

BIAS AND INCLUSION IN ADVERTISING: AN ANALYSIS OF 2018 CANNES LIONS FILM CRAFT ADS



Geena Davis Institute  on Gender in Media
If she can see it, she can be it.™



EXECUTIVE SUMMARY

This study examines representations of gender, race/ethnicity, sexual orientation (LGBTQ), and ability (with a focus on people with disabilities) among characters in Cannes Lions ads from 2018. Furthermore, we compare findings from 2018 to gender representations in Cannes Lions Film Craft advertisements from the previous decade (2006-2017). This executive summary presents our major findings.

GENDER

- Male characters outnumber female characters in 2018 ads (59.8% men, 39.9% women), but by a smaller margin than in the previous decade.
- Male characters are twice as likely as women to be shown working in 2018 ads (26% compared with 11%), an even larger discrepancy than that found in ads from 2006 to 2017.
- Overall, gender differences within stereotypical activities and settings in 2018 ads showed marked improvement when compared with ads from the previous decade.
- Male characters in 2018 ads are twice as likely as female characters to be shown as having an occupation (25.6% compared with 13.8%), comparable to findings from the previous decade.
- Male characters in 2018 ads are twice as likely as female characters to be portrayed as leaders (16.4% compared with 8.3%), comparable to findings from the previous decade.
- Although female characters remain sexualized in 2018 ads, there is marked improvement compared with ads from 2006 to 2017.

- Male characters in 2018 ads are 1.5 times more likely than female characters to be shown performing physical comedy (13.3% compared with 8.5%), and twice as likely to be shown performing verbal comedy (9.8% compared with 4.1%).

RACE

- People of color make up 43.1% of characters in 2018 ads.
- White characters are more likely than characters of color to be depicted as having an occupation (22.9% compared with 17.9%).

SEXUAL ORIENTATION

- LGBTQ characters are underrepresented in 2018 ads, making up only 1.9% of characters despite the fact that 10% of people identify as LGBTQIA globally.

DISABILITY

- Characters with disabilities are underrepresented in 2018 ads, making up only 0.8% of characters. This is in sharp contrast to the 15% of people globally with some form of cognitive or physical disability.

We begin this report with a brief description of our research methodology. Next, we present a profile of our 2018 sample. In the third section, we summarize our findings from a content analysis of 133 Cannes Lions Film Craft advertisements from 2018. In the fourth section, we compare findings from 2018 to 2,137 Cannes Lions Film Craft advertisements from 2006 to 2017.

METHODOLOGY

Content analysis is a research method ideal for systematically analyzing the content of communications, such as advertisements. Content analysis is used by social scientists to quantify and examine the presence of certain themes or concepts. We used automated coding (GD-IQ) and hand coding to assess representations of gender, race, sexuality, and ability in Cannes Lions Film Craft advertisements. We analyzed 133 English-speaking Cannes Lions Film Craft advertisements from 2018. Ads originated from 23 countries, including Argentina, Australia, Brazil, Canada, Chile, Czech Republic, Denmark, France, Germany, India, Ireland, Italy, Japan, Luxembourg, Mexico, New Zealand, Norway, South Africa, Spain, Sweden, Thailand, the UK, and USA. Furthermore, we compare these findings to 2,137 Cannes Lions Film Craft advertisements from 2006 to 2017. English-speaking ads from 2006 to 2017 originate from Australia, Canada, New Zealand, the UK, and USA. All reported differences are statistically significant at the .05 level unless otherwise indicated. The unit of analysis for this study is characters in Cannes Lions Film Craft advertisements.

For automated coding, we employed the Geena Davis Inclusion Quotient (GD-IQ), a ground-breaking software tool developed by the Geena Davis Institute on Gender in Media at Mount Saint Mary's University to analyze audio and video media content. Funded by Google.org, the GD-IQ incorporates machine learning technology as well as the University of Southern California's audio-visual processing technologies, and is the only software tool in existence with the ability to measure screen and speaking time through the use of automation. This revolutionary tool was co-developed by the Institute and led by Dr. Shrikanth (Shri) Narayanan and his team of researchers at the University of Southern California's Signal Analysis and Interpretation Laboratory (SAIL).

For hand-coding, a team of seven researchers systematically evaluated the 1,115 characters in Cannes Lions Film Craft advertisements. Prior to initiating the work, the research team engaged in a total of 29 hours of training and codebook development. The team also performed a test to measure inter-coder reliability. Initial inter-coder reliability tests were performed on 10 ads to ensure that members of the research team reached agreement on evaluations. Inter-coder reliability was achieved in terms of both absolute agreement and Cohen's Kappa measures.

SAMPLE OVERVIEW

TABLE ONE:

Gender in 2018 Cannes Lions Film Craft Ads

DEMOGRAPHIC BREAKDOWN	PERCENT OF SAMPLE
MALE	59.8%
FEMALE	39.9%
TRANSGENDER	0.4%

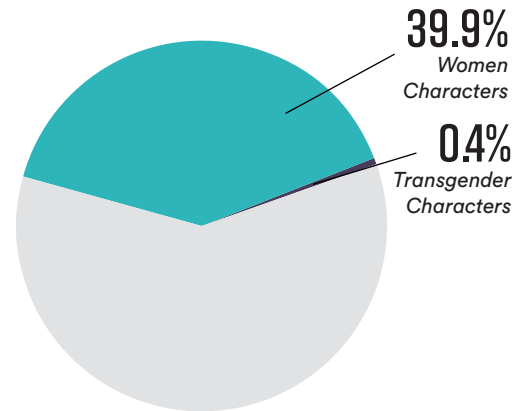
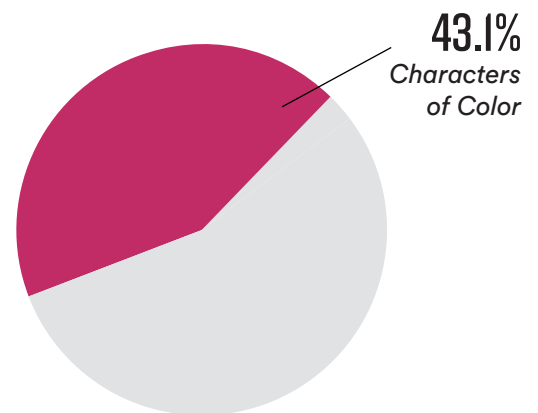


TABLE TWO:

Race in 2018 Cannes Lions Film Craft Ads

DEMOGRAPHIC BREAKDOWN	PERCENT OF SAMPLE
WHITE	54.5%
LATINO	5.8%
BLACK	17.1%
NATIVE AMERICAN/INDIGENOUS	0.9%
ASIAN	16.9%
MIDDLE EASTERN	1.7%
MIXED	0.7%



Summary: White vs. People of Color

WHITE	54.5%
PEOPLE OF COLOR	43.1%

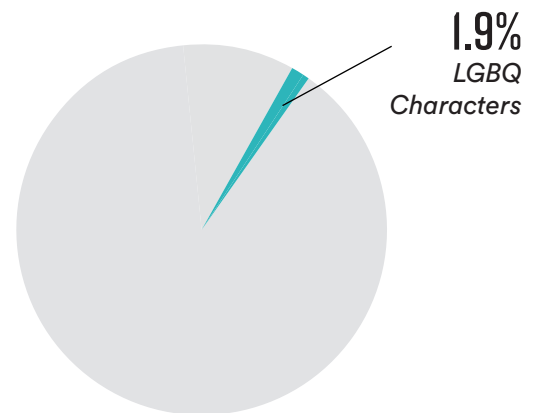


TABLE THREE:

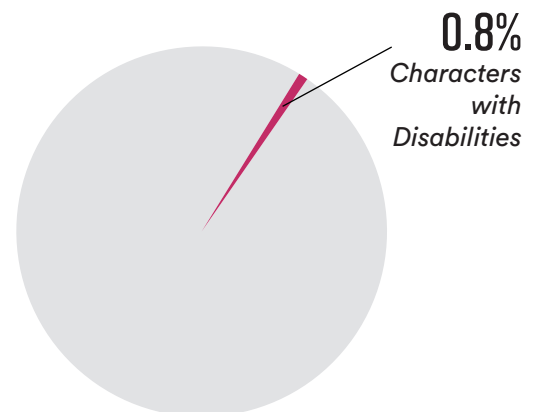
Sexual Orientation in 2018 Cannes Lions Film Craft Ads

DEMOGRAPHIC BREAKDOWN	PERCENT OF SAMPLE
HETEROSEXUAL	98.1%
GAY/LESBIAN	1.3%
BISEXUAL/QUEER	0.6%

TABLE FOUR:

Ability Status in 2018 Cannes Lions Film Craft Ads

DEMOGRAPHIC BREAKDOWN	PERCENT OF SAMPLE
ABLE-BODIED	99.2%
PHYSICALLY DISABLED	0.7%
COGNITIVELY DISABLED	0.3%
COMMUNICATION DISABLED	0.1%



CANNES LION FILM CRAFT AD ANALYSIS

Our analysis and findings for 2018 ads are broken down into four main sections: representations of gender, race, sexual orientation, and ability status. All reported differences are statistically significant at the .05 level unless otherwise indicated.

GENDER

In this section, we analyze the representation of gender by examining prominence, screen time, speaking time, activities, settings, portrayals of occupation and leadership, personal attributes, and sexualization.

Prominence

Research has consistently found that women are underrepresented on screen, despite making up 51% of the U.S. population.¹ In advertisements from 2018, male characters outnumbered female characters (59.8% compared with 39.9%), but by a smaller margin than found in previous studies.²

Screen Time & Speaking Time

Using the GD-IQ, we analyzed the amount of time characters appeared on screen (screen time) and speaking time by gender. Results indicate that women appear on screen for 38% of ad duration and account for 39.2% of speaking time.

Age

Female characters are 2.5 times more likely than male characters to be shown in their late teens (8.9% compared with 3.5%). In contrast, male characters are 2 times more likely than female characters to be shown in their 50s (12.1% compared with 5.3%). No significant gender differences emerged among other age groups.

Overall, only 47.5% of female characters are shown over the age of 30 whereas 60.1% of male characters are portrayed over the age of 30.

ONLY 47.5% OF WOMEN ARE SHOWN OVER THE AGE OF THIRTY (COMPARED TO 60.1% OF MEN)

Activity

We examined the types of activities that men and women are shown engaging in. Across all 2018 ads, men and women are equally likely to be shown shopping, cleaning, cooking, socializing, eating or drinking, working out, and using technology. In contrast, male characters are three times as likely to be shown driving (7.2% compared to 2.5%), and more than twice as likely to be shown working (26% compared with 11%).

Setting

We also examined the settings in which men and women are presented. Men and women are equally likely to be shown in the kitchen, office, car, store, living room, restaurant or bar, gym, bathroom, and at a sporting event.

In contrast, men are more likely to be shown outdoors (45.7% compared with 36.5%), and women are almost twice as likely to be shown in the bedroom (6.2% compared with 3.5%).

Work and Leadership

In addition, we examined the representation of men and women with regards to occupation and leadership. Here, several gender differences emerged. First, men are twice as likely as women to be shown with an occupation (25.6% compared with 13.8%). Second, several occupational categories show a gender imbalance. Men are more likely than women to be depicted in the service industry (4.9% compared with 1.5%), sales (2.6% compared with 0.8%), military positions (2.5% compared with 0.3%), and criminal occupations (2.1% compared with 0%). In contrast, men and women are equally likely to be shown in management, professional positions, and administrative positions. And lastly, men are twice as likely as women to be shown in a leadership position (16.4% compared with 8.3%).

Personal Attributes

We examined the representation of men and women with regards to intelligence and humor. Men are more likely to be depicted as less intelligent (4.1% compared to 1.2%), and women are more likely to be depicted as smart (13.8% compared with 9.9%). Men are also more likely to be depicted performing physical comedy (13.3% compared with 4.1%).

Sexualization

We also examined the sexualization of women in 2018 ads. Sexualization occurs when a person's value is primarily derived from their sexual appeal, when physical beauty is equated with sexiness, when sexuality is inappropriately imposed on someone, or when a person is sexually objectified.³ Sexual objectification refers to the process of treating someone like a sexual object, such as by focusing in on sexualized parts of someone's body. We measured sexualization and objectification of characters through revealing clothing, degree of nudity, visual, and verbal objectification. Visual objectification occurs when the camera focuses on specific body parts, pans up and the down the character's body, or when slow motion is used to accentuate the body in a sexual manner. In contrast, verbal objectification includes cat calling and making comments about a person's physical appearance to others. Gender differences emerged among several dimensions.

Women are more than twice as likely as men to be depicted in revealing clothing (7.8% compared with 2.8%), and three times as likely to be verbally objectified (3.7% compared with 0.9%). No gender differences emerged with regards to nudity or visual objectification.

WOMEN ARE MORE THAN TWICE AS LIKELY TO BE SHOWN IN REVEALING CLOTHING (7.8% COMPARED TO 2.8%)

RACE

In this section, we analyze the representation of race in prominence, activities, settings,

portrayals of occupation and leadership, and personal attributes.

Prominence

Overall, representation of people of color in 2018 advertisements exceeded that of U.S. population numbers (43.1% people of color compared with 38% of U.S. population).⁴ Characters of color are equally likely as white characters to be featured in both speaking roles and visually prominent roles.

Activity

Across all 2018 ads, white characters and characters of color are equally likely to be shown shopping, driving, cleaning, cooking, working, socializing, working out, and using technology. The only racial difference that emerged was in regard to characters shown eating and drinking. White characters are more likely than characters of color to be shown eating and drinking (13.4% compared with 9.2%).

Setting

Characters of color and white characters were equally likely to be shown in the kitchen, in a car, store, living room, restaurant or bar, bathroom, and at sporting events.

In contrast, white characters are more likely to be shown in an office (6.3% compared with 2.9%) and characters of color are more likely to be shown outdoors (49.4% compared with 36.6%), in a gym (1.5% compared with 0.2%), and in a bedroom (6.5% compared with 3.8%).

Work and Leadership

White characters are more likely than characters of color to be depicted as having an occupation (22.9% compared with 17.9%). White characters are more likely to be shown in sales positions (2.7% compared with 0.9%), while characters of color are more likely to be shown in military positions (3.0% compared with 0.2%). No racial difference emerged with regards to leadership.

CHARACTERS OF COLOR ARE LESS LIKELY TO BE SHOWN HAVING AN OCCUPATION (17.9% COMPARED WITH 22.9%)

Personal Attributes

When it came to intelligence, white characters are more likely than characters of color to be depicted as less intelligent (4.5% compared with 1.3%). Characters of color are more likely to be depicted as visually funny (e.g., the audience is asked to laugh at the character's physical appearance) compared with white characters (8.5% compared with 4.8%). Characters of color are also less likely to perform physical comedy (e.g., falling down, slapstick comedy) than white characters (8.8% compared with 13.2%).

SEXUAL ORIENTATION

In this section, we analyze the representation of LGBTQ characters in prominence, activities, settings, portrayals of occupation and leadership, and personal attributes.

Prominence

When it comes to sexual orientation in 2018 Cannes Lions ads, gay and bi/queer characters are virtually nonexistent despite the fact that 10.0% of people globally identify as LGBTQIA.⁵ Of the 1008 characters in 2018 ads with a discernible sexual orientation, only 19 (1.9%) are LGBTQ. Of these characters, 63.2% are shown in speaking roles and 36.8% in visually prominent roles.

ONLY 1.9% OF CHARACTERS WITH A DISCERNIBLE SEXUAL ORIENTATION ARE LGBTQ

Activity and Setting

Because of the small number of LGBTQ characters across multiple activity and setting categories, we cannot make appropriate comparisons with regards to activity or setting.

Work and Leadership

Straight characters and LGBTQ characters were equally likely to be depicted as having an occupation. Because of the small number of LGBTQ characters, we cannot make appropriate comparisons within occupational categories.

LGBTQ characters are more likely than heterosexual characters to be depicted as leaders (36.8% compared with 14%).

Personal Attributes

No differences emerged with regard to intelligence or humor.

ABILITY

In this section, we analyze the representation of those with disabilities with regards to prominence, activities, settings, portrayals of occupation and leadership, and personal attributes.

Prominence

15.0% of people globally have some form of cognitive or physical disability,⁶ but people with disabilities make up 0.8% of characters in 2018 Cannes Lions ads. Across all 2018 ads, only 9 characters are depicted with a disability. Of these characters, there are 8 instances of physical disabilities, 3 instances of cognitive disabilities, and 1 instance of communication impairments.

ONLY 0.8% OF CHARACTERS IN 2018 ADS ARE SHOWN WITH A DISABILITY

Activity and Setting

Because of the small number of characters with disabilities across multiple activity and setting categories, we cannot make appropriate comparisons.

Work and Leadership

No characters with disabilities are shown as having an occupation or in leadership positions.

Personal Attributes

Three characters with disabilities are depicted as smart, and one is depicted as funny. Because of the small number of characters with disabilities in these categories, we cannot make appropriate comparisons to characters without disabilities.

COMPARATIVE ANALYSIS

In this section, we compare gender representation in 2018 Cannes Lions Film Craft advertisements with advertisements from 2006 to 2017 with regards to prominence, screen time, speaking time, activities, settings, work and leadership, personal attributes, and sexualization.

Prominence

Gender representation in 2018 showed improvement from the previous decade (2006-2017) of advertisements, in which women accounted for an average of 34.9% of characters and men represented an average of 65.1% of characters. Representation of women peaked at 40.2% in 2014, followed by 39.9% in 2018, and 38.8% in 2015. Figure 1 presents the gender breakdown in our 2018 dataset side-by-side the previous decade.

Screen Time and Speaking Time

Advertisements in 2018 showed a significant improvement from previous years. Data from 2006 to 2017 indicate that men were 5 times as likely to be depicted alone compared with women (25% versus 5%), and that men received 4 times as much screen time as women (20.1% female screen time). Similarly, ads with male voices were far more common than ads with female voices (18% compared with 3%), and men spoke 7 times more than women (12.5% female speaking time).

Age

Advertisements from 2018 showed a similar age distribution as ads from the previous decade, with female characters being younger than male characters. In the previous decade, female characters were 2 times more likely than male characters to be shown in their 20s (40.7% compared with 18.8%). In contrast, male characters were more likely than female characters to be shown in their 40s (16.8% compared with 9.8%) and above age 60 (8.1% compared with 2%). Table 5 presents the age breakdown our 2018 dataset side-by-side with the previous decade.

FIGURE ONE:

Gender Across Time in Cannes Lions Film Craft Ads

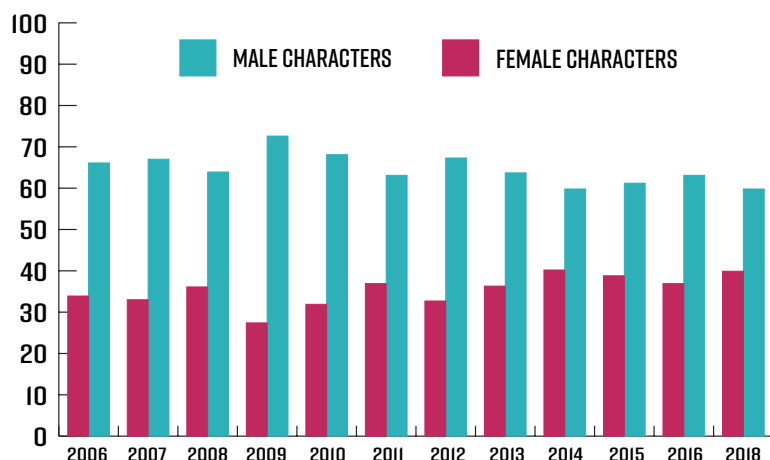


TABLE FIVE:

Age Across Time in Cannes Lions Film Craft Ads

AGE	2006-2017		2018	
	WOMEN	MEN	WOMEN	MEN
CHILD (1-12)	7.3%	7.2%	10.6%	9.7%
TWEEN (13-14)	3.3%	4.7%	3.4%	2.6%
TEEN (15-19)	10.2%	8.1%	8.9%	3.5%
20'S (20-29)	40.7%	18.8%	26.1%	21.2%
30'S (30-39)	22.0%	26.4%	20.6%	24.2%
40'S (40-49)	9.8%	16.8%	15.6%	16.1%
50'S (50-59)	3.3%	6.7%	5.3%	12.1%
60 AND OLDER	2.0%	8.1%	6.0%	7.7%

Activity

Advertisements in 2018 showed fewer gender differences than ads from the previous decade. Across advertisements from 2006 to 2017, women were more likely than men to be shown shopping (3.5% compared with 1.7%), cleaning (0.8% compared with 0.2%), cooking (1.1% compared with 0.5%), and socializing (27.7% compared with 21.2%). In contrast, men were more likely than women to be shown working (29% compared with 20.6%) and using technology (65.1% compared with 34.9%). Overall, this contributed to a pattern in which women were typically seen as engaging in domestic and leisure activities.

Setting

Advertisements in 2018 also showed fewer gender differences with regards to setting compared with ads from 2006 to 2017. In advertisements from the previous decade, women were more likely than men to be shown in the kitchen (5% compared with 3.3%), in the living room (11% compared with 7.9%), and in the bedroom (4.5% compared with 3%). Men were more likely than women to be shown in active settings, such as outdoors (29.5% compared with 25.8%), in the gym (1.9% compared with 1%), and at a sporting event (4.4% compared with 2.3%). Together, these images reinforce the notion that women are domestic caretakers, while men are more active and physical.

Work and Leadership

Portrayals of work and leadership from 2006 to 2017 showed similar patterns as 2018. Men

were more likely than women to be shown as having an occupation (38.5% compared with 24.5%). Within occupational categories, men were more likely to be shown in management (3.5% compared with 1.3%), professional occupations (16% compared with 8.1%), service (7.3% compared with 5.6%), military positions (1.5% compared with 0.1%), and criminal positions (1.0% compared with 0.2%). Furthermore, men were more likely than women to be shown as leaders (14.2% compared with 9.2%).

Personal Attributes

Findings from 2018 indicate some progress when it comes to depictions of intelligence. Across ads from 2006 to 2017, men were twice as likely as women to be shown as smart (0.8% compared with 0.4%).

Representations of humor remained similar, with men being more likely to be shown performing physical comedy (9.2% compared with 3.9%) and verbal comedy (8.4% compared with 6.7%).

Sexualization

Overall, advertisements in 2018 showed some improvement over ads from the previous decade. Compared with men, women in ads from 2006 to 2017 were six times as likely to be shown in revealing clothing (12.6% compared with 2.1%), three times as likely to be visually objectified (6.8% compared with 2%) and verbally objectified (1.6% compared with 0.5%), and more than twice as likely to be shown in a state of partial nudity (7.9% compared with 3.3%).

CONCLUSION

This study systematically examined representations of gender, race, sexual orientation, and ability status in 2018 Cannes Lions Film Craft advertisements. Furthermore, we compared gender representation in ads from 2018 to results from a previous study of Cannes Lions Film Craft Advertisements from 2006 to 2017.

The findings highlight various areas with positive trends and suggest clear areas for improvement. With regards to positive trends, 2018 ads showed a smaller difference between the number of male and female characters than that found in ads from 2006 to 2017, in which men outnumber women two-to-one. Female screen time and speaking time also showed improvement but failed to reach parity. Furthermore, some improvement was seen over time with regards to characters shown in gender stereotypical activities and settings.

RECOMMENDATIONS

To further improve diversity in advertisements, we recommend the following considerations:

PRIORITIZE WOMEN WHEN CREATING NEW AD CHARACTERS TO CONTINUE CLOSING THE GENDER GAP

WHEN CREATING FEMALE CHARACTERS, FOCUS ON CHALLENGING GENDER STEREOTYPES REGARDING OCCUPATIONAL STATUS, LEADERSHIP CAPACITY, AND SEXUALIZATION

INCLUDE FEMALE CHARACTERS THAT REPRESENT WOMEN ACROSS ALL STAGES OF THE LIFESPAN, WITH PARTICULAR ATTENTION TO WOMEN OVER 30

WHEN CREATING CHARACTERS OF COLOR, FOCUS ON CHALLENGING STEREOTYPES REGARDING WORK ETHIC AND OCCUPATIONAL STATUS

PRIORITIZE LGBTQ INDIVIDUALS WHEN CREATING NEW CHARACTERS, AIMING TO INCLUDE NUMBERS PROPORTIONAL WITH U.S. DEMOGRAPHICS (OR THAT OF THE COUNTRY IN WHICH THE AD WILL AIR)

PRIORITIZE INDIVIDUALS WITH PHYSICAL, COGNITIVE, AND COMMUNICATION DISABILITIES, AIMING TO INCLUDE NUMBERS PROPORTIONAL WITH U.S. DEMOGRAPHICS (OR THAT OF THE COUNTRY IN WHICH THE AD WILL AIR)

APPENDIX A

The GD-IQ was funded by Google.org. Using machine learning technology and the University of Southern California's audio-visual processing technologies, this tool was co-developed by the Institute and led by Dr. Shrikanth (Shri) Narayanan and his team of researchers at the University of Southern California's Signal Analysis and Interpretation Laboratory (SAIL), with additional analysis from Dr. Caroline Heldman and Soraya Giaccardi.

To date, most research investigations of media representations have been done manually. The GD-IQ revolutionizes this approach by using automated analysis, which is not only more precise, but makes it possible for researchers to quickly analyze massive amounts of data, which allows findings to be reported in real time. Additionally, the GD-IQ allows for more accurate analysis, and because the tool is automated, comparisons across data sets and researchers are possible, as is reproducibility. Automated analysis of media content gets around the limitations of human coding. Beyond the significant advantage of being able to efficiently analyze more in less time, the GD-IQ can also calculate content detail with a level of accuracy that eludes human coders. This is especially true for factors such as screen and speaking time, where near exact precision is possible. Algorithms are a set of rules of calculations that are used in problem-solving. For this report, we employed two automated algorithms that measure screen and speaking time of characters by their gender. Here is an overview of the procedures we used for each algorithm.

SCREEN TIME ANALYSIS

We compute the screen time of female characters by calculating the ratio of female faces to the total number of faces in the ad's visuals. The screen time is calculated using online face detection and tracking with tools provided by Google's machine learning technology. In the interest of precision and time, we estimate screen time by computing statistics over face-tracks (boxes tracking the general outline of each face) instead of individual faces. The face-tracks returned by technology include different attributes of the face with the corresponding time of occurrence in the video. Among the attributes returned for each of the detected faces, we use two parameters - the confidence of the detected face and the system's posterior probability for gender prediction. A threshold of 0.25 was empirically chosen for determining confident face detection.

Due to multiple characters appearing on screen simultaneously, the face-tracks can be overlapping. A gender label is then assigned to each track using the average gender posterior associated with the confident faces in the track. If the average gender posterior probability of the track is greater than 0.5, the track is classified as a "female track," otherwise, it is a "male track." The number of frames with confident face detections in each track is summed up across all tracks to get the total number of faces. The number of female tracks is aggregated to get the total number of faces predicted as female. Finally, the screen time is computed as the ratio between the number of female face detections to the total number of face detections across the length of the movie. Supplementary analysis shows that screen time estimated at frame-level (individual faces) instead of using face-tracks was not significantly different and was comparable. Furthermore, computing the average of gender posterior over tracks has an added benefit of "smoothing out" some of the local gender prediction errors. Face-tracking incorporates temporal contiguity information to reduce transient errors in gender prediction that may occur with analyzing individual faces independently.

SPEAKING TIME ANALYSIS

Using movie audio, we compute the speaking time of male and female characters to obtain an objective indicator of gender representation. The algorithm for performing this analysis involves automatic voice activity detection, audio segmentation, and gender classification.

Voice Activity Detection: Movie audio typically contains many non-speech regions, including sound effects, background music, and silence. The first step is to eliminate non-speech regions from the audio using voice activity detection (VAD) and retain only speech segments. We used a recurrent neural network based VAD algorithm implemented in the open-source toolkit OpenSMILE to isolate speech segments.

Segmentation: We then break speech segments into smaller sections in order to ensure each segment includes speech from only one speaker. This is performed using an algorithm based on Bayes Information Criterion (BIC), available in the KALDI toolkit. Thirteen dimensional Mel Frequency Cepstral Coefficient (MFCC) features are used for the automatic speaker segmentation. This step essentially decomposes continuous speech segments obtained in the VAD step into smaller segments to make sure no segment contains speech from two different speakers.

Gender Classification: The speech segment is then classified into two categories based on whether it was likely spoken by a male or female character. This is accomplished with acoustic feature extraction and feature normalization.

- *Acoustic Feature Extraction:* We use 13-dimensional MFCC features for gender classification because they can be reliably extracted from movie audio, unlike pitch or other high-level features where extraction is made unreliable by the diverse and noisy nature of movie audio.
- *Feature Normalization:* Feature normalization is deemed necessary to address the issue of variability of speech across different movies and speakers, and to reduce the effect of noise present in the audio channel. Cepstral Mean Normalization (CMN) is a standard technique popular in Automatic Speech Recognition (ASR) and other speech technology applications. Using this method, the cepstral coefficients are linearly transformed to have the same segmental statistics (zero mean).

Classification of the speaker as either male or female is based on gender-specific Gaussian mixture models (GMMs) of the acoustic features. These models are trained on a gender-annotated subset of general speech databases used for developing speech technologies using frame-level features for each gender. The GMM we use in this system has 100 mixture components and is optimized by tuning the parameters in a held-out evaluation set. For a new input segment whose gender label is to be predicted, the likelihoods of the segment belonging to a male or female class are computed based on this pre-trained model. The class with higher likelihood is assigned to the segment as the estimated gender prediction. The total speaking time by gender is then computed by adding together the durations for each utterance classified as Male/Female. This gives us the male and female speaking time in a movie.



ENDNOTES

1. The Henry J. Kaiser Family Foundation (2015). "Population Distribution by Gender," State Health Facts, <http://kff.org/other/state-indicator/distribution-by-gender/?currentTimeframe=0&sortModel=%7B%22colId%22:%22Location%22,%22sort%22:%22asc%22%7D>.
2. The Geena Davis Institute on Gender in Media and J. Walter Thompson Intelligence (2017). "Gender Bias in Advertising: Research, Trends & New Visual Language," <https://seejane.org/wp-content/uploads/gender-bias-in-advertising.pdf>.
3. Zurbriggen, E. L., Collins, R. L., Lamb, S., Roberts, T. A., Tolman, D. L., & Ward, L. M. (2007). APA Task Force on the Sexualization of Girls. American Psychological Association.
4. Hunt, Darnell Ana-Christina Ramón, Michael Tran, Amberia Sargent, and Vanessa Días (2017). 2017 Hollywood Diversity Report: Setting the Record Straight, Ralph J. Bunche Center for African American Studies at UCLA, February, <http://www.bunchecenter.ucla.edu/wp-content/uploads/2017/04/2017-Hollywood-Diversity-Report-2-21-17.pdf>.
5. Spiegelhalter, David (2015). "Is 10% of the Population Really Gay?" The Guardian. June 19, <https://www.theguardian.com/society/2015/apr/05/10-per-cent-population-gay-alfred-kinsey-statistics>.
6. The World Bank (2019). Disability Inclusion, <https://www.worldbank.org/en/topic/disability>.

HOW TO CITE THIS STUDY

Giaccardi, S., Cooper, R., Heldman, C., Cooper-Jones, N., McTaggart, N., Juliano, L., Phillips, H., Esparza, P., & Conroy, M. (2019). *Bias and inclusion in advertising: An analysis of 2018 Cannes Lions Film Craft ads*. The Geena Davis Institute on Gender in Media at Mount Saint Mary's University.

FOR QUESTIONS OR INFORMATION ON THIS REPORT, CONTACT:

SORAYA GIACCARDI
Associate Director of Research
soraya@seejane.org

DR. CAROLINE HELDMAN
Senior Research Advisor
caroline@seejane.org