2010). While scholars have examined gender differences in Twitter behavior, such as comparing male and female journalists' interactivity (Parmelee, Roman, Beasley, & Perkins, 2017), the consequences of whether one is perceived as a man, woman, or unattractive online remains undertheorized. We build upon scholarship on credibility of a source and their content as a perception (Fogg & Tseng, 1999) by examining the potential consequences of Twitter users' perceived social characteristics through MTurk workers' evaluations.

First, we examine whether workers are more likely to express a desire to follow trustworthy Twitter accounts. Second, we explore how Twitter users' perceived race, gender, and attractiveness influence evaluations of trust and potential visibility. Lastly, we study the effect of MTurk workers' self-reported race since in-groups social boundaries influence evaluations of trust and individuals often exhibit a preference for similar others offline (Lazarsfeld & Merton, 1954; McPherson, Smith-Lovin, & Cook, 2001).

Data and methods

Through the crowdsourcing platform Amazon Mechanical Turk, workers were paid to evaluate 816 Twitter profiles. Each worker evaluated a total of three profiles and three workers randomly evaluated each profile. Workers were US residents and at least 18 years old and answered questions about the accounts' username, name, profile description, profile image, and an attention check question asking whether a profile image had been uploaded. We obtained a sample stratified by followers from an initial sample of Twitter's 'Gardenhose' stream in February 2015 and who selected English as their language. The 'Gardenhose' stream is a 10% sample of Twitter's entire public stream (McKelvey & Menczer, 2013).

We used mixed effects logistic regression models with weights for workers and Twitter accounts and ran models including an interaction between race, gender, and physical attractiveness to compare types of Twitter accounts. Our variables of interest were how trustworthy MTurk workers perceived a Twitter account and the likelihood that they would follow the account. Because many factors can affect these outcomes, we incorporate the following control variables: account's total number of tweets, account age (years), whether the user had a description, had a URL, profile provided a location, the number of accounts that the user followed, and whether the account user was a verified account. Models examining whether the worker would follow the account include measures of trustworthiness, whether they could use the profile image to identify this person in public, whether they could use the account to establish a person's real identity, and workers' gender and race. A summary of how variables were coded is provided in Table 1.

Table 1: Summary of Variables

	Description	Variable Categories	Notes
Trust	*How trustworthy is the man/woman?*	0-untrustworthy and very untrustworthy, 1-above average, trustworthy, and very trustworthy	
Desire to Follow	*How likely are you to follow this account?*	0- unsure, unlikely, and very unlikely, 1- very likely and likely	
Gender	*Workers estimate gender based on the profile image, description, username, and name.*	0-Male, 1-Female, 2-Ambiguous	If the combined characteristics had 50% or more agreement then the account was classified into binary gender categories. Accounts that did not match this threshold were classified as ambiguous.
Race	*Please indicate whether the adult pictured in the profile appears to be: (list of racial categories)*	0-White, 1-Black, 2-Other racial category	Other racial category consists of American Indian, Alaska Native, Asian, or Native Hawaiian Pacific Islander.
Attractiveness	*How physically attractive is the man/woman?*	1-Very Unattractive, 2-Unattractive, 3-about average, 4-Attractive, 5-Very attractive	We created variables of female, male, and overall attractiveness.
White Homophily	*N/A*	0-White worker evaluated White account, 1- White worker evaluated Black accounts, 2-White worker evaluated account classified as American Indian, Alaska Native, Asian, or Native Hawaiian Pacific Islander	Racial homophily limited to the evaluations of White workers because our worker sample was predominantly White.
Person in Public	*How easy or difficult would it be to use this profile image to identify this person in public?*	0-difficult or very difficult, 1- very easy, easy, or neither	
Real Identity	*How easy or difficult would it be to use this account to establish a person's real identity?*	0-difficult or very difficult, 1- very easy, easy, or neither	

Results

Approximately 50% of our sample was classified into binary gender categories. Table 2shows that about 25% are perceived as female and 25% as male. 35% of our sample was classified as White, 54% of the sample was Black.

Table 2: Descriptive Statistics

	Mean	Std	Minimum	Maximum
Male	0.25	-	0	1
Female	0.25	-	0	1
Gender Ambiguous	0.50	-	0	1
White	0.35	-	0	1
Black	0.54	-	0	1
Other Races	0.10	-	0	1
Male Attractiveness	3.14	0.75	1	5
Very Unattractive	0.02	-	-	-
Unattractive	0.12	-	-	-
About Average	0.58	-	-	-
Attractive	0.25	-	-	-
Very Attractive	0.03	-	-	-
Female Attractiveness	3.66	0.85	1	5
Very Unattractive	0.01	-	-	-
Unattractive	0.06	-	-	-
About Average	0.33	-	-	-
Attractive	0.44	-	-	-
Very Attractive	0.15	-	-	-
Male Trustworthiness	3.09	0.73	1	5
Very Untrustworthy	0.02	-	-	-
Untrustworthy	0.13	-	-	-
About Average	0.60	-	-	-
Trustworthy	0.22	-	-	-
Very Trustworthy	0.02	-	-	-
Female Trustworthiness	3.24	0.69	1	5
Very Untrustworthy	0.01	-	-	-
Untrustworthy	0.07	-	-	-
About Average	0.62	-	-	-
Trustworthy	0.26	-	-	-
Very Trustworthy	0.04	-	-	-
Worker Follow	1.80	0.98	1	5
Very Unlikely	0.49	-	-	-
Unlikely	0.30	-	-	-
Unsure	0.13	-	-	-
Likely	0.06	-	-	-
Very Likely	0.01	-	-	-
Friends	733.31	1648.11	1	24,868
Account Age (years)	3.61	1.83	0	8
Status Updates	9793.12	17879.36	1	189,485
Verified Account	0.01	-	0	1
Has Description	0.96	-	0	1
Has URL	0.29	-	0	1
Has Location	0.91	-	0	1
Has Image	0.99	-	0	1
Person in Public	3.32	0.98	1	5
Real Identity	3.49	0.81	1	5
N=2,448				

Table 3 describes our worker sample: 54% are male and 45% female, 48% have a bachelor's degree or higher, and approximately 85% of the sample is White. Our sample had higher levels of education and was less racially diverse than the general population of the United States. About 39% of respondents were 18–29 years old. In addition, nearly every worker reported using social networking sites; and 81% of the workers self-reported having a Twitter account (Table 3).

Table 3: Demographic of Mturk Workers

Variables	Mean
Sex	
Male	0.54
Female	0.45
Prefer not to answer	0.01
Relationship Status	
Married	0.28
Divorced	0.07
Single	0.48
Living with partner	0.13
Separated	0.03
Widowed	0.01
Age	
18-29 years old	0.39
30-44 years old	0.45
56-59 years old	0.12
over 60 years old	0.04
Race	
Hispanic	0.08
Black	0.08
White	0.85
Asian	0.08
American Indian or Alaska Native	0.03
Hawaiian or Pacific Islander	0.01
Level of Education	
Less than High School	0.13
Some College	0.38
Bachelor's degree	0.40
Graduate School	0.08
Income	
Less than 15,000	0.12
15,000-39,000	0.22
40,000-59,000	0.22
60,000-79,000	0.12
80,000 or more	0.15
N=956	

Table 4 provides a summary of our findings, which outline an important association between perceived characteristics and evaluations of trust and hypothesized visibility.

Table 4: Summary of Findings

	Overall
Workers are more likely to follow trustworthy Twitter accounts.	N/A
Evaluations of Trust	
Workers are more likely to evaluate attractive accounts as trustworthy.	✓
Workers are more likely to evaluate White accounts as trustworthy compared to other racial categories	✓
Potential Visibility	
Workers are more likely to follow attractive Twitter accounts	✓
Workers are more likely to follow White Twitter accounts compared to Black Twitter	N/A
White workers are more likely to follow White accounts compared to other racial categories	✓

Results based upon mixed-effect logistic regression analyses ✓ Significant Finding; N/A -not a statistically significant relationship

Trust and desire to follow twitter account

We found that while trustworthiness was positively associated with whether the worker indicated they would follow the account, this relationship was not statistically significant.

Trust

We find that attractiveness was positively associated with evaluations of trust ($p < 0.001$). But while a beauty premium appears to lead to positive evaluations of trust for all types of Twitter accounts, the effect of attractiveness carries greater importance for some Twitter users than others. Based on our estimates, the premium of good looks was considerably smaller for White males than they are for White females or Black males.

Very attractive White females had a higher predicted probability of being viewed as trustworthy than very unattractive White females, holding other variables at their means ($p < 0.05$). Similarly, very attractive Black male and Black female Twitter accounts had a higher predicted probability of being perceived as trustworthy than very unattractive Black accounts ($p < 0.001$). Very attractive Black female accounts had a higher probability compared to their unattractive counterparts ($p < 0.001$).

Our results also suggest that workers were more likely to evaluate White Twitter accounts as trustworthy. Black male and female Twitter accounts were viewed as less trustworthy compared to their White counterparts. White male accounts had a higher probability of being viewed as trustworthy compared to accounts believed to belong to Black male accounts ($p < 0.01$). Likewise, White female accounts had a higher probability of being evaluated as trustworthy compared to Black female accounts, holding other variables at their means ($p < 0.001$).

Desire to follow

Overall, we find a beauty premium, with workers more likely to follow Twitter accounts with images of very attractive individuals compared to accounts evaluated as very unattractive. However, while attractive Twitter accounts fare better, the association between physical attractiveness and whether the worker would follow the account is only significant for White

male Twitter accounts and Black female Twitter accounts. Very attractive White males have a higher predicted probability that the worker would follow the account compared to very unattractive White males ($p < 0.05$). Similarly, very attractive Black females have a higher probability of the worker following their account than very unattractive Black female accounts, holding other variables at the means ($p < 0.001$). In terms of perceived race, while Black accounts were negatively associated with the likelihood that workers would follow the account this relationship was not statistically significant.

White homophily

We found evidence of a significant association between workers' evaluations and racial homophily. White workers were more likely to evaluate White accounts as trustworthy. In this case, racial homophily, defined as White workers evaluating profile images of White individuals, takes on greater significance for unattractive accounts. Black male and female Twitter accounts evaluated as unattractive were less likely to be viewed as trustworthy than their White counterparts by White workers ($p < 0.01$).
Also, White workers were more likely to follow Twitter accounts perceived as White compared to other racial categories. We find that workers' race is significantly related to the hypothesized action of following both male and female Twitter accounts ($p < 0.05$). While our analyses cannot fully test for in-group racial preference, these findings reveal the importance of exploring the relationship between racial homophily and online visibility.

Limitations

When exploring differences in trustworthiness and potential visibility across race, attractiveness, and gender we lost statistical power because we were examining smaller subgroups. Given this restriction we focused on racial homophily and include worker's self-reported gender as a control variable. Future analysis will explore how social media users' perceived social characteristics fully relates to the demographics of their online audience. While further experimental work is needed to confirm a causal effect of Twitter users' race, gender, or attractiveness, our comparisons illustrate the importance of considering the intersectionality of social media users' self-presentation when studying online interactions. Trust may be influenced by other variables such as Tweet content or sentiment.

Another limitation is that workers' evaluations do not capture a live Twitter audience. In our study, an active link was embedded that populated a new window for workers to view Twitter profiles and they could read through previous tweets. To address this shortcoming, we tracked the amount of time workers spent evaluating three profiles. Workers spent roughly three and a half minutes looking at each profile and answering all of the corresponding questions. However, MTurk workers are not the same as social media users. In this study, we explore individual preferences of workers as a source of homophily, but cannot examine biases defined by the chances that people interact with individuals or test the effects of homogeneity of Twitter as a whole. Understanding how trustworthiness relates to social categorization carries implications for interventions to reduce biases. For instance, light-skinned participants assigned to enact a dark-skinned body in virtual reality had reduced implicit racial bias (Peck, Seinfeld, Aglioti, & Slater, 2013).

Despite these limitations, our study calls for understanding how biases may affect online behavior. Social media profiles provide cues that produce evaluations based on social characteristics. These evaluations have the power to impact interactions online. This also means that online platforms have the opportunity to destabilize such perceptions. Future work

can assess how perceptions of trust are potentially mediated by gender, race, and attractiveness and the malleability of these perceptions.

In the previous paragraphs, we analyzed the differences in whether MTurk workers evaluated a Twitter account as trustworthy or the likelihood that they would follow the account. We examine possible social psychological drivers that reflect broad hypotheses of this variation: gender, race, physical attractiveness, and racial homophily. By showing a beauty premium and that White MTurk workers are more willing to trust same-race Twitter accounts, we hope to shed light on various explanations of visibility and potential discrimination online. This study calls for additional work on understanding how Twitter users' race or attractiveness may influence evaluations of trust and indirectly their online visibility. Understanding these mechanisms carries important implications for designing and fostering trust in online platforms (Bos, Olson, Gergle, Olson, & Wright, 2002).

Conclusion

In principle, online platforms can reduce the scope of biases. But how social media users evaluate each other each other remains undertheorized. Small heuristic cues provided in Twitter accounts lead to evaluations of social characteristics. In turn, these evaluations affect trust and whether users might follow their account. Our study sets the groundwork for testing this premise by exploring how social factors such as race or attractiveness relate to Twitter users' trustworthiness or visibility. Our results suggest that in many ways, online social characteristics may activate many of the same perceptions they do offline. We find a beauty premium with attractive Twitter accounts being positively associated with evaluations of trust and workers being more likely to follow their account compared to unattractive accounts. We also found that accounts perceived as Black were negatively associated with these outcomes. White male and female accounts had a significantly higher probability of being viewed as trustworthy compared to Black accounts. We attribute this to racial homophily with assessments being impacted by racial similarity. In our study, White workers were more likely to positively assess Twitter accounts with White profile images than accounts depicting nonwhite individuals. Our results provide compelling evidence that greater attention must be paid to understanding the role perceived characteristics play in shaping online visibility and trust. While our findings need to be experimentally validated, this study highlights the necessity of understanding how perceived race structure online interactions.

References

Bos, N., Olson, J., Gergle, D., Olson, G., & Wright, Z. (2002). *Effects of four computer-mediated communications channels on trust development*. Proceedings of the SIGCHI conference on human factors in computing systems, pp. 135–140. ACM.

Fogg, B. J., & Tseng, H. (1999). *The elements of computer credibility*. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '99, pp. 80–87. New York, NY, USA: ACM.

Lazarsfeld, P. F., & Merton, R. K. (1954). Friendship as a social process: A substantive and methodological analysis. *Freedom and Control in Modern Society*, *18*(1), 18–66.

Lin, X., Spence, P. R., & Lachlan, K. A. (2016). Social media and credibility indicators: The effect of influence cues. *Computers in Human Behavior*, *63*, 264–271.
McKelvey, K. R., & Menczer, F. (2013). *Truthy: Enabling the study of online social networks*. Proceedings of the 2013 conference on Computer supported cooperative work companion, pp. 23–26. ACM.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, *27*(1), 415–444. doi: 10.1146/annurev.soc.27.1.415

Metzger, M. J., Flanagin, A. J., & Medders, R. B. (2010). Social and heuristic approaches to credibility evaluation online. *Journal of Communication*, *60*(3), 413–439.

Parmelee, J. H., Roman, N., Beasley, B., & Perkins, S. C. (2017). Gender and generational differences in political reporters' interactivity on Twitter. *Journalism Studies*. Advance online publication.

Peck, T. C., Seinfeld, S., Aglioti, S. M., & Slater, M. (2013). Putting yourself in the skin of a black avatar reduces implicit racial bias. *Consciousness and Cognition*, *22*(3), 779–787.

Segal-Caspi, L., Roccas, S., & Sagiv, L. (2012). Don't judge a book by its cover, revisited: Perceived and reported traits and values of attractive women. *Psychological Science*, *23*(10), 1112–1116.

Silverman, C., & Singer-Vine, J. (2016). Most americans who see fake news believe it, new survey says.

Walther, J. B., & Jang, J. W. (2012). Communication processes in participatory websites. *Journal of Computer-Mediated Communication*, *18*(1), 2–15.

Zarghooni, S. (2007). *A study of self-presentation in light of facebook*. Institute of Psychology, University of Oslo.