



Geena Davis Inclusion Quotient™



The Reel Truth: Women Aren't Seen or Heard

An Automated Analysis of
Gender Representation in Popular Films



Geena Davis Institute  on Gender in Media



If she can see it, she can be it.™





Geena Davis
Inclusion Quotient™



“The GD-IQ is an extraordinary tool that gives us the power to uncover unconscious gender bias with a depth that had never been possible to date. Our hope is that we can use this technology to push the boundaries of how we identify the representation imbalance in media. Media that is more representative of our society not only fosters a more inclusive industry, but by increasing the number and diversity of female leaders and role models on screen, content creators are affecting the ambitions and career aspirations of young girls and young women everywhere. If she can see it, she can be it.” - Geena Davis

The Geena Davis Inclusion Quotient (GD-IQ) is 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 and incorporating Google’s machine learning technology, and the University of Southern California’s audio-visual processing technologies, GD-IQ 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), with additional analysis from Dr. Caroline Heldman.

Previous studies find that female characters are vastly underrepresented in film, and this has not changed much in the last half a century.¹ In this report, we use the GD-IQ to not only analyze gender representation but also analyze screen time and speaking time in the top 100 grossing films of 2014 and 2015. In addition, we have analyzed results by box office revenue.

To date, most research investigations of media representations have been done manually. The GD-IQ revolutionizes this approach by using automated analysis of media content with a precision that is not possible with the human eye or ear. It makes it possible for researchers to quickly analyze massive amounts of data, which allows findings to be reported in real time.

The GD-IQ was developed to more accurately measure gender representation in film. We find that female characters continue to be unrepresented in popular film, and when they are present, they have far less screen time and speaking time. This means that simply adding more female characters into films is not enough. To truly address gender inequity, female characters need to be seen and heard as often as their male counterparts.



Geena Davis Institute  on Gender in Media



If she can see it, she can be it.™





Geena Davis
Inclusion Quotient™

Methodology

We analyzed a total of 200 films for this report: the top grossing (non-animated)² films of 2014 and 2015, as reported by *Variety*. Findings for 2015 are presented here, and findings for 2014 are included in Appendix C. We use the GD-IQ, a revolutionary new automatic audio-visual tool— the first of its kind developed specifically to analyze media content.

Existing research on gender, race, and other representations in media almost exclusively employ content analysis performed manually by research assistants who view media content and score it “by hand.” The approach limits the number of films or other media content that can be analyzed due to time constraints. It is also subject to human error and lacks precision because it relies on a human capacity to record character and scene details.

Automated analysis of media content gets around the limitations of human coding. Beyond the significant advantage of being able to efficiently analyze more films in less time, the GD-IQ can also calculate content detail with a level of accuracy that eludes human coders.

For this report, we measure on-screen time by partitioning the movie into face-tracks by tracking the detected faces locally in time. Gender is computed for each face-track separately. We then calculate total screen time by gender for each film using the track duration. We measure speaking time by applying an automatic speech detection program that classifies the speaker as female or male. For further information about this automated processing tool, please refer to Appendix A.

In addition to the automated data, we include a general overview of gender in each film using hand coded data. For more information on the hand coding process, please refer to Appendix B.



Geena Davis Institute  on Gender in Media



If she can see it, she can be it.™





Geena Davis
Inclusion Quotient™

The Invisible Woman: Female Characters Not Seen or Heard

In 2015, 17% of the top grossing films had a female lead. Women had a particularly strong presence in the comedy and action genres. Amy Schumer, Melissa McCarthy, Tina Fey, Amy Poehler, and Anna Kendrick had hit movies, demonstrating that funny women are bankable in Hollywood.

For this analysis, we compare films with male leads, female leads, and male-female co-leads. The co-lead category includes ensemble casts where both men and women are featured roughly equally. Films with multiple male leads or multiple female leads were folded into the male lead and female lead categories, respectively. The GD-IQ shows that male characters dominated both the screen time and speaking time in the top grossing films of 2015.



Screen Time: Women Missing in Action

Male characters received two times the amount of screen time as female characters in 2015 (**28.5% compared to 16.0%**).



When a film has a male lead, this gender gap is even wider, with male characters appearing on screen nearly three times more often than female characters (**34.5% compared to 12.9%**).



In films with a female lead, male characters appear about the same amount of time as female characters (**24.0% compared to 22.6%**). This means that even when women are featured in a leading role, male characters appear on screen just as often.



In films with male and female co-leads, male characters receive significantly more screen time (**24.8%**) than female characters (**16.0%**). So even when men and women are both featured as leads in a film, male characters are far more prominent than female characters.



Geena Davis Institute  on Gender in Media



If she can see it, she can be it.™





Geena Davis
Inclusion Quotient™

Speaking Time: Women, the Silent Minority

In 2015, male characters spoke two times as often as female characters in the top box office movies (**28.4% compared to 15.4%**).



The gender gap in speaking time is even larger in films led by men. Male characters spoke three times more often than female characters (**33.1% compared to 9.8%**) in films with male leads.



In films with female leads, male characters spoke about the same amount as female characters (**23.9% compared to 26%**). In other words, in films with male leads, male characters dominate the speaking time, but in films with female leads, men speak as much as women.



In films with both male and female co-leads, male characters spoke far more often than female characters. Male characters spoke **25.5%** of the time compared to **16.7%** for female characters.



■ Female Screen Time
■ Male Screen Time

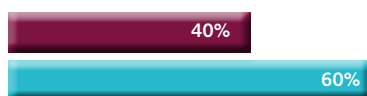
Female Lead



Male Lead



Co-Lead



■ Female Speaking Time
■ Male Speaking Time

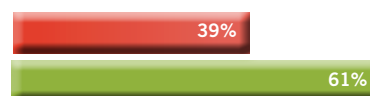
Female Lead



Male Lead



Co-Lead



Geena Davis Institute  on Gender in Media



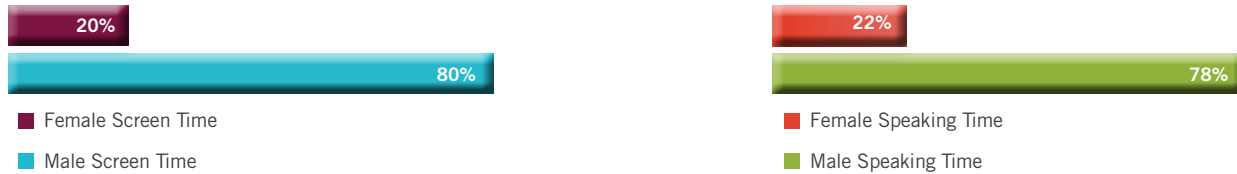
If she can see it, she can be it.™





Geena Davis
Inclusion Quotient™

Gender gaps in screen time and speaking time were even bigger in action films, a film genre that is typically dominated by men. Even though women played leading roles in action blockbusters such as *Star Wars: The Force Awakens* (Daisy Ridley), *The Hunger Games Series: Mockingjay Part 2* (Jennifer Lawrence), and *The Divergent Series: Insurgent* (Shailene Woodley), overall, male characters appeared and spoke on screen three times more often than female characters in action films.



Box Office: Gender Balance Brings In the Bucks

On average, the top 100 grossing non-animated films of 2015 earned \$90,660,000 each. Films with female leads made considerably more on average than films with male leads – \$89,941,176 for female leads compared to \$75,738,095 for male leads. Films led by women grossed **15.8%** more on average than films led by men.



Films featuring male and female co-leads earned \$108,317,073 – **23.5%** more on average than films with male or female leads alone.



Many factors determine the box office revenue of a given film, but these numbers are revealing. Our findings debunk the idea female leads are not bankable. Films with female leads actually earned more money than films with male leads, and casts with both male and female leads perform even better. Gender balance in casting produces sound financial returns.



Geena Davis Institute  on Gender in Media



If she can see it, she can be it.™





Geena Davis
Inclusion Quotient™

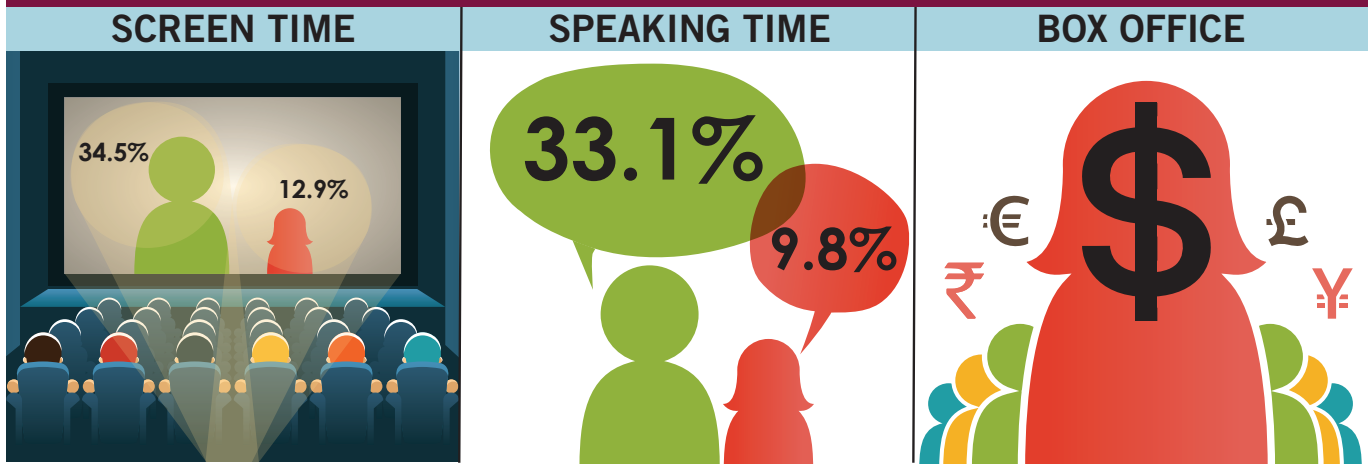
Conclusion

The revolutionary technology powering GD-IQ gives the tool the ability to uncover unconscious gender bias with a depth that has not been possible to date. The GD-IQ was designed to push the boundaries of how we identify the imbalance of the representation of specific demographics and stereotypes in media. Content creators from the worlds of film, television, advertising, publishing, digital and more will be able to identify and recognize the issues contributing to the problem and correct the course.

The GD-IQ's capabilities go well beyond simply analyzing gender bias. The Institute, in partnership with Google and USC Viterbi School of Engineering will present the research findings from this investigation along with additional automated methods to analyze individual-level character attributes, such as representations of animated characters and the composition of background scenes, at our Global Symposiums on Gender in Media in Los Angeles and in New York.

This first report concludes that women are underrepresented in film, and when they do appear, they are seen and heard far less than their male counterparts.

Summary of Key Findings



- Male characters received two times the amount of screen time as female characters in 2015 (**28.5% compared to 16.0%**).
- In films with a male lead, male characters appearing on screen nearly three times more often than female characters (**34.5% compared to 12.9%**).

- Male characters spoke two times as often as female characters (**28.4% compared to 15.4%**).
- In films with male leads, male characters spoke three times more often than female characters (**33.1% compared to 9.8%**).

- Films led by women grossed **15.8% more** on average than films led by men.



Geena Davis Institute  on Gender in Media



If she can see it, she can be it.™





**Geena Davis
Inclusion Quotient™**

Special Thanks

Julie Ann Crommett, Entertainment Industry Educator in Chief, Google

Jacqueline Fuller, Director, Charitable Giving and Advocacy, Google.org

Jason Lambert, Executive Director, Content Licensing & Metadata, Sony Pictures International Television

Jennie Peters, Founder, 82 Rogue

Megan Smith

GD-IQ Team



Geena Davis
Founder & Chair
Geena Davis Institute on Gender in Media

Academy Award® winner Geena Davis, one of Hollywood’s most respected actors, is recognized for her tireless advocacy of gender equality in media nearly as much as for her acting accomplishments. She is the Founder and Chair of the Geena Davis Institute on Gender in Media, which is successfully influencing film and television content creators to dramatically increase the percentages of female characters — and reduce gender stereotyping — in media targeting children 11 and under.

In 2015, Davis launched the Bentonville Film Festival (BFF), an unprecedented initiative in support of women and diversity in the entertainment industry and serves as its Co-Founder and Chair. BFF provides a platform to significantly boost the commercial value of content produced and starring minorities and women. This initiative is the only film festival in the world to provide guaranteed theatrical, television, digital and home entertainment distribution for its winners.

In 2012, Davis was appointed Special Envoy for Women and Girls in ICT for the UN’s International Telecommunication Union (ITU). She is an official partner of UN Women, working toward their goal of promoting gender equality and empowering women worldwide. Davis is also an appointee to the California Commission on the Status of Women. As an actor, Davis has appeared in several roles that have become cultural landmarks. Earning the 2006 Golden Globe Award for Best Performance by an Actress in a Television Series Drama, Davis broke ground in her portrayal of the first female President of the United States in ABC’s hit show *Commander in Chief*.

In 1989, Davis received the Academy Award® for Best Supporting Actress for her role in Lawrence Kasdan’s *The Accidental Tourist*. She was again nominated for an Academy Award® and Golden Globe for her performance as Thelma in Ridley Scott’s *Thelma and Louise*. Davis went on to receive a Golden Globe nomination for Best Actress for her work in *A League of Their Own*. Other credits include iconic films such as *Tootsie*, *The Fly*, *Beetlejuice*, *The Long Kiss Goodnight*, and *Stuart Little*.

Few have achieved such remarkable success in as many different fields as Davis has: she is not only an award winning actor, but a world-class athlete (at one time the nation’s 13th-ranked archer, and a semi-finalist in the Olympic Trials), and a member of the genius society Mensa. Davis holds honorary degrees from Boston University, Bates College and New England College.

Geena stars in the upcoming feature film *Marjorie Prime* with Tim Robbins and Jon Hamm and the new fall FOX series *The Exorcist*.




Geena Davis Institute on Gender in Media



If she can see it, she can be it.™





**Geena Davis
Inclusion Quotient™**



Madeline Di Nonno
Chief Executive Officer
Geena Davis Institute on Gender in Media

Madeline Di Nonno is the Chief Executive Officer of the Geena Davis Institute on Gender in Media, the only research based organization working in the entertainment and media industry to accelerate gender and diversity representation in children’s entertainment through cutting-edge research, education and advocacy programs. Di Nonno leads the Institute’s strategic direction, management, financial and operational activities. Di Nonno brings over thirty years of executive leadership experience in the entertainment, non-profit, digital and consumer packaged goods industries.

Previously, Di Nonno served as President and CEO of On The Scene Productions, a leader in digital video creation and distribution. Di Nonno lead the company’s new business development and video content platforms for the entertainment, healthcare, and consumer products clients such as Nike, Gatorade, Iconix, and PepsiCo.

Prior to On The Scene, Di Nonno served in executive marketing positions for Anchor Bay Entertainment/Starz Media and Echo Bridge Home Entertainment where she led global brand marketing, branded franchise development, acquisitions, digital media initiatives for home entertainment and limited theatrical releases in North America, United Kingdom and Australia.

As Executive Vice President and GM for Nielsen EDI, a leading research provider of Theatrical distribution measurement and information, Di Nonno drove innovations in digitizing product offering, new business and product development and studio theatrical distribution and exhibition client management.

Previously, Di Nonno served as Senior Vice President, Marketing Alliances and Digital media at the *Hallmark Channel*, and launched the cable channel and established marketing, digital media, e-commerce, and corporate alliance functions. Di Nonno developed revenue-generating, integrated marketing programs with companies such as Johnson & Johnson, Fuji, Mail Boxes Etc., Universal Pictures, Target, Sony Pictures, Chrysler, and Baskin Robbins. Di Nonno pioneered the Channel’s digital media initiatives.

Di Nonno served eight years at Universal Studios Home Video as Vice President, Strategic Marketing where she established the company’s first Digital Platforms and marketing campaigns. Di Nonno supervised marketing campaigns for all theatrical, direct-to-video and library releases and spearheaded Universal’s launch into DVD and interactive.

Di Nonno began her career at ABC Television Network in corporate publicity working on the marketing communications campaigns for mini-series, sports and daytime, including “*The Winds of War*,” “*The Thorn Birds*” and the 1984 Olympics. Di Nonno is a requested speaker at major events, companies, and organizations such as The White House, United Nations, UNESCO, CinemaCon, CTAM, CES, and featured in media publications such as *The New York Times*, *Variety*, *The Hollywood Reporter*, *Huffington Post*, *Kidscreen*, *Fast Company*, and *Wired Magazine*.

Madeline holds a bachelor’s degree from Boston University.

Madeline serves on the Board of Directors for the Television Academy Foundation; Promundo U.S.; Board of Advisors Los Angeles Press Club; GameChanger Films and the Rutgers University Institute for Women’s Leadership. Madeline served as President of the Glass Lions Jury for Cannes Lions 2016.



If she can see it, she can be it.™





**Geena Davis
Inclusion Quotient™**



Elizabeth Kilpatrick
Vice President, Development
Geena Davis Institute on Gender in Media

Elizabeth Kilpatrick has 20 years management, fundraising, marketing, research management and strategic planning experience with a range of organizations from major universities to strictly volunteer-driven grassroots organizations. Prior to working with the Geena Davis Institute on Gender in Media, she held the position of National Director of Development for Cure Autism Now (now Autism Speaks) and Director of Development for the Alzheimer's Association of Orange County. She worked in management roles with Goodwill Industries of Southern California and the University of Southern California School of Engineering.

Prior to her work in the non-profit field, Kilpatrick worked as Manager of Quality Program overseeing total quality management programs for Archive Corporation and Conner Peripherals. She is a Handley Walker certified lead assessor of management systems and a member of the founding administrative committee for the International Meeting for Autism Research. She attended Western Carolina University and the University of Southern California, earning a Bachelor's degree in English Literature.



Shrikanth (Shri) Narayanan, Ph.D.
Andrew J. Viterbi Professor of Engineering
University of Southern California

Dr. Narayanan is the Andrew J. Viterbi Professor of Engineering at the University of Southern California (USC), and holds appointments as Professor of Electrical Engineering, Computer Science, Linguistics, Psychology, Neuroscience and Pediatrics and is the founding director of the Ming Hsieh Institute. He also directs the USC Signal Analysis and Interpretation Laboratory (SAIL).

His research focuses on human communication and information processing with applications including media, education and health informatics. He is a Fellow of the Acoustical Society of America, IEEE, International Speech Communication Association and the American Association for the Advancement of Science (AAAS) and a recipient of numerous research and education awards. He has published over 700 papers and has been granted seventeen U.S. patents, and his work has been featured widely in national and international print and broadcast media.



Caroline Heldman, Ph.D.
Geena Davis Institute on Gender in Media
Associate Professor of Politics, Occidental College, Los Angeles

Dr. Heldman is the Research Director for the Geena Davis institute on Gender in Media and an Associate Professor of Politics, Occidental College, Los Angeles. Her research specializes in the presidency, gender, race, and sexuality in U.S. politics. She co-edited *Rethinking Madame President: Are We Ready for a Woman in the White House?* (2007), and she has been published in the top journals in her field. Her work has been featured in the *New York Times*, *U.S. News and World Report*, *Ms. Magazine*, *The Huffington Post*, and *The Daily Beast*, and on the *Katie Couric Show* and the *Rosie Show*.



If she can see it, she can be it.™





**Geena Davis
Inclusion Quotient™**

Dr. Heldman's new book, *Women, Power, and Politics: The Fight for Gender Equality in the United States*, will be published with Oxford University Press this year. Dr. Heldman has also worked as a regular commentator for *MSNBC*, *Fox News*, *Fox Business News*, *CNBC*, *RT America*, and *Al Jazeera America*, and writes for *Ms. Magazine Blog*.

Dr. Heldman has been featured in many popular documentary films, including *Aftermath: New Orleans in the Wake of Hurricane Katrina*, *Miss Representation*, *Equal Means Equal*, *Informant*, and *The Mask You Live In*. She has also been active in "real world" politics as a congressional staffer and campaign manager.

Dr. Heldman drove to New Orleans to help with rebuilding efforts the week that Hurricane Katrina hit the Gulf Coast, and co-founded the New Orleans Women's Shelter and the Lower Ninth Ward Living Museum. Dr. Heldman is also an active leader in the campus anti-rape movement. She co-founded End Rape on Campus (EROC) and Faculty Against Rape (FAR), and currently co-chairs the statewide campaign to abolish the statute of limitations for rape in California.



Hartwig Adam, Ph.D.
Senior Staff Engineer
Google

Hartwig Adam is a Senior Staff Engineer at Google, where he leads the Mobile Vision team in the Machine Intelligence group. His group focuses on developing computer vision and machine perception systems for mobile devices and Google's internal and external developer platforms. Previous projects include Photo Search, Google Goggles and Glass. Prior to Google, Hartwig was a co-founder and VP of Platform Development at Neven Vision, an early pioneer in mobile computer vision.



Ian Breckinridge-Jackson

Ian Breckinridge-Jackson is a visiting instructor at California State University, Los Angeles. He was awarded a National Science Foundation Graduate Research Fellowship, the American Sociological Association Community Action Research Initiative Award, and a research grant from the University of California Center for New Racial Studies. Mr. Breckenridge-Jackson's research has been published in *Politics, Groups, and Identities*, *Policy Matters*, *The Oxford Handbook of U.S. Women's Social Movement Activism*, and *The Routledge International Handbook of World Systems Analysis*. His research has also been featured in *Time Magazine* and *U.S. News & World Report*.



Rebecca Cooper

Rebecca manages the Lower Ninth Ward Oral History Project and works as a researcher for the Mossville Oral History Project at Louisiana State University, T. Harry Williams Center for Oral History. Ms. Cooper's research has been featured in the *Oxford University Press Oral History Review* and the *Contemporary Justice Review*.



Geena Davis Institute  on Gender in Media



If she can see it, she can be it.™





**Geena Davis
Inclusion Quotient™**



Tanaya Guha, Ph.D.

Tanaya is an assistant professor (electrical engineering) at IIT Kanpur, India. She was with SAIL USC as a postdoctoral fellow from 2013 - 2015. She has received her Ph.D. (2013) from the University of British Columbia (UBC), Vancouver, Canada. She was a recipient of Mensa Canada Woodham's memorial scholarship, Google Anita Borg scholarship, and Amazon Grace Hopper celebration scholarship.



Ian Hall

Ian Hall, a research assistant on this project for the Geena Davis Institute on Gender in Media, is an undergraduate studying English and Film at Santa Monica College with an interest in documentary work and gender issues.



Che-wei Huang

Che-Wei Huang is a Ph.D. student at SAIL USC. His research interests fall into the intersection of machine learning and pattern recognition. In particular, his research topics include manifold learning applied to speaker clustering, emotion recognition, and information-theoretical analysis of deep neural networks.



Naveen Kumar, Ph.D.

Naveen received a Ph.D in Electrical Engineering from USC in 2016. He received a B.Tech. degree in instrumentation engineering from the Indian Institute of Technology, Kharagpur, India, in 2009. He was awarded the Viterbi School Dean's Doctoral Fellowship at USC in 2009. His research interests include machine learning and signal/image processing for applications to speech and multimedia problems. Naveen received his Ph.D. in Electrical Engineering from USC in 2016.



Nikalos Malandrakis

Nikos is a Ph.D. candidate in the Department of Computer Science, USC. He received a Diploma and M.Sc. in Electronic and Computer Engineering, in 2007 and 2012 respectively, from the Technical University of Crete, Greece. His research revolves around the human perception of multimedia.



Geena Davis Institute  on Gender in Media



If she can see it, she can be it.™





**Geena Davis
Inclusion Quotient™**



Md Nasir

Md Nasir is a Ph.D. student in the USC Department of Electrical Engineering. His research interests lie in the field of speech processing and machine learning with applications in behavioral signal processing and media content analysis.



Anil Ramakrishna

Anil Ramakrishna is a third year Ph.D. student in the USC Department of Computer Science. His research interests are in Machine Learning and Natural Language Processing. For MICA, he is currently focusing on building character networks for movies to automatically answer questions related to Bechdel test, character cliques and gender representations thereof.



Krishna Somandepalli

Krishna is a Ph.D. student in the USC Department of Electrical Engineering, USC. He received his master's degree in Electrical Engineering from University of California, Santa Barbara in 2012 and worked as a Jr. Research Scientist at NYU Langone Medical Center from 2012-2015. His research interests include computer vision, affective signal processing and media content analysis.



Yan Zhu

Yan is an undergraduate Electrical Engineering student and a research assistant with the media informatics and content analysis (MICA) group at SAIL. She started working with MICA as a freshman because of the societal value and implications of its research and continue to work with discovering more from movie content analysis.



If she can see it, she can be it.™





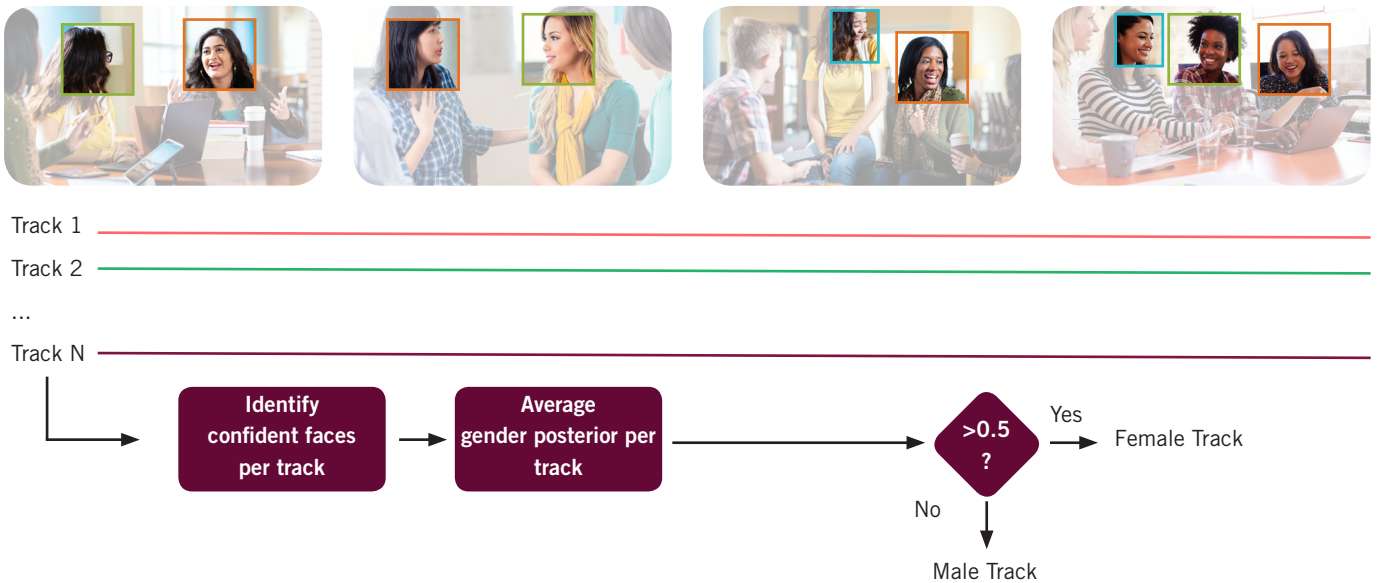
Appendix A

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 film’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. An overview of the on-screen time estimation process is shown in Figure 1.

Figure 1.



Due to multiple characters appearing on screen simultaneously, the face-tracks can be overlapping, as illustrated in Figure 1. 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





Geena Davis
Inclusion Quotient™

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 a 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.



Geena Davis Institute on Gender in Media



If she can see it, she can be it.™





Geena Davis
Inclusion Quotient™

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.

Appendix B

Data on character prominence was produced through hand coding. For this study, leading characters are defined as the major force driving the story. Co-leads are classified as two (or more) characters that share roughly equivalent screen time and are equally involved in driving the story. Some analysts require that characters appear within the first five minutes of a film to be counted as a lead or co-lead, but for our analysis, we evaluate the entire film to determine the prominence of the character.

A team of three researchers conducted a content analysis that produced these statistics. Prior to initiating the work, the research team met twice for training sessions and performed multiple statistical tests to ensure that their analysis was in agreement. They calculated inter-coder reliability on 10 films (10% of the sample) that were not in the top 100 grossing of 2015 and 2014 to ensure agreement. Inter-coder reliability was achieved in terms of both absolute agreement and Cohen's Kappa measures.



Geena Davis Institute  on Gender in Media



If she can see it, she can be it.™





Geena Davis
Inclusion Quotient™

Appendix C

Top Films of 2014

Of the top 100 grossing films of 2014 compiled by *Variety*, 11% featured a female lead. According to our automated analysis, male characters dominated both the screen time and speaking time in the top grossing films of 2014. The findings for 2014 look remarkably similar to the findings for 2015. (We note one difference below.)

Screen Time

Male characters received twice as much screen time as female characters in 2014. Male characters were on screen **29.6%** of the time compared to **15.9%** of the time for female characters.



In films with male leads, female characters only appear on screen **12.3%** of the time, while male characters appear **32.8%** of the time. This means that male characters are almost three times more likely than female characters to appear on screen in films with male leads.



In films with female leads, male characters (**20.5%**) get about the same screen time as female characters (**21.6%**). Having a female lead does not translate into more screen time for female characters than male characters.



Male characters get significantly more screen time in films that have male leads than female characters get in films that have female leads (**32.8% compared to 21.6%**). When men play the leading role, male characters dominate the screen time, but when women play the leading role there is no screen time advantage for female characters.



In films with male and female co-leads, male characters appear on screen more often than female characters (**26.9% compared to 20.6%**). Adding a woman as a co-lead in a film does not mean female characters get equal screen time with male characters.





Geena Davis
Inclusion Quotient™

Speaking Time

Male characters are twice as likely to speak as female characters in the top grossing films. Overall, male characters spoke **31.8%** of the time in films compared to **14.5%** of the time for female characters.



In movies with male leads, male characters speak four times as much as female characters (**36.7% compared to 8.8%**). When men play the lead in a film, male characters have a strong speaking time advantage.



In movies with female leads, female characters speak **29.5%** of the time compared to **17.7%** for male characters. This speaking time advantage for female characters is not as large as the advantage male characters get in films with male leads, and this advantage disappears in films from 2015.



In films with male and female co-leads, male characters speak significantly more than female characters (**27.8% compared to 19.7%**). This means that when men and women lead a cast together, male characters still dominate the dialogue.



Revenue

On average, the top 100 grossing non-animated films of 2014 earned \$72, 228,000 each. Films with female leads made considerably more on average than films with male leads or male and female co-leads. Films with female leads earned \$108,909,090 compared to \$70,500,000 for films with male leads and \$66,161,290 for films with male and female co-leads.



- Female Lead Revenue
- Male Lead Revenue
- Co-Lead Revenue



If she can see it, she can be it.™





Geena Davis
Inclusion Quotient™

[1] See the Geena Davis Institute on Gender in Media, 2014. "Gender Bias Without Borders: An Investigation of Female Characters in Popular Films Across 11 Countries." Stacy L. Smith, Marc Choueiti, & Dr. Katherine Pieper with assistance from Yu-Ting Liu & Christine Song Media, Diversity, & Social Change Initiative USC Annenberg, retrieved from <http://seejane.org/wp-content/uploads/gender-bias-without-borders-executive-summary.pdf>

[2] Our datasets do not include animated films since the automated tool is not yet able to read animated characters. Future reports will include animated films.



Geena Davis Institute  on Gender in Media



If she can see it, she can be it.™

