COMPUTATIONAL COMMUNICATION RESEARCH 4.1 (2022) 208-253
HTTPS://DOI.ORG/10.5117/CCR2022.1.006.MALI

ARTICLE

Representations of Racial Minorities in Popular Movies

A Content-Analytic Synergy of Computer Vision and Network Science

Musa Malik
University of California Santa Barbara, Department of Communication, Media Neuroscience Lab

Frederic R. Hopp
Amsterdam School of Communication Research, University of Amsterdam

René Weber
University of California Santa Barbara, Department of Communication, Media Neuroscience Lab

Abstract

In the Hollywood film industry, racial minorities remain underrepresented. Characters from racially underrepresented groups receive less screen time, fewer central story positions, and frequently inherit plotlines, motivations, and actions that are primarily driven by White characters. Currently, there are no clearly defined, standardized, and scalable metrics for taking stock of racial minorities' cinematographic representation. In this paper, we combine methodological tools from computer vision and network science to develop a content analytic framework for identifying visual and structural racial biases in film productions. We apply our approach on a set of 89 popular, full-length movies, demonstrating that this method provides a scalable examination of racial inclusion in film production and predicts movie performance. We integrate our method into larger theoretical discussions on audiences' perception of racial minorities and illuminate future research trajectories towards the computational assessment of racial biases in audiovisual narratives.

Keywords: inclusion, film, computer vision, network science, computational communication research

In today's saturated media landscape, the film industry remains a central branch for attracting media users and generating revenue (MPAA, 2019; UNIC, 2017). By the end of 2019, global box-office sales topped a record-breaking, annual revenue of $42.5 billion (McClintock, 2020); the online streaming service Netflix reported 8.8 million new subscribers (Spangler, 2020); and 814 movies entered production in the United States alone, an increase of eleven percent compared to 2015 (MPAA, 2019). Clearly, the film industry continues to touch the lives of myriad media consumers. At the same time, a significant portion of voices remain underrepresented in cinematographic productions. Most notably, mounting evidence demonstrates a critical gender gap in the film industry: Movie directors are primarily male (Smith et al., 2017), thereby reinforcing the overall dominance of males in movies (Ramakrishna et al., 2017; Smith & Choueiti, 2010), while simultaneously underrepresenting or stereotypically portraying females (Lauzen, 2018b; University, 2017; Wood, 1994). In fact, Lauzen (2018a) showed that there are twice as many male speaking characters as female in the average movie.

Recently, Smith and colleagues (2020) showed that these biases extend beyond gender, with films featuring racial minorities in lead roles receiving significantly less production budget, a key predictor of box-office sales (Eliashberg et al., 2007; Eliashberg et al., 2014; Gopinath et al., 2013; Stimpert et al., 2008). Yet, research into the cinematographic representation of racial minorities remains sparse and methodically limited in important ways: First, the majority of studies have focused on simple frequency counts among cast members (e.g., Smith et al., 2020), thus providing limited insights into the story centrality and the dynamic portrayal of non-White characters. Second, efforts to examine the *visual* portrayal of racial minorities have largely relied on manual, costly, and non-scalable content-analysis, for instance, to measure the percentage of screen time that a given character receives (Smith & Choueiti, 2011).

We herein improve upon these shortcomings via the following, innovative steps: First, we leverage advancements in computational face recognition to detect specific cast members on-screen. Second, utilizing a suite of image processing tools, we combine our face recognition metrics with visual features that capture White and non-White characters' dynamic portrayal in motion-pictures. In particular, we extract cinematographic features that capture how filmmakers place and orient the camera to portray characters. Next, by considering the position of characters in a movie's social (character) network, we compute detailed statistics about the *structural* representation of racial minorities. The application of these network metrics has recently provided valuable insights into the network position of female characters (e.g., Kagan et al., 2020), highlighting that female characters are more likely to be assigned peripheral, decentralized character network positions. We build on and extend this line of research by examining the character network positions of White and non-White characters. Furthermore, we combine characters' race with their social network position and cinematographic representation to predict how these variables influence the critical and financial performance of a movie. Finally, we probe the scalability and accuracy of computational race coding by performing an evaluation of the FairFace (Kärkkäinen & Joo, 2021) computer vision package which assigns individual cast members to one of seven racial groups: White, Black, Native American, East Asian, Southeast Asian, Middle East, and LatinX.

Concentration of Media Ownership

In recent years, Hollywood has been increasingly criticized for lacking diversity in talent recruitment and content production. Racial and ethnic minorities still face significant barriers to entering the entertainment industries, mostly expressed by a noticeable absence on-screen (Molina-Guzmán, 2016). Concentration of media ownership–increasing industry consolidation and the dominance of a small number of companies over content production (Noam, 2009)–contributes to such homogeneity in thought and practice (Kim & Brunn-Bevel, 2020). This concentration is exemplified by the fact that Hollywood, for all practical purposes, refers to the six production houses that collectively dominate market share; Disney (38%), Warner Bros. (13.8%), Universal (13.4%), Sony (11.7%), Lionsgate (6.8%), and Paramount (5%). The remaining companies constitute 11.2% altogether (Statista, 2020). Theoretical frameworks from political economy suggest that financial incentives often take priority over racial justice considerations. Any meaningful assessment of the latter is mostly in relation to the potential of economic returns such efforts may contribute to the enterprise (Croteau & Hoynes, 2014). Thus, production houses seek to minimize economic risk by investing in projects that have historically attracted large audiences. Following such profit-driven logic has traditionally implied catering to the White demographic–at-least within domestic U.S. audiences–which reflect the racial majority (76.3%) in the United States (U.S. Census, 2019). Here, the assumption is that the presence of racial-ethnic minorities, including the narration of their life experiences, will not appeal to the White majority audience which will subsequently choose not to watch movies that feature racially diverse casts. In addition, projects that do integrate racial-ethnic minorities are frequently supplemented with White cast members, whose inclusion may be intended to dilute any negative feelings elicited in audience members due to the presence of an overall diverse cast (Kim & Brunn-Bevel, 2020). The trade-off between meeting diversity targets as well as ensuring maximum profits is a challenging task for filmmakers who strive to "move the needle" (AII, 2020) towards a more inclusive representation of racial-ethnic minorities. One strategy of mitigating this financial risk is to reduce budget allocation towards such projects (Smith et al., 2020). At the same time, constraining production and marketing resources negatively impacts the cinematographic quality of inclusive narrative projects and furthermore causes these productions to only reach small audiences, limiting greater demand for similar projects in the future.

Race-Related Media Effects

In view of this media ownership concentration, a growing tradition of race-related research is emerging that explores the effects that media portrayals of racial-ethnic minorities have on audiences. Social Identity Theory (SIT; Tajfel & Turner, 1986) posits that individuals experience a collective identity based on their membership to a specific group, including race-related groups. This theoretical framework states that audiences rely on positive media depictions to support their group identity and associate unfavorable media depictions with negative judgments of non-group members. Accordingly, research indicates that negative media depictions of minorities tend to increase adverse social responses among White audiences, including negative stereotyping (Figueroa-Caballero et al., 2019; Mastro & Tukachinsky, 2011) and expressions of beliefs in intergroup threat (Atwell Seate & Mastro, 2016). Building on SIT, self-categorization theory (Turner et al., 1987) argues that individual identity is sustained via the cognitive placement of the self and others into group-based categories. Collective group identities emerge when individuals perceive themselves as a member of a group on the basis of shared beliefs, or on observable traits such as race and gender. These traits distinctly mark a group's characteristics as different and separate from outsiders (Hornsey, 2008). Indeed, research has suggested that the more similar race-based media depictions are to perceived White norms, evaluations of minority behaviors become more positive as these are viewed as acceptable among White audiences. On the other hand, when race-based media depictions deviate from perceived White norms, favorability towards minorities decreases among White individuals (Mastro & Kopacz, 2006;

Ramasubramanian, 2010). The formation of such collective group identities may then also exacerbate existing media selection biases. Viewers may prefer watching content that favorably emphasizes their ingroup, and in turn, maintain a positive social identity (Trepte, 2006). Racial representation in Hollywood films holds the potential to mitigate polarization in content preferences as viewers are unable to exhaustively avoid exposure to promotional film materials that highlight the presence of a racially diverse cast. Such exposure may indeed influence their media consumption habits and positively shift attitudes towards racial minorities.

Research also shows that exposure to favorable media depictions of communities of color decreases perceived threat and social distance amongst White audiences (Dalisay & Tan, 2009; Ortiz & Harwood, 2007). At the same time, unfavorable media depictions, such as associations with criminality, can enhance negative perceptions of minority groups and further marginalize their members (Abraham & Appiah, 2006; Hurley et al., 2015). Furthermore, exposure to racial media depictions over a longer time horizon can shape audience expectations in real life. Consistent negative portrayals of underrepresented minorities may augment biased attitudes of White audiences that have minimal real-life contact with these communities and can influence White audiences to assume that media portrayals of specific racial groups are accurate, leading to a heavier reliance on stereotypical heuristics for judgement (Mastro et al., 2007).

Cinematographic Techniques

As argued previously, substantial research indicates that the overrepresentation of White characters in media and the underrepresentation of people of color plays an important role in influencing racial perceptions in society. However, whether this literal representation bias translates into cinematographic choices remains an open question. In the cinematographic literature, it is well known that visual representations of characters shape perceptions of these characters with consequences for attitude formation. For instance, screen time has a significant impact on shaping audience attitudes towards story content. Exposure to specific characters over longer time horizons may strengthen perceptions of their importance to the development of the plot for viewers. Likewise, the scale of character framing can be conceptualized as the apparent distance of characters from the camera and is one of the most effective visual devices in regulating the relative size of characters' faces to the background (Rooney & Balint, 2018). Particularly the close-up is one of the most recognizable units of cinematographic discourse which allows a filmmaker to direct the attention of their audience towards the subject on-screen. The emphasis of all facial expressions enacted by the subject, and its subsequent influence on the efficacy of communicating character intent, helps stimulate an empathetic response in the viewer. As a result, the importance of the subject is momentarily enhanced (Doane, 2003). Gradation along the shot scale allows filmmakers to cinematographically portray characters in a way which most accurately supports their artistic intentions. In fact, research suggests that these choices of shot scale do in fact influence the psychology of audiences, with patterns of shot types effectively impacting levels of self-reported arousal (Canini et al., 2011), inducing feelings of empathetic care towards victims (Cao, 2013), and even regulating attitudes towards oppositional viewpoints (Mutz, 2006).

Furthermore, the two-dimensional positioning of a character depends on how central the filmmaker intends their character to be on-screen. A character may be situated closer to the center of the frame or deviate along the horizontal and vertical planes. These deviations are typically a function of intentional changes in camera angles by the director, and such character positioning has long been argued to impact audience perception (Millerson, 1961). Indeed, video production fundamentals often highlight the idea that lower viewpoints can impart strength to the character while elevated viewpoints may induce perceptions of character weakness, thereby eliciting completely different attitudes towards the character from the audience (Tiemens, 2009).

Moreover, crowd scenes, where a large number of characters are simultaneously present on-screen, is a relevant cinematographic choice that shifts attention away from the role an individual character plays and instead emphasizes characteristics that are typically associated with herd-like behavior (Tratner, 2003). This perspective is particularly informative in the context of movies as the visual representation of groups suggests that the moods, practices, and opinions of characters existing within those groups may not be under their rational control. Indeed, the reduction of individual actions being framed as savagelike, animalistic, and hysterical within group settings is a frequent theme in both film and literature (Kølvraa, 2013). In contrast, individual-specific shots can increase the level of identification an audience may experience in relation to the character on-screen, thereby encouraging them to evaluate the developing plot from their perspective as well (Katz, 1991). Keeping in mind the psychological consequences of visual representations, it is thus important to determine how often characters appear on screen, both alone and in a group context, and what the specific visual representations of characters are.

Evaluating Diversity, Equity, and Inclusion

Ethnic and racial minorities currently comprise 40% of the U.S. population and demographers predict that by 2050 the U.S. will be majority-minority (Desilver, 2015). Notably, White audiences currently account for less than 50% of moviegoers, while Black and African-American and LatinX communities engage in traditional and digital entertainment in greater proportions than their population shares in the U.S. (Gonzalez, 2014; MPAA, 2014). Given the value of promoting equal representation, there now exist economic arguments for establishing standards regarding diversity, equity, and inclusion (DEI) in Hollywood image production. Accordingly, a number of tests for measuring gender and racial representation in film, television, and literary works have recently been introduced. Here, the majority of literature discusses the Bechdel Test, a popular measure to detect gender asymmetries in artistic productions by establishing a simple rule that content "has to contain at least two women in it, who talk to each other, about something besides a man" (Bechdel, 1985). Notably, a few research attempts have already been made to computationally determine levels of gender bias in Hollywood movies (Agarwal et al., 2015; Kagan et al., 2020). In the past few years, a suite of new tests has been constructed to serve as DEI benchmarks for Hollywood movies, particularly aimed to benefit female minorities (Hickey et al., 2017). Given their recent introduction, research implementing these tests remains sparse, although they provide starting guidelines on how structural representation of minorities can be conceptualized and operationalized in empirical terms (see Table 1 in the *SI*).

While the above tests provide useful benchmarks for understanding how racial minorities should be represented on-screen, they remain constrained in two important ways. First, they are limited to certain categories of race (or gender) and hence only specific storylines, written for characters belonging to these communities, are able to successfully pass them. Second, given the number of scripts produced every year, the problem of scalability and efficiency for implementing and performing these tests becomes ever-more salient. Currently, prominent DEI scholars like the Annenberg Inclusion Initiative (AII; https://annenberg.usc.edu/research/aii) focus on various metrics (e.g., frequency counts) to demonstrate that films misrepresent women and people of color (Smith et al., 2020). While these initiatives are frequently confronted with the problem of scalability, they additionally face challenges for providing deeper insights into structural biases that are propagated within a film, some of which the recently developed tests have attempted to uncover. However, as manual annotation of racial and gender biases in films is a costly and labor intensive task, it is no surprise that large-scale attempts at understanding the structural and visual biases, latent in Hollywood movies, have not been undertaken.

Computational Approaches

Building computational solutions for detecting racial biases in films provides researchers with the advantage to circumvent various of the above mentioned challenges based on clearly defined, empirical metrics whilst ensuring scalability and efficiency. A recently introduced framework for the scalable examination of character representations adopts computational tools from network science. In this line of research, narratives are typically computationally parsed, abstracted, and parsimoniously encoded in a mathematical object called a graph (Bollobás, 2001), where the set of parts (network nodes) frequently represent a story's characters and connections (network edges) between characters denote some quality or strength of their (social) relation (e.g., Hopp et al., 2020; Jones et al., 2020; Kagan et al., 2020; Skowron et al., 2016). Conceptualizing narratives in the form of such social character networks (SCN) affords the examination of characters' *structural* representation via distinct graph-based properties. For example, a character's degree centrality (i.e., the total number and/or weight of connections this character has to other characters in the story) reveals information about the character's overall (social) importance to the story, whereas a character's betweenness centrality (i.e., the number of shortest paths between any couple of characters in the network that passes through the target character) indicates the extent to which this character mediates between other characters. Notably, research has started to rely on these network metrics for identifying structural biases in film productions. For instance, a recent study by Kagan and colleagues (2020) demonstrated that 3.5 times more relationship triangles in SCNs of movies have a majority of men, and in the top-10 most central movie roles, there is a majority of men. Yet, the authors also report an improvement in equality over the years, with women inheriting more important roles in contemporary movies than in the past.

Computer Vision

Building on and extending this line of research, we examine how such structural biases extend beyond the gender gap to characters with a racial minority background. Furthermore, although a character's social network position affords examinations of various *structural* biases, they do not provide information about the *visual* marginalization of these characters. Here, the computational extraction of visual features such as the presence, location, and cinematographic portrayal of human characters in motion pictures is made possible through significant advancements in computer vision. A growing line of research has attempted to understand semantics in movies via the synergy of audio-visual-text information and has focused on understanding the relationships between movie characters (Bamman et al., 2013; Park et al., 2009; Weng et al., 2009), reconstructing high-level storylines from visual content (Huang et al., 2018; Zhu et al., 2015), graphing and clustering emotional arcs of a movie (Chu & Roy, 2015), and classifying genres of movies based on input from trailer content (Simoes et al., 2016; Zhou et al., 2010). Research in this domain has primarily benefited from the advent of real-time object detection capabilities made possible by convolutional neural networks.

Existing object detection approaches, which are also a boon for modern face detection solutions, are mainly categorized into two primary domains, two-stage detectors and one-stage detectors. Two-stage detectors such as Region Based Convolutional Neural Networks (R-CNN; Girshick et al., 2014), and its extended generations Fast R-CNNs (Girshick, 2015) and Faster R-CNNs (Ren et al., 2015), divide the detection task into two stages: extracting region of interest (RoI) and then doing classification and bounding box regression tasks based on the RoI. Despite high accuracy rates, these detectors still involve expensive computations and offer relatively slower inference speeds. On the other hand, one-stage detectors such as the YOLO (You Only Look Once) series (Redmon et al., 2016), and its extended generations YOLOv2 (Redmon & Farhadi, 2017), YOLOv3 (Redmon & Farhadi, 2018), and YOLOv4 (Bochkovskiy et al., 2020) eliminate the RoI extraction stage entirely and directly classify and regress the candidate anchor boxes. While they do provide improved speed and precision, one-stage detectors demonstrate this optimized trade-off between accuracy and speed only on high performance Graphics Processing Units (GPUs), making their implementation harder for some projects. Our present research leverages tools and approaches that rely on both one-stage and two-stage detectors for extracting features from visual content.

Computational solutions that allow for the automated detection of race and ethnicity are another novel application of modern computer vision techniques. However, widespread adoption of such solutions has been obstructed by a lack of datasets that are 1) exhaustively reflective of the diversity in facial features across multiple racial groups, and 2) robustly annotated based on standardized facial coding procedures (Fu et al., 2014). Popular datasets leveraged for the delivery of automated solutions towards race classification include FERET (Phillips et al., 1998), MORPH-II (Ricanek & Tesafaye, 2006), CAFE (LoBue & Thrasher, 2015), LFWA+ (Liu et al., 2015), among others. While these datasets independently consist of thousands of images, recent computational experiments have demonstrated that models trained on images from a single dataset are unable to generalize to others (Kärkkäinen & Joo, 2021). This limitation has been attributed to the fact that existing public face datasets are strongly imbalanced, typically biased toward Caucasian faces, and this imbalance contributes towards the racial asymmetric accuracy observed against colored communities (Buolamwini & Gebru, 2018; Raji & Buolamwini, 2019). Furthermore, as there exist few standardized annotation procedures in face analysis tasks, annotator-induced racial biases currently remain unchecked. For instance, humans tend to perform significantly better when categorizing faces of people of their own race (Fu et al., 2014). Thus, it is critical for the success of robust race classification models that exhaustive and independent annotation procedures are implemented. FairFace (Kärkkäinen & Joo, 2021) is arguably the most complete dataset available to researchers. The FairFace dataset contains 108,501 images with an emphasis on balanced race composition. Furthermore, the recently developed FairFace computer vision package achieves state-of-the-art generalizability in automated race classification tasks. Thus, in the present manuscript, for the evaluation of a computational approach to assess characters' race, we also use the FairFace computer vision package to automatically assign individual characters to unique racial groups.

Research Questions

Keeping the aforementioned arguments and explanations in mind, our contribution seeks to address the following two central research questions:

RQ1: What is the relationship between movie characters‘ race, their social network position, and their cinematographic representation?
RQ2: Does movie characters' race, their social network position, and their cinematographic representation relate to movie performance?

Methods

Data, code, and supplemental information (*SI*) for all reported analyses have been made available on the Open Science Framework (OSF) at https://osf.io/3rmtu/.

Movie Selection

We selected a total of 103 popular movies released between 2007 and 2018 (Figure 1; Table 2, *SI*). To ensure variation in the racial composition of our movie collection, we selected movies that vary in their proportional representation of underrepresented minorities in the overall cast. This information had already been collected by a team of human coders at the Annenberg Inclusion Initiative (AII; https://annenberg.usc.edu/research/aii; see Smith et al., 2020). Due to various complications in obtaining and computationally analyzing each movie's corresponding, full-length motion picture (described below), our final collection consisted of 89 movies.

Cinematographic Feature Extraction

To explore the capabilities of computer vision approaches to algorithmically analyze and extract features from full length movies, we bought the Digital Video Disc (DVD) for the movies in our selection. Given that the algorithmic processing of movie content requires a digital format, we used the popular and open-source video transcoder *HandBrake* (HandBrake; https://handbrake.fr) to convert all movies to the digital multimedia container format (MPEG-4)[1].

As a first step, we determined which cast member appears on screen and for how long. For this purpose, we leveraged the open-source library *Face Recognition* (Face Recognition; https://github.com/ageitgey/face_recognition) built on top of the popular C++ toolkit *Dlib* (Dlib; http://dlib.net). *Dlib* provides rich deep learning models to assist with face detection, face feature extraction, and face recognition tasks. Its pre-trained facial landmark detector estimates the location of 68 (x, y) coordinates that map to landmark facial structures. These estimates are then compared with a database of known images that we provide the Face Recognition library with to train on (see section on *Computational Race Coding* for details on image collection protocols) to determine the identity of the character in the movie frame. Furthermore, we utilize the *Yolov4* object detection model (Bochkovskiy et al., 2020), built on *Darknet*–an open source neural network framework (Darknet; https://pjreddie.com/darknet/)–to predict the (x, y) coordinate locations and counts of all humans detected within a single frame. While custom editing of the source code for the previously mentioned libraries can also provide us with similar inferences, leveraging Yolov4 has clear advantages to it. First, Yolov4 is designed with the intention to optimize parallel computing and improve the speed of object detection. Yolov4 provides the capacity to process up to 180 frames/second. Second, our results showed that Yolov4 gives highly accurate results for our task of face detection with minimal errors (see the instructional notebook made available on OSF for model specifications). Such detailed estimates are necessary for the extraction of character-specific visual features. We leveraged the extracted facial coordinates at unique time-stamps and quantified the cinematographic measures discussed previously as follows:

Screen Time. This measure indicates the number of frames a character appears in a full-length movie. We conceptualize screen time as $(ST)_{n,\,i}$ which simply refers to the total number of frames where the individual character appears in.

Scale. This measure refers to the relative area that the bounding box around a detected face encapsulates within a single frame. We conceptualize scale $(s)_i$ as $\frac{(x_{\max} - x_{\min}) \times (y_{\max} - y_{\min})}{F_{area}}$ where $(x_{\max} - x_{\min})$ and $(y_{\max} - y_{\min})$ respectively refer to the width and height of the bounding box associated with an individual character and F_{area} is a frame dimensional constant equal to (1280×720).

Frame Centrality . This measure quantifies how much the character's face, on a 2-dimensional plane, deviates from the center of the frame. We conceptualize horizontal frame centrality $(FC)_{h,\,i}$ as $\left|x_{mid} - \left(\frac{Fx_{\max}}{2}\right)\right|$ and vertical frame centrality $(FC)_{v,\,i}$ as $\left|y_{\mathrm{mid}} - \left(\frac{Fy_{\max}}{2}\right)\right|$ where x_{mid} and y_{mid} respectively refer to the axial center of the bounding box around a detected face and $\left(\frac{Fx_{\max}}{2}\right)$ and $\left(\frac{Fy_{\max}}{2}\right)$ respectively refer to the central coordinates of the frame along the (x, y) axes. Positive values along the x-axis indicate that the detected face is to the right side of the frame while negative values show that the detected face is to the left side of the frame. In a similar fashion, positive values along the y-axis indicate that the detected face is present in the lower quadrants while negative values show that the detected face is present in the upper quadrants. Consequently, values closer to 0 demonstrate greater frame centrality along the relevant axis.

Groups. This measure highlights the total number of detected faces within a single frame. We conceptualize groups as $(G)_{n,\,i}$ which simply refers to the number of additional characters associated with the unique character recognized by our facial recognition pipelines.

Singular Shots. This measure specifies the number of frames a unique character appears on screen, independent of additional characters. We conceptualize singular shots $(SS)_{n,\,i}$ as $(ST)_{n,i}|D_{n,t} = 1$ where $(ST)_{n,\,i}$ refers to the total number of frames unique to a character while $D_{n,t}$ refers to the total number of faces D_n detected at time equals t. We consider $(SS)_{n,\,i}$ equals 1, if and only if, it is also observed that $D_{n,t}$ equals 1 when a unique character is recognized by our facial recognition pipelines.

UR/White Frame Ratio. This measure signals the average ratio of Underrepresented to White characters present when a unique character i appears on screen. We conceptualize this ratio $(R)_i$ as $\frac{(UR)_{n,i}}{(W)_{n,i}}|D_{n,t} \geq 2$ where $(UR)_n$ and $(W)_n$ respectively refer to the total number of Underrepresented and White characters recognized.

We rely on the general image processing capabilities provided by the library open source computer vision (OpenCV; https://opencv.org/) to assist in the processing of the above measures.

Character and Cast Members

We obtained the characters for each of our 89 movies via the *Subs2Network* (Kagan et al., 2020) dataset (described in more detail below). To obtain the corresponding actor/actress name for each character, we first retrieved each movie's unique Internet Movie Database identifier (IMDb; https://www.imdb.com/) by entering the title as query parameter into the application programming interface (API) of the Open Movie Database (OMDb; https://www.omdbapi.com/). Thereafter, we entered the movie's IMDb ID into a Uniform Resource Locator (URL) linking to the movie's full cast (e.g., https://www.imdb.com/title/tt1205489/fullcredits points to *Gran Torino*) and then computed the intersection between Kagan et al.'s (2020) characters and the IMDb cast. This produced a total of 1,559 unique characters (1,406 unique cast members).

Human Coding of Cast Race

Two trained undergraduate research assistants (RA) were instructed to independently research background information on cast members in an effort to assign each character to one of the following, commonly studied racial categories: White, Black and African American, East Asian, South Asian, Middle Eastern, Multiracial, LatinX (Hispanic), and Native American. RAs overlapped in their judgment of race in 1,319 cases (84.60%). When RAs disagreed, the first author resolved the disagreement with the RAs and assigned the character to one of the eight racial categories. No information could be obtained for two cast members and were subsequently dropped from the analysis. In addition, as we could not obtain reliable images for the purposes of our computer vision pipeline (described below) for each individual cast member, the total number of characters for our analysis reduced to 1,377 (1,235 unique cast members).

Computational Race Coding

To probe the scalability and accuracy of computational race coding, we obtained high quality images for each cast member using Google image search. We observed that for 20 unique cast members no reliable images could be found and thus we dropped those cast members from our computational visual analysis. These cast members were often not well-known and peripheral to plot development. Next, we utilized FairFace (Kärkkäinen & Joo, 2021) on each cast member's high-resolution photograph and every cast member was assigned to one of seven racial groups: White, Black, Native American, East Asian, Southeast Asian, Middle East, and LatinX. It is important to note that FairFace does not provide any metrics on Multiracial individuals, a category we found to be frequent during our manual-annotation process.

Social Character Network Metrics

We used the *Subs2Network* (Kagan et al., 2020) dataset to obtain the social character network (SCN) for each movie. Kagan and colleagues (2020) created SCNs by analyzing subtitles via the construction of "the movie social network G := ⟨V, E⟩, where V is the network's vertices set, and E is the set of links among the network's vertices. Each vertex v ∈ V is defined to be a character in the movie. Each link e := (u, v, w) ∈ E is defined as the interaction between two movie characters u and v, w times," (Kagan et al., 2020). A link was created between characters u and v if they appeared in the movie in a time interval less than 60 seconds. Finally, all edges with weight lower than 3 were removed to reduce the number of false positive edges (for further information and validations, please refer to Kagan et al., 2020, but see limitations below). Next, we computed the following graph metrics to index each character's centrality:

Degree Centrality. The degree centrality $C_d(v)$ of a character is simply the number of connections this character has to other characters. The higher the number of connections, the more central the character. $C_d(v)$ is formulated as $\frac{|\Gamma(v)|}{|V|-1}$ where $\Gamma(v)$ is defined as a set of v characters (Brandes & Erlebach, 2005).

Closeness Centrality. Closeness centrality $C_c(v)$ indicates how close a character is to all other characters in the network. It is calculated as the average of the shortest path length from the character to every other character in the network. $C_c(v)$ is formulated as $\frac{1}{\sum_s d(v,s)}$ where $d(v,s)$ refers to the distance between vertices v and s (Brandes & Erlebach, 2005).

Pagerank. Pagerank provides a network centrality score which indicates the influence of a particular node in the context of the entire network. It is measured for a particular node by taking into account the number of nodes, as well as their individual centrality measures, pointing towards that node. The Pagerank centrality score for node *v,* i.e. (x_v), is formulated as $\alpha\left(\sum_s a_{s,v} \frac{x_s}{L(s)}\right) + \frac{1-\alpha}{N}$ where α is an attenuation factor ∈ (0, 1), *N* indicates the number of nodes within a set of nodes *V, L*(*s*) refers to the number of neighbouring nodes for node *s*, and ($a_{s,v}$) is an adjacency matrix, i.e. ($a_{s,v}$)= 1 if vertex *s* is linked to vertex *v* and ($a_{s,v}$)= 0 otherwise (Gleich, 2015). Lastly, for every movie in our selection, we mapped each character's manually-obtained race as node attributes to the corresponding social character networks in the *Subs2Network* dataset.

Movie Performance Metrics

MetaCritic Score. For each movie in our collection, we obtained the *MetaCritic* (MetaCritic; https://www.metacritic.com/) score, which reflects a movie's average rating by professional movie critics and is commonly considered as an indicator of story strength.

International Rent. As an indicator of a movie's final performance, we used international rent, which captures the international revenue generated from the fees theaters pay to rent a film's print for exhibition. In other words, this indicator is the amount the production studio receives after exhibitors take their portion of the box office.

Results

Race Distribution Among Cast Members

We first analyzed the race distribution across our manually coded characters (Figure 1A). White characters are in the majority, followed by Black and African American, East Asian, and Multiracial characters. Furthermore, due to class imbalances across characters' races and the fact that non-White characters are frequently considered to be underrepresented on screen (Smith et al., 2020), we additionally grouped all non-White characters into an underrepresented category (*UR*, Figure 1B). Next, we computed the ratio of *UR* to White characters represented in the cast of our movies (Figure 1C). Notably, our distribution of *UR* characters in casts correlates strongly with Smith et al.'s (2020) reported distribution of *UR* characters in casts (r = .73, p < .000), speaking for the validity of our manual coding of characters' race.

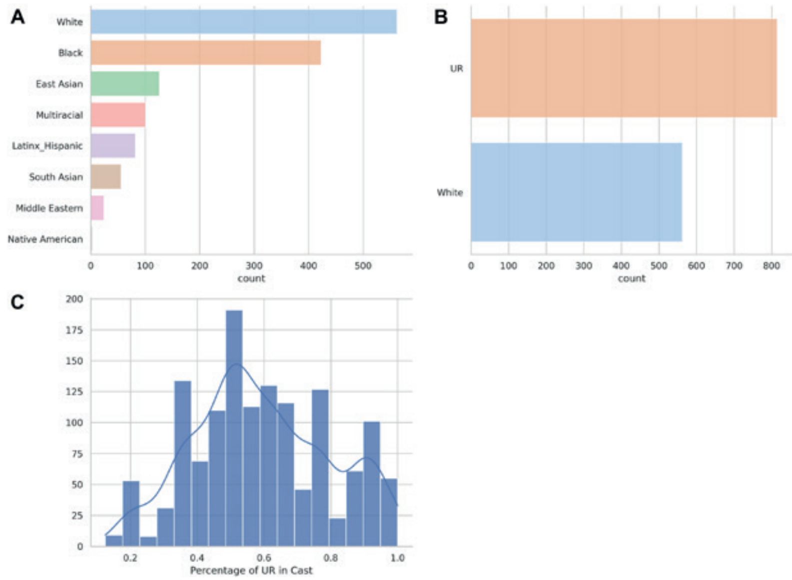


Figure 1. Distribution of Characters' Race Across Movies.

Note. **A**) Frequency plot of manually-coded character race across movies. **B**) Frequency plot of underrepresented (*UR*; non-White) and White characters. **C**) Distribution of percentage of *UR* characters in cast across movies.

Character Network Position

We examined the unique effect of race on characters' network centrality. Due to the right-skewed distribution of pagerank, we applied the natural logarithm to pagerank and applied a min-max scaling to center pagerank between 0 and 1. Figure 2 provides descriptive contrasts for each network measure using characters' race as the independent variable.

To robustly account for the nested nature of our data–character-specific race variables (first level) that are nested within movies (second level)–we relied on Multilevel Linear Models (MLMs) to further test the effect of individual race categories on different indices of character network position.[2] We found that at the first level, as suggested by Figure 2, the composite race variable (*UR*) was not significant in predicting any of the three centrality variables. However, the original race variable with its eight racial categories, which vary from movie to movie, may reveal notable differences. Accordingly, we found that East Asian characters are represented with significantly lower degree centrality $C_d(v)$ as compared to White characters ($\beta = -.244$, $p = .025$), while the effect for Multiracial (compared to White) characters

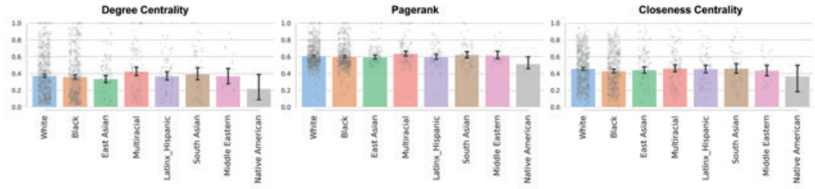


Figure 2. Barplots and Stripplots for Character Network Measures.

Note. Error bars represent 95% confidence intervals based on 10,000 bootstrapped iterations for all characters across all movies (i.e., nested nature of data not reflected).

indicates higher $C_d(v)$ with only marginal significance (β = .182, p = .074). Interestingly, upon comparing the standardized parameter estimates' confidence intervals between East Asian (95% CI [-.456, -.032]) and Multiracial (95% CI [-.017, .381]) characters we found that these are not overlapping which indicates a significant difference in their $C_d(v)$. Regarding closeness centrality $C_c(v)$, we found that East Asian characters are actually represented with significantly lower $C_c(v)$ as compared to White characters (β = -.251, p = .016), while the effect for Multiracial (compared to White) characters indicates higher $C_c(v)$ with again marginal significance (β = .169, p = .080). Our pairwise comparisons indicated that the mean difference for East Asian (95% CI [-.455, -.047]) and Multiracial (95% CI [-.020, .357]) characters of $C_c(v)$ was significant. We found no evidence supporting race differences in predicting pagerank centrality.

Cinematographic Representation

We examined the unique effect of race on how characters are being represented on screen. With the exception of horizontal and vertical centrality, all visual measures (screen time, singular shots, groups, scale, horizontal and vertical centrality, and *UR/White* Frame ratio) were heavily right-skewed, log-normal distributed (see *SI* for distributions) and thus we applied a natural logarithm transformation followed by a min-max scaling to screen time, singular shots, groups, scale, and *UR/White* Frame ratio. Figure 3 provides descriptive contrasts for each cinematographic measure using characters' *UR* status as the independent variable.

As above, we estimated additional MLMs to further test the effect of individual race categories on character-associated visual features (on character level, not aggregated to movie level). Interestingly, we found that Black and African American characters are represented with significantly higher screen time $(ST)_{n,i}$ as compared to White characters (β = .137, p = .047).

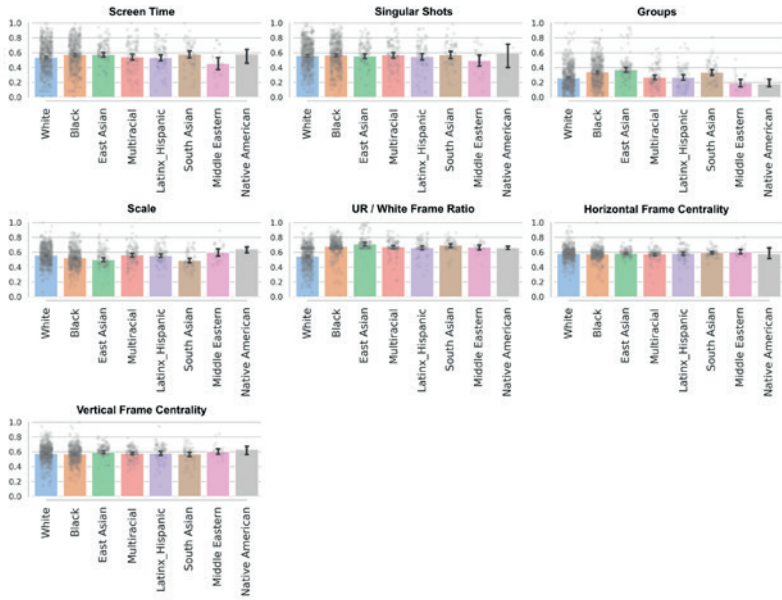


Figure 3. Barplots and Stripplots for Characters' Cinematographic Measures.

Note. Error bars represent 95% confidence intervals based on 10,000 bootstrapped iterations for all characters across all movies (i.e., nested nature of data not reflected).

However, we found no evidence for the involvement of race in predicting singular shots $(SS)_{n,\,i}$. Furthermore, *UR* status was important in predicting higher group counts $(G)_{n,\,i}$(β = .220, p = .00) with significantly higher $(G)_{n,\,i}$ of Black and African American (β = .226 , p = .00), East Asian (β = .648 , p = .00), and South Asian (β = .288, p = .028) characters compared to White characters. Subsequent pairwise comparisons further indicated that mean differences between Black and African American (95% CI [.108, .344]) and East Asian (95% CI [.459, .836]) characters and also between East Asian (95% CI [.459, .836]) and LatinX (95% CI [-.179, .214]), Middle Eastern (95% CI [-.447, .285]), and Multiracial (95% CI [-.181, .168]) characters were significant for predicting $(G)_{n,\,i}$.

We also found that *UR* status was significant in predicting lower scale $(S)_i$ (β = -.173, p = .00). We followed this observation up with an MLM which indicated that Black and African American (β = -.136, p = .027), East Asian (β = -.538, p = .00), and South Asian (β = -.328, p = .014) characters are represented with significantly lower $(S)_i$ as compared to White characters. Our pairwise comparisons further demonstrated that the mean differences between

Black and African American (95% CI [-.256, -.015]) and East Asian (95% CI [-.730, -.346]) characters and also between East Asian (95% CI [-.730, -.346]) and LatinX (95% CI [-.257, .143]) and Multiracial (95% CI [-.154, .201]) characters are significant for predicting $(S)_i$. Further analyses built on this strong interplay between race and visual features and demonstrated that *UR* was significant (β = .817, p = .00) in predicting a higher *UR/White* Frame ratio $(R)_i$ in individual frames throughout the movies (which we again remind here should not be confused with a *UR/White* ratio for the entire cast at the movie level). Interestingly, our analysis enabled us to discover that, with the exception of Native American characters (due to unfortunately only 3 Native American characters represented in our dataset), all character race categories–Black and African American (β = .751, p = .00), East Asian (β = .904, p = .00), LatinX (β = .655, p = .00), Middle Eastern (β = .930, p = .00), Multiracial (β = .758, p = .00), and South Asian (β = 1.223, p = .00) –are represented with significantly higher $(R)_i$ as compared to White characters. This means that *UR* characters were more frequently accompanied by other *UR* characters in the same frame but not as frequently with White characters. In our pairwise comparison, significant mean differences further emerged between South Asian (95% CI [.970, 1.476]) and Black and African American (95% CI [.633, .869]), LatinX (95% CI [.454, .855]), and Multiracial (95% CI [.580, .936]) characters for predicting $(R)_i$. Finally, we found no strong evidence to suggest that either *UR* or any race category was significant in predicting horizontal frame centrality $(FC)_{h,i}$ and vertical frame centrality $(FC)_{v,i}$. To examine how characters' race, their network centrality, and cinematographic representations interplay, we computed a correlation matrix (Figure 4). Notably, we observed that structural and visual features of characters' story centrality correlate in meaningful ways (e.g., pagerank and screen time; r = .50, p <.000 ; pagerank and scale; r = .328, p < .000). In addition, we also observe that singular shots more frequently appear in close-ups (singular shots and scale; r = .398, p < .000), whereas groups appear less frequently in close-ups (singular shots and groups; r = –.648, p < .000).

To assess the synergistic effect of combining visual and structural features, we regressed the structural network measures onto the visual features across White and *UR* characters (Figure 5). We found that for screen time, singular shots, and scale, there are positive relationships with characters' measures of network centrality across both *UR* and White characters, while there exists a negative relationship between groups and network centrality across both *UR* and White characters. Interestingly, we observed that an increase in characters' network centrality does not affect the ratio of *UR*

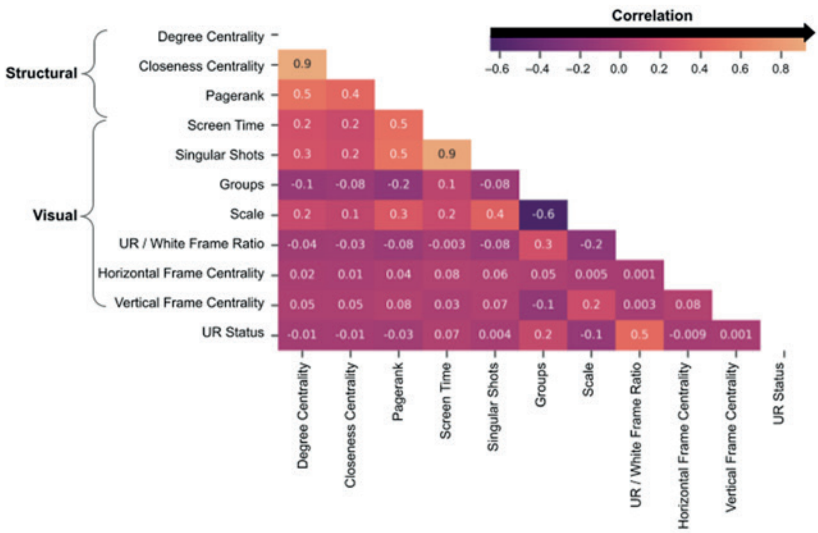


Figure 4. Correlation Matrix Relating Character UR Status, Structural, and Visual Features

Note. UR Code: 1=UR, 0=White.

to White characters on-screen for White characters, but as *UR* characters become more central, the *UR/White* Frame ratio decreases meaning that *UR* characters are less frequently associated with other *UR* characters in the same frame.

Altogether, in response to RQ1, we find convincing evidence for the existence of robust linkages between the structural and visual features of characters' stories and we further observe how character race is particularly meaningful in shaping their subsequent interactions.

Computational Classification of Character Race

To examine the accuracy of character race classification via the FairFace computer vision package, we constructed a confusion matrix between our manual character race codings and FairFace's automated prediction (Figure 6). The agreement between manual and computational codings is indeed promising (Cohen's $\varkappa$ = .63; Kronbach's α = .63), particularly for the top three featured race categories White, Black, and East Asian. However, it must be noted that FairFace does not include a "Multiracial" racial category as our manual codings did. Here, we observed that FairFace's predictions, particularly of White, Black, and LatinX cast members, are frequently assigned a Multiracial race category by our human coders. We believe that

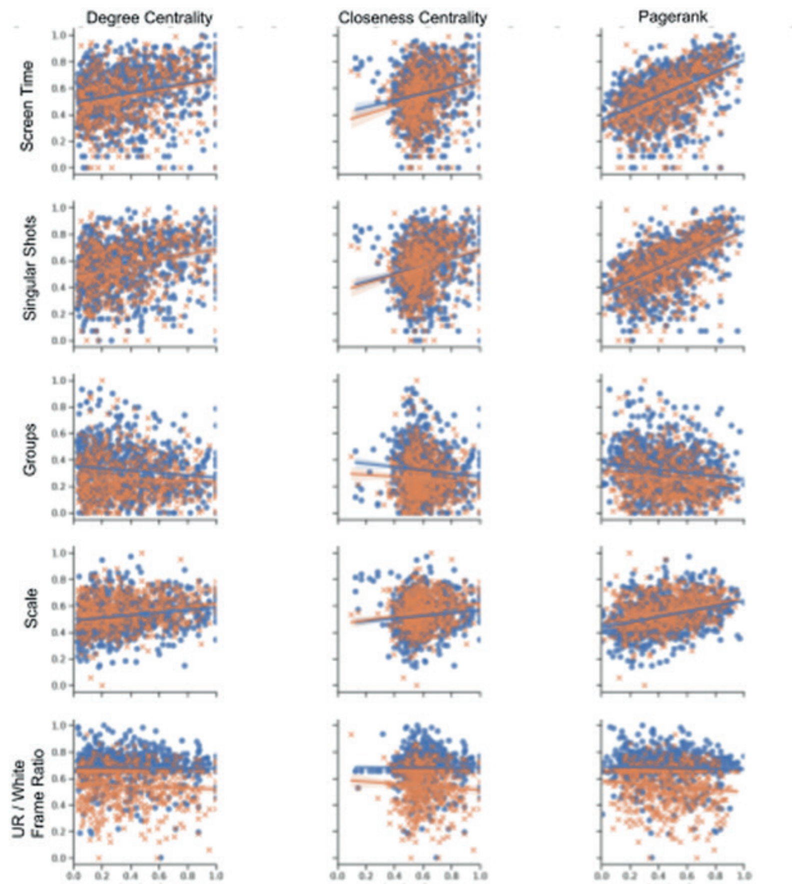


Figure 5. Relationships Between Characters' Race, Structural, and Visual Representation.
Note. Blue dots: White characters; Brown dots: UR characters.

our manual codings are not only reflective of real demographic changes taking place in the United States where the multiracial population is indeed expanding rapidly (Gardner & Hughey, 2017) but they may also be reflective of the generally increasing prevalence of multiracial individuals in American media (Jones, 2008).

Movie Performance Models

For addressing RQ2, we fitted a General Linear Model (GLM). Alongside our eight race categories, we relied on the structural and visual features used in addressing RQ1 above to serve as additional predictor variables.

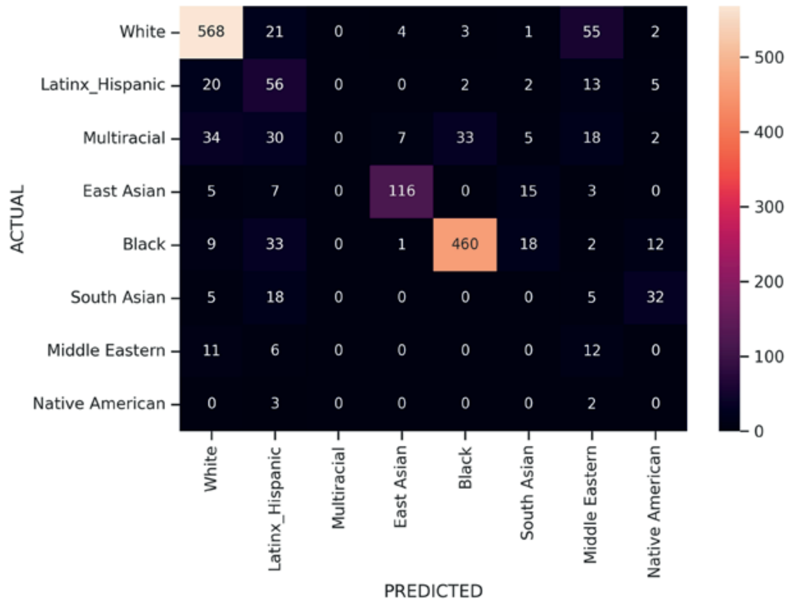


Figure 6. Confusion Matrix Contrasting Manual and Computational Race Classification

Note. Actual = Human coding; Predicted = FairFace prediction.

We used *MetaCritic* as a dependent performance variable [3] and explained 12.0% of the variance (adjusted 10.9%). Most interestingly, we find that movies with Black and African American characters (as compared to White characters) significantly predict a higher *MetaCritic* rating (β = .279, p = .000) and movies with South Asian characters (as compared to White characters) predict a higher *MetaCritic* rating (β = .237, p = .082) with marginal significance. Meanwhile, movies with East Asian characters (as compared to White characters) significantly predict a lower *MetaCritic* rating (β = -.277, p = .006). In fact, upon subsequent pairwise comparison, we observed significant mean differences between Black and African American (95% CI [.146, .411]) and East Asian (95% CI [-.473, -.080]), LatinX (95% CI [-.352, .096]), and Multiracial (95% CI [-.272, .142]) characters for the prediction of *MetaCritic* ratings. Moreover, upon adding interactions of race category with scale and the *UR/White* Frame ratio, we observe that the interaction of East Asian characters with *UR/White* Frame ratio (β = .412, p = .000) and South Asian characters with scale (β = .465, p = .002) significantly predicts a greater *MetaCritic* score (see Figure 4 in *SI* for interaction plots).

Next, we used International Rent (*IntRent*) as our financial performance dependent variable. Due to a strong right-skew in its distribution, we took the natural logarithm of *IntRent* so as to obtain a distribution approaching normality. We were able to successfully explain 12.1% of the variance (adjusted 11.0%).[4] Interestingly, we find that movies with East Asian (β = .602, p = .000) and South Asian (β = .487, p = .000) characters significantly predict greater *IntRent* as compared to White characters. In contrast, movies with Black and African American (β = -.215, p = .002) and *LatinX* (β = -.340, p = .003) characters significantly predict lower *IntRent* as compared to White characters. In our subsequent pairwise comparison, we again observed significant mean differences between Black and African American (95% CI [-.350, -.080]) and East Asian (95% CI [.404, .800]) characters and Black and African American (95% CI [-.350, -.080]) and Multiracial (95% CI [-.062, .358]) characters indicating that while movies featuring Black and African American characters are more positively reviewed than East Asian and Multiracial character movies, those movies fall short of earning revenue in the international market at par with them. Interestingly, we observe that movies with South Asian characters not only receive greater *MetaCritic* ratings but also earn greater revenue internationally as compared to White characters. Moreover, we also notice significant mean differences between LatinX (95% CI [-.566, -.114]) and East Asian (95% CI [.404, .800]) characters, LatinX (95% CI [-.566, -.114]) and South Asian (95% CI [.218, .755]) characters, and LatinX (95% CI [-.566, -.114]) and Multiracial (95% CI [-.062, .358]) characters for predicting *IntRent*.

Upon the addition of interactions of race with scale and the *UR/White* Frame ratio, we find that it is the interactions of scale with Multiracial (β = -.222, p = .041) and South Asian (β = .265, p = .076) characters that predict *IntRent* with significance and marginal significance respectively. We further observe that it is the interactions of *UR/White* Frame ratio with East Asian (β = .295, p = .006) and LatinX (β = .256, p = .081) characters that predict *IntRent* with significance and marginal significance respectively (see Figure 5 in *SI* for interaction plots). Lastly, our pairwise comparisons of Black and African American (95% CI [-.331, -.051]) with South Asian (95% CI [.270, .940]) characters, East Asian (95% CI [.256, .705]) with Black and African American (95% CI [-.331, -.051]) and LatinX (95% CI [-.639, -.159]) characters, and LatinX (95% CI [-.639, -.159]) with Multiracial (95% CI [-.059, .409]) and South Asian (95% CI [.270, .940]) characters demonstrate significant mean differences for predicting *IntRent*.

Discussion

In light of movies' enduring impact on audiences' perception of racial minorities (Ramasubramanian, 2010), including racial minorities' own self-image and identity (Mastro & Kopacz, 2006), identifying biases in the representation of racial minorities remains a critical task for social scientific research. We herein presented results for pushing the envelope of this research by making the following contributions: First, we introduced a computational pipeline for a detailed and scalable examination of group representations in film across multiple modalities of film productions (i.e., subtitles and motion-pictures). By combining the network and visual representations of White and *UR* characters in movies we demonstrated that variations in characters' visual portrayal can be directly linked to their structural position and race status via screen time, scale, groups, and racial homogeneity measures. Notably, we observed that Black and African American and East Asian characters are typically associated with portrayals that emphasize a group structure, giving way to audience perceptions susceptible to influence by latent psychological biases (Kølvraa, 2013). We argue that such cinematic portrayals may further perpetuate stereotypes that stem from unfounded cultural frames that overstate notions of homogeneity, interdependence, and a collectivist focus of identity (Celius & Oyserman, 2001) within marginalized communities.

Moreover, we also find evidence for cinematographic techniques, such as a reduced scale, that deemphasize the importance of Black and African American, East Asian, and South Asian characters in cinematic formats. One of the few techniques that a cinematographer uses in screen media is to portray an enhanced visual of characters' faces which is instrumental in conveying not just performance but also establishing an empathetic relationship between characters and audience (Doane, 2003). This simple enlargement of the face allows for a momentary increase in the importance of the character (and by consequence their race). Surprisingly, across our entire sample of popular movies it was actually White characters who received increased scale portrayal speaking even more strongly to the representational bias this study addresses. Furthermore, the *UR/White* Frame ratio was particularly informative in helping us understand patterns of racial homogeneity that are prevalent on-screen. Surprisingly, characters from underrepresented groups were more frequently associated with members of their own, or another minority group, implying that cinematographers either work on scripts that are frequently racially homogeneous in terms of their cast members or that they deliberately portray *UR* characters in the absence of White characters. While we believe that these visual representations may actually be a reflection of the type of narratives that are scripted and funded for production, we provide in this study a computational framework and potential toolkit for future researchers and practitioners to use in their own work, push for increased awareness, and help chart the way towards greater representation in the cinematic arts.

Second, as our validation of FairFace demonstrates, cast members' race is not always correctly inferred using automated methods. In a past approach, we had provided the FairFace algorithm with Google images scraped from the Internet and had thus hypothesized that inaccuracies stemmed from the low-resolution of the images we obtained. In the present paper, we circumvented that limitation by collecting high-quality, manually-verified images. This resulted in an improved accuracy for automatically classifying the race of movie cast members. Yet, before using FairFace for the scalable, automated classification on a larger set of cast members, we argue that FairFace (or similar automated race detection solutions) may be enhanced by (a) adding a Multiracial race category, and (b) curating a cast member database that adjusts for fluctuations in images' lighting, camera angles, and make-up. To this end, we look forward to contributing towards the improvement of automated race detection technologies, particularly in the highly complex context of cinematic portrayal, and building supplementary datasets, as we do in this study, that can advance the racial categories FairFace is currently capable of predicting.

Finally, in response to RQ2, we have demonstrated that our vision-based metrics possess the capability to predict two important movie performance indicators, one for the quality of movies (*MetaCritic*) from the perspective of film critics, and one for international financial performance (*IntRent*). Notably, we established linkages between the cinematographic metrics of scale and *UR/White* Frame ratio, and their interactions with the Black and African American, East Asian, South Asian, and White character race categories to improve predictions of *MetaCritic* ratings. In fact, *MetaCritic* ratings are considered a reliable proxy for "story strength" (Smith et al., 2020) and are among the strongest predictors of financial performance when other structural variables such as production costs, distribution network, and advertising budgets are controlled for. As we have also established the predictive capability of our computer-vision metrics for movies' financial performance via a "proof of principle" analysis, we are confident that adding more fine grained analyses of story strength at the scene and character level (within movies) will not only substantially improve financial performance models (like in Smith et al., 2020), but also provide recommendations for the artistic creation of movies as well as new insights into the representations of racial minorities in popular movies.

Limitations, Ongoing Work, and Future Directions

As with all studies, our study has limitations. First, our analysis of cast members was limited to characters included in Kagan et al.'s (2020) dataset. Upon manual inspection during our race coding procedures, we discovered that Kagan et al.'s (2020) character list was incomplete across multiple movies. For instance, the movie *After Earth* has a full cast member count of 57 based on information listed on IMDb, with 14 of them listed as being credited for their performance. The prominent actor Will Smith and his son Jaden Smith star in leading roles. Surprisingly, the character network Kagan et al. (2020) reported only includes four of the IMDb listed cast members (all credited) and excludes the two leading actors mentioned above.

Similarly, reported characters from the movie *12 Years A Slave* do not mention the leading antagonist Edwin Epss, played by popular actor Michael Fassbender. While the cast is more representative in the case of *12 Years A Slave*, the reliability of a network that excludes a plot-centric character is questionable. Thus, we recommend care in generalizing the findings reported in this study to the broader movie industry. However, we argue that our introduced computer vision framework identified an important methodological constraint of extant character networks constructed from movie subtitles. When creating character networks using subtitle files (like in Kagan's study), previous research has linked characters if their names co-occur in some predefined window. Here, our framework suggests that solely relying on subtitles may miss important character interactions. Although mentioning a character by name is undoubtedly an important aspect in the formation of character relationships, characters' co-occurrence on-screen may be a much more fine-grained and salient feature for learning how characters relate to each other, particularly because screenplay writers frequently present different communities of characters alternately (Hopp et al., 2020). In turn, we argue that researchers interested in the construction of movie character networks may benefit from adopting a multimodal approach and consequently be able to identify a more representative selection of cast members using automated procedures. Accordingly, for future research, we will supplement, validate, and extend text-based character networks with networks extracted from motion pictures based on our applied face detection and face recognition techniques.

Second, we acknowledge that the computational framework introduced in this paper is complex and necessitates advanced technical expertise.

Thus, researchers interested in the automated evaluation of racial diversity and inclusion in cinematic formats may argue that using screenplays (textual data) would offer a technically less challenging and computationally more efficient alternative. As mentioned earlier, we build on previous work that explored gender representation in film through the application of network science techniques on textual data, specifically subtitles (Kagan et al., 2020). Indeed, an apparent advantage of screenplays over subtitles would be that screenplays contain richer, more descriptive information which may provide greater insights into the context surrounding characters. In fact, for the construction of specific measures such as screen-time, screenplays offer a straight-forward alternative to our approach above. Researchers would simply have to count the number of times a character spoke (already tagged by textual cues) and weigh each observation with the length of the corresponding dialogue. However, beyond this rather straightforward measure for screen-time, we may not be able to proxy other cinematographic measures. For instance, scale is the apparent distance of characters from the camera. Screenplays *may* offer structured text that utilizes film jargon (ECU; extreme close-up, VO; voice-over, POV; point of view) to identify broad temporal moments when the cinematographer is instructed by the screenwriter to emphasize a specific character's face. However, the idiosyncratic nature of both professions may violate this assumption as it is not necessary for every screenwriter to adhere to a single, structured format of writing, just as it is not necessary for a cinematographer to adhere to every single instruction. Relatedly, using screenplays for constructing cinematographic measures also rests on the questionable assumption that there are no substantial deviations in narrative structure across the pre-production (screenplays) and the post-production (final movie) stages. We argue that computer vision techniques are practically simple to apply and can allow us to accurately construct the scale measure even for an individual frame. Similarly, the two-dimensional positioning of a character depends on how central the cinematographer intends the character to be on-screen. They may be situated closer to the center of the frame or deviate along the horizontal and vertical planes. Therefore, frame centrality is again a cinematographic feature that would be most easily extracted using computer vision techniques as it is very likely that such information would be very difficult to articulate precisely in a textual format. Finally, the portrayal of groups, the emphasis on singular shots, and the average ratio of *UR* to White characters present when a unique character appears on screen, are idiosyncratic choices that a screenwriter may not have influence over. These choices are likely dependent on the artistic inclinations of the cinematographer.It is also not necessary for a character to speak in order to simply be present in the camera frame. Similarly, it is not necessary for the camera to focus on the speaking character. The cinematographer may choose to divert the audience's attention towards the recipient(s) of the dialogue or even intersperse a panoramic shot while a character speaks in the background. As the reader may imagine, these visual patterns are not identifiable via textual data.

In addition to the conceptual limitations referenced above, computational modeling of screenplays based on robust natural language processing techniques may likewise be technically complex for many researchers. We assume here that original screenplays that were used on-set for the film production are available online. Thus, a research project of this nature would require the automated scraping of screenplays from multiple public repositories (e.g., Internet Movie Script Database (IMSDb), Daily Scripts, Simply Scripts, and Write to Reel). These screenplays would then need to be parsed into consistent, tabular formats for further analyses which would most certainly involve text-character decoding issues, dependent on the programming language and operating system. Finally, network science techniques may be applied to extract conceptual and linguistic relationships that can assist in the valid and quantifiable estimation of the cinematographic features we have identified above. We argue that a relatively straight-forward computer vision approach (as we have outlined in this manuscript) offers an acceptable compromise between technical complexity and the reliable estimation of cinematographic features. Nonetheless, we strongly encourage future work to integrate these different modalities and robustly identify the structural and visual representation of characters in film narratives.

Third, we understand that our second analytic block (RQ2) about the relationship between representation metrics and film performance (as proxied via MetaCritic scores and BoxOffice returns) remains under-explored. While recent work has indeed demonstrated that the gender and race of cast members can influence a film's success, it only does so after taking into account important structural factors such as production value, advertising budget, and star power (Smith et al., 2020). Thus, we explicitly acknowledge that movie performance models that do not include these structural variables are not to be considered exhaustively robust. However, even in the absence of stricter control, we demonstrate in this paper that visually extracted information alone has the potential to explain a substantial amount of variance in performance outcomes. We believe that our findings inspire the further questioning of the many connections between representation metrics and performance variables and also motivate future work that incorporates computationally extracted visual features in the construction of more robust performance models.

Finally, this study remains largely descriptive with a strong content-analytic focus. Hence, we do not derive any conclusions about the cognitive and behavioral outcomes of the portrayals of racial minorities in film productions. In future work, we thus seek to utilize and combine both graph-based and computer vision movie features to predict audiences' responses to a movie, for instance, by examining stereotyping and judgment across characters from different racial groups or by examining the financial performance and film critics as well as lay audiences' evaluation of a movie based on its diversity, equality, and inclusion.

Conclusion

This paper provides further steps towards the computational implementation of tests geared towards the inclusion of racial minorities, such as the DuVernay, Ko, Villalobos, or Waithe Test (Hickey et al., 2017). Importantly, our herein proposed structural and visual measures provide precise anchor points for evaluating racial biases on multiple levels of analysis, which subsequently can inform concrete decisions during screenplay writing, direction, and production. Taken together, we believe that the herein introduced work will have implications that extend well beyond the study of inclusion in film. Most notably, multimodal content-analysis combining structural, visual, and semantic narrative features is likely a boon for advancing myriad social science theories. For example, classical selective exposure and salience predictions rooted in Social Identity Theory (Tajfel & Turner, 1986; for an overview, see Trepte, 2006) can be tested with novel, more fine-grained measures of character portrayals, and may be extended to both large-scale and field research. Likewise, dynamic-transactional models such as the Model of Intuitive Morality and Exemplars (Tamborini, 2011; Tamborini & Weber, 2020) have largely ignored the interplay between media characters' and audiences' sex, race, and age when examining entertainment outcomes. Accordingly, we think that our contribution provides numerous starting points to investigate these relationships with unprecedented detail, scale, and methodical rigor.

Supplemental Information

Representations of Racial Minorities in Popular Movies: A Content-Analytic Synergy of Computer Vision and Network Science

Table 1. Diversity, Equity, and Inclusion Tests and Criteria.

Tests	Criterion
DuVernay	Minority characters must not be peripheral to the narrative but must have their own full-fledged stories that do not simply support the advancement of White characters' storylines.
Waithe	The film has to have a black woman who exerts authority within a position of power and has a healthy relationship.
Villalobos	The Latina lead — or another Latina character — must be shown as professional or college educated, to be able to speak unaccented English, and must not be portrayed as highly sexualised.
Ko	The film has to show a non-White, female-identifying person who speaks English with dialogues in at least five scenes.

Table 2. Movie Collection

Title	Year
12 Strong	2018
12 Years a Slave	2013
47 Meters Down	2017
47 Ronin	2013
A Haunted House	2013
A Very Harold & Kumar 3D Christmas	2011
A Wrinkle in Time	2018
After Earth	2013
All Eyez on Me	2017
American Gangster	2007
Barbershop: The Next Cut	2016
Beverly Hills Chihuahua	2008
Big Mommas: Like Father, Like Son	2011
Black Panther	2018
Black or White	2015
Body of Lies	2008
Boo! A Madea Halloween	2016
Brooklyn's Finest	2010
Central Intelligence	2016
Chef	2014
Collateral Beauty	2016

College Road Trip	2008
Crazy Rich Asians	2018
Creed	2015
Dance Flick	2009
End of Watch	2012
Epic Movie	2007
Exodus: Gods and Kings	2014
Fast Five	2011
Fences	2016
First Sunday	2008
For Colored Girls	2010
Freedom Writers	2007
Furious 7	2015
Girls Trip	2017
Gran Torino	2008
Joyful Noise	2012
Jumanji: Welcome to the Jungle	2017
Jumping the Broom	2011
Lakeview Terrace	2008
Lee Daniels' The Butler	2013
Life of Pi	2012
Lion	2016
Machete	2010
McFarland, USA	2015
Million Dollar Arm	2014
Ninja Assassin	2009
No Escape	2015
Norbit	2007
Paranormal Activity: The Marked Ones	2014
Peppermint	2018
Predators	2010
Prince of Persia: The Sands of Time	2010
Push	2009
Rambo	2008
Red Tails	2012
Ride Along	2014
Ride Along 2	2016
Rush Hour 3	2007
Safe House	2012
Selma	2014
Seven Pounds	2008
Sex and the City 2	2010

Sicario: Day of the Soldado	2018
Skyscraper	2018
Slumdog Millionaire	2008
Snatched	2017
Snitch	2013
Star Wars: The Last Jedi	2017
Stomp the Yard	2007
Straight Outta Compton	2015
Suicide Squad	2016
Taken 2	2012
The Best Exotic Marigold Hotel	2012
The Best Man Holiday	2013
The Big Sick	2017
The Dictator	2012
The Expendables	2010
The First Purge	2018
The Forbidden Kingdom	2008
The Forest	2016
The Gambler	2014
The Great Debaters	2007
The Great Wall	2017
The Hate U Give	2018
The Kingdom	2007
The Last Airbender	2010
The Meg	2018
The Mummy: Tomb of the Dragon Emperor	2008
The Purge: Anarchy	2014
The Purge: Election Year	2016
The Second Best Exotic Marigold Hotel	2015
The Shallows	2016
The Sisterhood of the Traveling Pants 2	2008
The Soloist	2009
The Wolverine	2013
Think Like a Man	2012
Think Like a Man Too	2014
This Christmas	2007
Uncle Drew	2018
Vantage Point	2008
Whiskey Tango Foxtrot	2016
Widows	2018

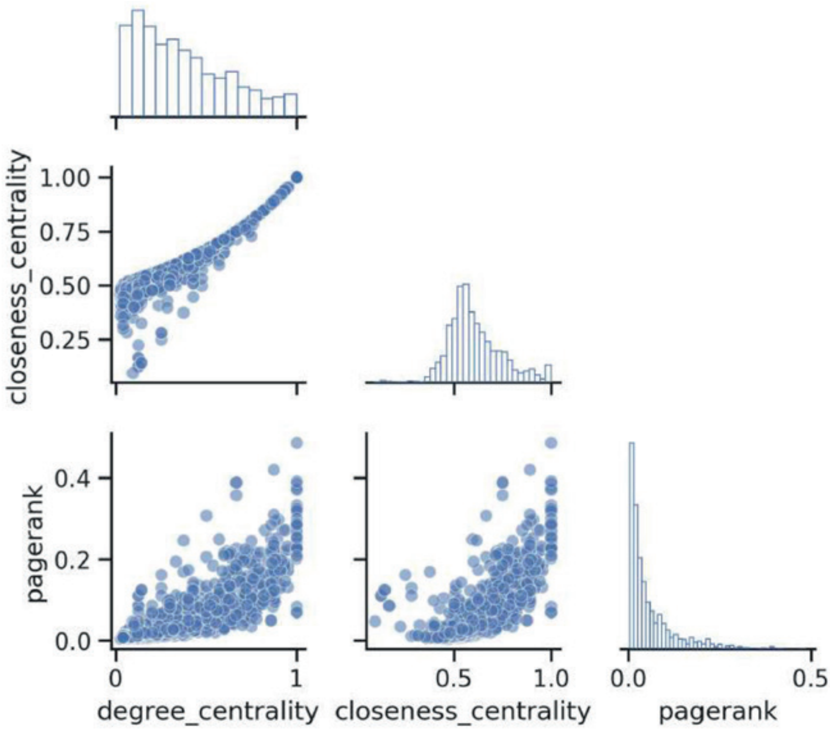


Figure 1. Pair Plot for Characters' Untransformed Network Metrics

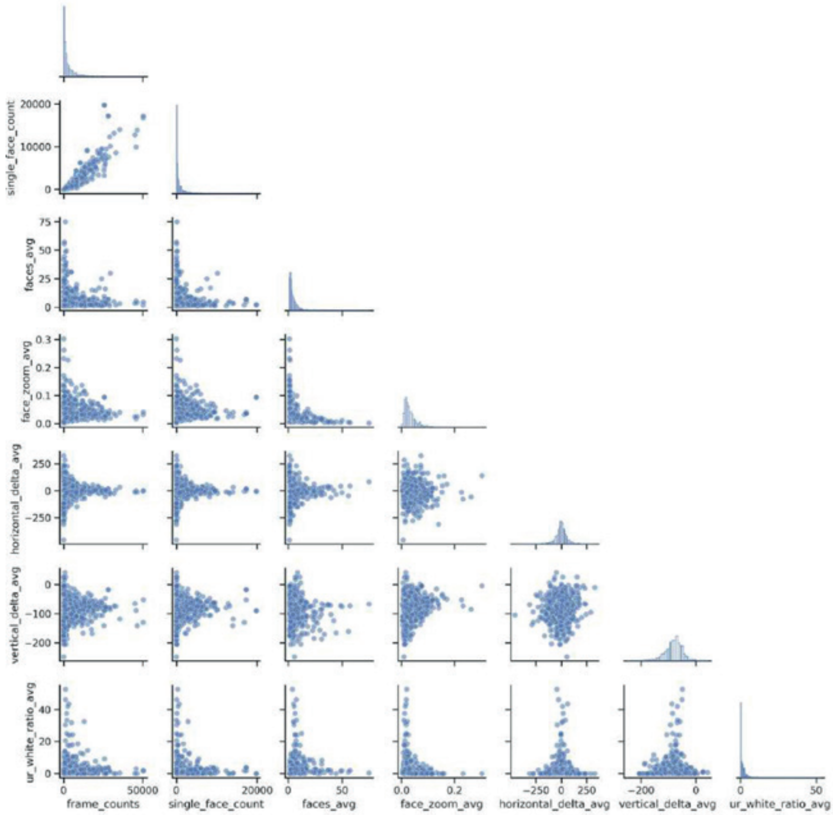


Figure 2. Pair Plot for Characters' Untransformed Visual Metrics

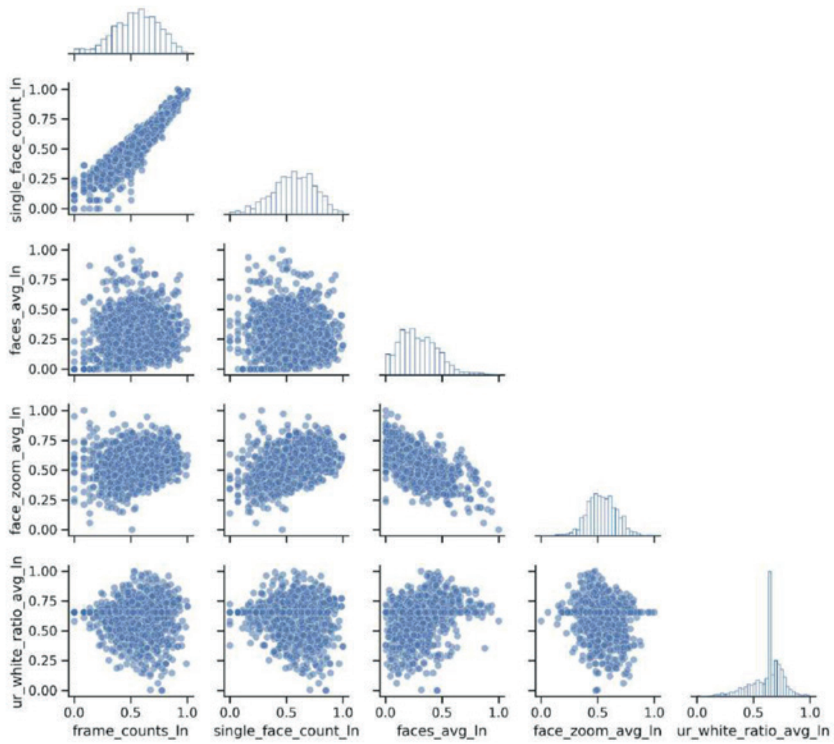


Figure 3. Pair Plot for Characters' Log-Transformed and Min-Max Scaled Visual Metrics

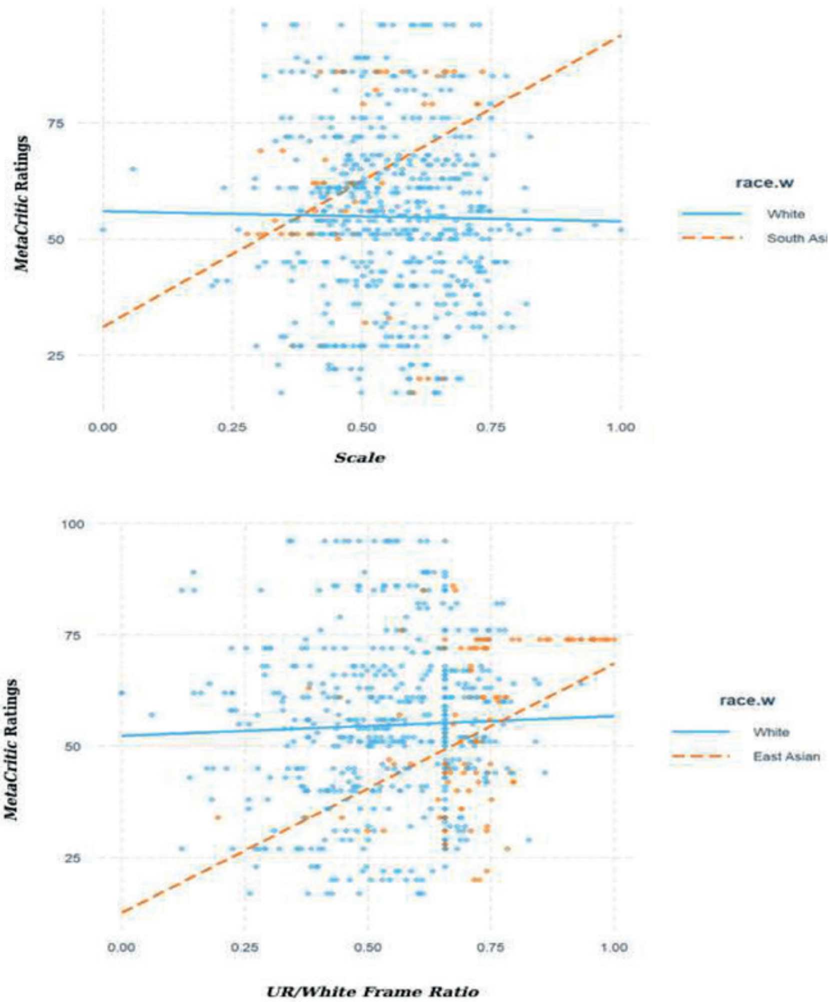
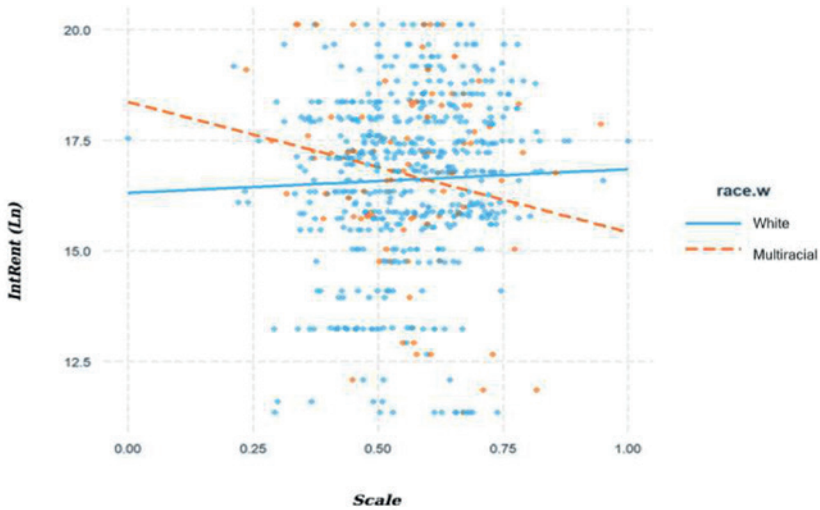


Figure 4. Interactions Between Races (South Asian, East Asian, and White) and Visual Features (Scale and UR/White Frame Ratio) to Predict MetaCritic Ratings.



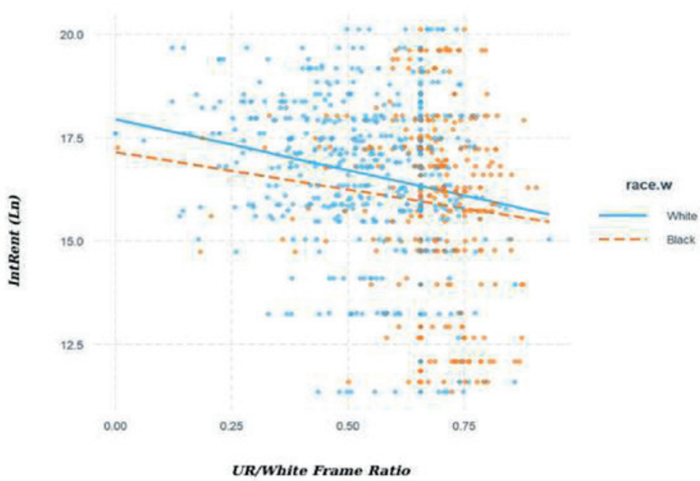


Figure 5. Interactions Between Races (Black, Multiracial, and White) and Visual Features (Scale and UR/White Frame Ratio) to Predict International Rent.

Citation Diversity Statement

Recent work in several fields of science has identified a bias in citation practices such that papers from women and other minority scholars are under-cited relative to the number of such papers in the field. Here we sought to proactively consider choosing references that reflect the diversity

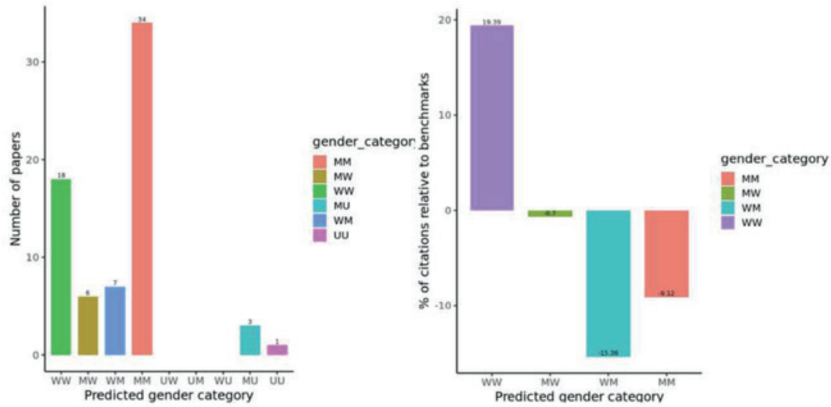


Figure 6. Gender Distribution Across Citations in Manuscript and Benchmark References
Note. **A**) Frequency counts of citations in the paper. **B**) Citations in paper relative to benchmarks

of the field in thought, form of contribution, gender, race, ethnicity, and other factors. First, we obtained the predicted gender of the first and last author of each reference by using databases that store the probability of a first name being carried by a woman. By this measure (and excluding self-citations to the first and last authors of our current paper), our references contain 26.29% woman(first)/woman(last), 10.68% man/woman, 10.49% woman/man, and 52.54% man/man. This method is limited in that a) names, pronouns, and social media profiles used to construct the databases may not, in every case, be indicative of gender identity and b) it cannot account for intersex, non-binary, or transgender people. Second, we obtained and predicted racial/ethnic category of the first and last author of each reference by databases that store the probability of a first and last name being carried by an author of color (7,8). By this measure (and excluding self-citations), our references contain 17.61% author of color (first)/author of color(last), 15.16% white author/author of color, 15.73% author of color/white author, and 51.5% white author/white author. This method is limited in that a) names and Florida Voter Data to make the predictions may not be indicative of racial/ethnic identity, and b) it cannot account for Indigenous and mixed-race authors, or those who may face differential biases due to the ambiguous racialization or ethnicization of their names. We look forward to future work that could help us to better understand how to support equitable practices in science.

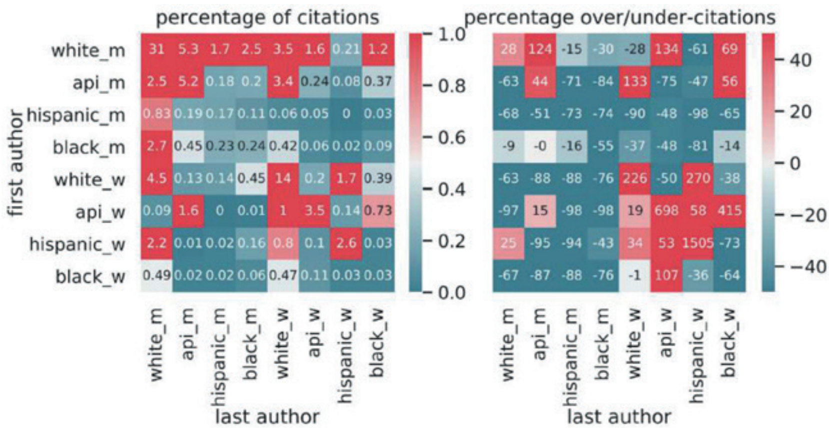


Figure 7. Gender and Race Distribution Across Citations in Manuscript and Benchmark

Notes

1. This conversion was conducted solely for academic purposes (fair use) and the extracted video content can only be accessed by the authors of this paper.
2. Intraclass correlations across models for character network and cinematographic measures varied between 0.15 and 0.30 indicating a data dependency across levels of analysis with about 15-30% of the measures' variance explained at the movie level (between movies).
3. Both dependent performance variables (MetaCritic and IntRent) are movie level measures and do not vary within movies (across characters). Hence, fitting MLMs is not possible.
4. As the focus of this manuscript is not on fitting sophisticated financial performance models, we did not control for production budget, distribution, advertising spend, etc. in our models, which would increase the fit of our models substantially and may even change the size and direction of our race coefficients. For more sophisticated financial performance models including gender and race as predictors, we refer to Smith et al. (2020).
5. A citation diversity statement is provided in the *SI*.

References[5]

Abraham, L. & Appiah, O. (2006). Framing news stories: The role of visual imagery in priming racial stereotypes. *Howard Journal of Communications*, 17(3), 183–203. https://doi.org/10.1080/10646170600829584

Agarwal, A., Zheng, J., Kamath, S., Balasubramanian, S., & Dey, S. A. (2015). Key female characters in film have more to talk about besides men: Automating the Bechdel test. *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 830-840. https://doi.org/10.3115/v1/n15-1084

AII. (2020). *Annenberg Inclusion Initiative*. https://annenberg.usc.edu/research/aii

Atwell Seate, A. & Mastro, D. (2016). Media's influence on immigration attitudes: An intergroup threat theory approach. *Communication Monographs*, 83(2), 194–213. https://doi.org/10.1080/03637751.2015.1068433

Bamman, D., O'Connor, B., & Smith, N.A. (2013). Learning latent personas of film characters *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics,* 1, 352–361. https://aclanthology.org/P13-1035

Bollobás, B. (2001). Random Graphs. Cambridge University Press. https://doi.org/10.1017/cbo9780511814068

Bechdel, A. (1985). The Rule. *Dykes to Watch Out For.*

Bochkovskiy, A., Wang, C.Y., & Liao, H. M. (2020). YOLOv4: optimal speed and accuracy of object detection. *CoRR*, abs/2004.10934. https://arxiv.org/abs/2004.10934v1

Brandes, U. & Erlebach, T. (Eds.). (2005). Network Analysis. *LNCS*, 3418. Springer.

Buolamwini, J. & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Conference on Fairness, Accountability and Transparency*, 77–91. https://proceedings.mlr.press/v81/buolamwini18a.html

Canini, L., Benini, S., & Leonardi, R. (2011). Affective analysis on patterns of shot types in movies. *Proceedings of the 7th International Symposium on Image and Signal Processing and Analysis*, 253-258. https://ieeexplore.ieee.org/abstract/document/6046615

Cao, X. (2013). The effects of facial close-ups and viewers' sex on empathy and intentions to help people in need. *Mass Communication and Society*, 16(2), 161–178. https://doi.org/10.1080/15205436.2012.683928

Celious, A. & Oyserman, D. (2001). Race From the Inside: An Emerging Heterogeneous Race Model. *Journal of Social Issues*, 57(1), 149–165. https://doi.org/10.1111/0022-4537.00206

Croteau, D. & Hoynes, W. (2014). *Media/Society: Industries, images, and audiences*. (5th ed.). SAGE Publications.

Chu, E. & Roy, D. (2017). Audio-visual sentiment analysis for learning emotional arcs in movies. *IEEE International Conference on Data Mining*. 829–834. https://doi.org/10.1109/icdm.2017.100

Dalisay, F. & Tan, A. (2009). Assimilation and Contrast Effects in the Priming of Asian American and African American Stereotypes through TV Exposure. *Journalism & Mass Communication Quarterly*, 86(1), 7–22. https://doi.org/10.1177/107769900908600102

Desilver, D. (2015). Share of counties where whites are a minority has doubled since 1980. *Pew Research Center.* http://www.pewresearch.org/fact-tank/2015/07/01/share-of-counties-where-whites-are-a-minority-has-doubled-since-1980/

Doane, M.A. (2003). The Close-Up: Scale and Detail in the Cinema. *differences: A Journal of Feminist Cultural Studies*. 14(3), 89-111. https://doi.org/10.1215/10407391-14-3-89

Eliashberg, J., Hui, S. K., & Zhang, Z. J. (2007). From storyline to box office: A new approach for green-lighting movie scripts. *Management Science*, 53(6), 881–893. http://dx.doi.org/10.1287/mnsc.1060.0668.

Eliashberg, J., Hui, S. K., & Zhang, Z. (2014). Assessing box office performance using movie scripts: A kernel-based approach. *IEEE Transactions on Knowledge and Data Engineering*, 26(11), 2639–2648. https://doi.org/10.1109/tkde.2014.2306681

Figueroa-Caballero, A., Mastro, D., & Stamps, D. (2019). An Examination of the Effects of Mediated Intragroup and Intergroup Interactions among Latino/a Characters. *Communication Quarterly*, 67(3), 271–290. https://doi.org/10.1080/01463373.2019.1573745

Fu, S., He, H., & Hou, Z.-G. (2014). Learning Race from Face: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(12), 2483–2509. https://doi.org/10.1109/tpami.2014.2321570

Gardner, S. K. & Hughey, M. W. (2017). Still the tragic mulatto? Manufacturing multiracialization in magazine media, 1961–2011. *Ethnic and Racial Studies*, 42(4), 645–665. https://doi.org/10.1080/01419870.2017.1380212

Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. *Computer Vision and Pattern Recognition*, 580–587. https://doi.org/10.1109/cvpr.2014.81

Girshick, R. B. (2015). Fast R-CNN. *CoRR*, abs/1504.08083.

Gonzalez, E. (2014). Engaging the evolving Hispanic consumers: Look at two distinct subgroups. *The Nielsen Company.* Retrieved from http://www.nielsen.com/us/en/insights/news/2014/engaging-the-evolving-hispanic-consu mers.html.

Gleich, D. F. (2015). PageRank Beyond the Web. *SIAM Review*, 57(3), 321–363. https://doi.org/10.1137/140976649

Gopinath, S., Chintagunta, P. K., & Venkataraman, S. (2013). Blogs, advertising, and local-market movie box office performance. *Management Science*, 59(12), 2635–54. https://doi.org/10.1287/mnsc.2013.1732.

Hickey, W., Koeze, E., Dottle, R., & Wezerek, G. (2017). The next bechdel test. *FiveThirtyEight*. https://projects.fivethirtyeight.com/next-bechdel/

Hopp, F. R., Fisher, J. T., & Weber, R. (2020). A Graph-Learning Approach for Detecting Moral Conflict in Movie Scripts. *Media and Communication*, 8(3), 164–179. https://doi.org/10.17645/mac.v8i3.3155

Hornsey, M. J. (2008). Social Identity Theory and Self-categorization Theory: A Historical Review. *Social and Personality Psychology Compass*, 2(1), 204–222. https://doi.org/10.1111/j.1751-9004.2007.00066.x

Hurley, R. J., Jensen, J. J., Weaver, A., & Dixon, T. (2015). Viewer Ethnicity Matters: Black Crime in TV News and Its Impact on Decisions Regarding Public Policy. *Journal of Social Issues*, 71(1), 155–170. https://doi.org/10.1111/josi.12102

Huang, Q., Xiong, Y., Xiong, Y., Zhang, Y., Lin, D. (2018). From trailers to storylines: An efficient way to learn from movies. https://arxiv.org/abs/1806.05341v1

Jones, P., Quinn, E., & Koskinen, J. (2020). Measuring centrality in film narratives using dynamic character interaction networks. *Social Networks*, 63, 21-37. https://doi.org/10.1016/j.socnet.2020.03.003

Jones, S. W. (2008). Tragic No More?: The Reappearance of the Racially Mixed Character. *American Fiction of the 1990s* (Prosser, J. Ed.), 89-103. Routledge.

Kagan, D., Chesney, T., & Fire, M. (2020). Using data science to understand the film industry's gender gap. *Palgrave Communications*, 6(1), 1-16. https://doi.org/10.1057/s41599-020-0436-1.

Kärkkäinen, K. & Joo, J. (2021). FairFace: Face Attribute Dataset for Balanced Race, Gender, and Age for Bias Measurement and Mitigation. *IEEE Winter Conference on Applications of Computer Vision*, 1547-1557. https://doi.org/10.1109/wacv48630.2021.00159

Katz, S. D. (1991). Film directing shot by shot—Visualizing from concept to screen. Michael Wiese Productions.

Kim, M. & Brunn-Bevel, R. J. (2020). Hollywood's Global Expansion and Racialized Film Industry. *Humanity & Society*, 44(1), 37–66. https://doi.org/10.1177/0160597619832045.

Kølvraa, C. (2013). Ideology and the crowd. *Distinktion: Journal of Social Theory*, 14(2), 114–133. https://doi.org/10.1080/1600910x.2012.745811

Lauzen, M. (2018a). It's a man's (celluloid) world: Portrayals of female characters in the 100 top films of 2017. *Center for the Study of Women in Television and Film*. https://womenintvfilm.sdsu.edu/wp-content/uploads/2019/02/2018_Its_a_Mans_Celluloid_World_Report.pdf

Lauzen, M. (2018b). Boxed in 2017–18: Women on screen and behind the scenes in television. *Center for the Study of Women in Television and Film*. https://womenintvfilm.sdsu.edu/wp-content/uploads/2018/09/2017-18_Boxed_In_Report.pdf

Liu, Z., Luo, P., Wang, X., & Tang, X. (2015). Deep Learning Face Attributes in the Wild. *2015 IEEE International Conference on Computer Vision* (*ICCV*). https://doi.org/10.1109/iccv.2015.425

LoBue, V. & Thrasher, C. (2015). The Child Affective Facial Expression (CAFE) set: validity and reliability from untrained adults. *Frontiers in Psychology*, 5. https://doi.org/10.3389/fpsyg.2014.01532

Mastro, D. E. (2003). A social identity approach to understanding the impact of television messages. *Communication Monographs*, 70(2), 98–113. https://doi.org/10.1080/0363775032000133764

Mastro, D., Behm-Morawitz, E., & Ortiz, M. (2007). The cultivation of social perceptions of Latinos: A mental models approach. *Media Psychology*, 9(2), 347-365. https://doi.org/10.1080/15213260701286106.

Mastro, D. E. & Kopacz, M. A. (2006). Media representations of race, prototypicality, and policy reasoning: An application of self-categorization theory. *Journal of Broadcasting & Electronic Media*, 50(2), 305-322. https://doi.org/10.1207/s15506878jobem5002_8.

Mastro, D. & Tukachinsky, R. (2011). The Influence of Exemplar Versus Prototype-Based Media Primes on Racial/Ethnic Evaluations. *Journal of Communication*, 61(5), 916–937. https://doi.org/10.1111/j.1460-2466.2011.01587.x

McClintock, P. (2020). 2019 global box office revenue hit record $42.5B despite 4 percent dip in U.S. *Billboard*. https://www.billboard.com/articles/news/8547827/2019-global-box-office-revenue-hit-record-425b-despite-4-percent-dip-in-us

Millerson, G. (1961). *The technique of television production*. Focal Press.

Molina-Guzmán, I. (2016). #OscarsSoWhite: how Stuart Hall explains why nothing changes in Hollywood and everything is changing. *Critical Studies in Media Communication*, 33(5), 438–454. https://doi.org/10.1080/15295036.2016.1227864.

MPAA. (2014). Theatrical market statistics. *Motion Picture Association of America*. https://www.motionpictures.org/wp-content/uploads/2015/03/MPAA-Theatrical-Market-Statistics-2014.pdf

MPAA. (2019). Theme report. *Motion Picture Association of America*. https://www.motionpictures.org/wp-content/uploads/2020/03/MPA-THEME-2019.pdf

Mutz, D. C. (2006). Effects of "In-Your-Face" Television Discourse on Perceptions of a Legitimate Opposition. *Hendricks Symposium–Department of Political Science*. http://digitalcommons.unl.edu/politicalsciencehendricks/16.

Noam, E. (2009). *Media ownership and concentration in America*. Oxford University Press.

Ortiz, M. & Harwood, J. (2007). A social cognitive theory approach to the effects of mediated intergroup contact on intergroup attitudes. *Journal of Broadcasting & Electronic Media*, 51(4), 615–631. https://doi.org/10.1080/08838150701626487

Park, S. B., Kim, Y. W., Uddin, M. N., Jo, G. S. (2009). Character-net: Character network analysis from video. *Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology*, 1, 305–308. https://doi.org/10.1109/WI-IAT.2009.54

Phillips, P. J., Wechsler, H., Huang, J., & Rauss, P. J. (1998). The FERET database and evaluation procedure for face-recognition algorithms. *Image and Vision Computing*, 16(5), 295–306. https://doi.org/10.1016/s0262-8856(97)00070-x

Raji, I. D. & Buolamwini, J. (2019). Actionable auditing: Investigating the impact of publicly naming biased performance results of commercial AI products. *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society,* 429–435. https://doi.org/10.1145/3306618.3314244

Ramakrishna, A., Martínez, V.R., Malandrakis, N., Singla, K., & Narayanan, S. (2017). Linguistic analysis of differences in portrayal of movie characters. *Proceedings of the 55th annual meeting of the association for computational linguistics*, 1, 1669–1678. https://aclanthology.org/P17-1153

Ramasubramanian, S. (2010). Television viewing, racial attitudes, and policy preferences: Exploring the role of social identity and intergroup emotions in influencing support for affirmative action. *Communication Monographs*, 77(1), 102–120. https://doi.org/10.1080/03637750903514300

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. *2016 IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR*),779-788. https://doi.org/10.1109/cvpr.2016.91

Redmon, J. & Farhadi, A. (2017). YOLO9000: Better, Faster, Stronger. *2017 IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR*), 6517-6525. https://doi.org/10.1109/cvpr.2017.690

Redmon, J. & Farhadi, A. (2018). YOLOv3: An Incremental Improvement. *ArXiv, abs/1804.02767*. https://arxiv.org/pdf/1804.02767.pdf

Ren, S., He, K., Girshick, R., & Sun., J. (2015) Faster r-cnn: Towards real-time object detection with region proposal networks. *ArXiv, abs/1506.01497*. https://arxiv.org/abs/1506.01497v3

Ricanek, K. & Tesafaye, T. (2006). MORPH: a longitudinal image database of normal adult age-progression. *7th International Conference on Automatic Face and Gesture Recognition* (*FGR06*), 341-345. https://doi.org/10.1109/FGR.2006.78

Rooney, B. & Bálint, K. E. (2018). Watching more closely: Shot scale affects film viewers' Theory of Mind tendency but not ability. *Frontiers in Psychology*, 8. https://doi.org/10.3389/fpsyg.2017.02349.

Simoes, G. S., Wehrmann, J., Barros, R. C., Ruiz, D. D. (2016). Movie genre classification with convolutional neural networks. *2016 International Joint Conference on Neural Networks* (*IJCNN*), 259–266. https://doi.org/10.1109/ijcnn.2016.7727207

Skowron, M., Trapp, M., Payr, S., & Trappl, R. (2016). Automatic identification of character types from film dialogs. *Applied Artificial Intelligence*, 30(10), 942-973. https://doi.org/10.1080/08839514.2017.1289311

Smith S. & Choueiti, M. (2010). Gender disparity on screen and behind the camera in family films. Executive report. *USC Annenberg Inclusion Initiative*. https://

seejane.org/wp-content/uploads/full-study-gender-disparity-in-family-films-v2.pdf

Smith, S. L. & Choueiti, M. (2011). Gender inequality in cinematic content? A look at females on screen & behind-the-camera in top-grossing 2008 films. *USC Annenberg Inclusion Initiative*. https://annenberg.usc.edu/sites/default/files/MDSCI_Gender_Inequality_in_TopGrossing_Films.pdf

Smith S., Pieper K., & Choueiti, M. (2017). Inclusion in the director's chair? Gender, race, & age of film directors across 1,000 films from 2007–2016. *Media, Diversity, & Social Change Initiative*. https://annenberg.usc.edu/sites/default/files/2017/04/06/MDSCI_Inclusion%20_in_the_Directors_Chair.pdf

Smith, S. L., Weber, R., Choueiti, M., Pieper, K., Case, A., Yao, K., & Lee, C. (2020). The ticket to inclusion: Gender and race/ethnicity of leads and financial performance across 1,200 popular films. *USC Annenberg Inclusion Initiative*. https://assets.uscannenberg.org/docs/aii-2020-02-05-ticket-to-inclusion.pdf

Spangler, T. (2020). Netflix adds 8.8 million subscribers in Q4, cites competition for lower U.S. gains. *Variety*. https://variety.com/2020/digital/news/netflix-q4-2019-earnings-results-1203474435/.

Stimpert, J. L., Laux, J. A., Marino, C., & Gleason, G. (2008). Factors influencing motion picture success: Empirical review and update, *Journal of Business and Economics Research*, 6(11), 39–52. https://doi.org/10.19030/jber.v6i11.2488.

Statista. (2020). Market share of leading film studios in North America in 2019. *Statista - The Statistics Portal*. https://www.statista.com/statistics/187171/market-share-of-film-studios-in-north-america-2010/

Tajfel, H. & Turner, J. C. (1986). The social identity theory of intergroup behaviour. S. Worchel & W. G. Austin (Eds.), *Psychology of Intergroup Relations*, 7–24. Nelson-Hall.

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Blackwell.

Tamborini, R. (2013). Model of intuitive morality and exemplars. R. Tamborini (Ed.), *Media and the moral mind*, 43–74. Routledge.

Tamborini, R., & Weber, R. (2020). Advancing the model of intuitive morality and exemplars. K. Floyd & R. Weber (Eds.), *The handbook of communication science and biology*, 456–469. Routledge.

Tiemens, R. K. (2009). Some relationships of camera angle to communicator credibility. *Journal of Broadcasting*, 14(4), 483–490. https://doi.org/10.1080/08838157009363614.

Tratner, M. (2003). Working the Crowd: Movies and Mass Politics. *Criticism*, 45(1), Article 4. https://digitalcommons.wayne.edu/criticism/vol45/iss1/4

Trepte, S. (2006). Social identity theory. J. Bryant & P. Vorderer (Eds.), *Psychology of Entertainment*, 255–271, Routledge.

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Blackwell.

Turner, J. & Oakes, P. (1986). The significance of the social identity concept for social psychology with reference to individualism, interactionism and social influence. *British Journal of Social Psychology*, 25(3), 237–252. https://doi.org/10.1111/j.2044-8309.1986.tb00732.x

UNIC. (2017). UNIC annual report 2018. *Union Internationale des Cinémas*. https://www.unic-cinemas.org/fileadmin/user_upload/wordpress-uploads/2017/06/UNIC_ AR2018_online.pdf

University, SDS (2017). Women remain underrepresented in Hollywood, study shows. *Phys.org*. https://phys.org/news/2017-09-women-underrepresented-hollywood.html

U.S. Census (2019). 2019 Population Estimates by Age, Sex, Race and Hispanic Origin. *U.S. Census Bureau*. https://www.census.gov/newsroom/press-kits/2020/population-estimates-detailed.html

Weng, C. Y., Chu, W. T., & Wu, J. L. (2009). Rolenet: Movie analysis from the perspective of social networks. IEEE Transactions on Multimedia, 11(2), 256-271. https://doi.org/10.1109/tmm.2008.2009684

Wood, J. T. (1994). Gendered lives: Communication, Gender, and Culture, 231-244. Wadsworth Publishing.

Zhou, H., Hermans, T., Karandikar, A.V., Rehg, J. M. (2010). Movie genre classification via scene categorization. *Proceedings of the 18th ACM international conference on Multimedia*, 747–750. https://doi.org/10.1145/1873951.1874068

Zhu, Y., Kiros, R., Zemel, R., Salakhutdinov, R., Urtasun, R., Torralba, A., Fidler, S. (2015). Aligning books and movies: Towards story-like visual explanations by watching movies and reading books. *Proceedings of the IEEE international conference on computer vision* (*ICCV*), 19–27. https://doi.org/10.1109/iccv.2015.11

Correspondence should be addressed to René Weber (renew@comm.ucsb.edu); University of California Santa Barbara, Department of Communication - Media Neuroscience Lab, Santa Barbara, CA 93106-4020; https://medianeuroscience.org