

Exploring Inclusion in Snap Content

Dr. René Weber, Dr. Stacy L. Smith, Dr. Katherine Pieper,
Musa Malik & Brooke Kong

Executive Summary

April 2023



Media Neuroscience Lab (MNL)
University of California, Santa Barbara
<https://medianeuroscience.org>

USC Annenberg
Inclusion Initiative

Annenberg Inclusion Initiative (AII)
University of Southern California
<https://annenberg.usc.edu/aii>

SNAPCHAT 

Introduction

It is imperative for media organizations to understand whether the content they deliver represents their users. This is particularly important in the case of large and influential technology companies, such as Snap and its platform Snapchat. Per Snapchat's internal information, the platform reaches 90% of U.S. teens and young adults, with 375 million daily active users globally. This means that Snapchat content reaches an audience that is highly diverse across race/ethnicity, gender, and sexuality. Yet, is the content that Snapchat provides to these users as diverse as its audience?

To answer this question, the Media Neuroscience Lab (MNL) at the University of California Santa Barbara, and the Annenberg Inclusion Initiative (AII) at the University of Southern California partnered with Snap to perform a first-of-its-kind Diversity Equity and Inclusion (DEI) analysis of Snapchat Partnered and Originals Content. Pairing MNL's computational methods with AII's human assessment techniques resulted in a scalable procedure to reliably evaluate a large quantity of Content in a relatively short period. The outcome is a deeper understanding of how Snapchat Content reflects its audience with regard to gender, race/ethnicity, and other inclusion metrics.

Below we share a brief, executive summary of our findings from a hybrid and computational approach, combining AII's human coding with MNL's algorithmic evaluation. Next, we turn to a fully computational assessment of Snap Content videos conducted by the MNL, which provides insights into the structural representations of female and underrepresented (UR) characters within Snapchat narratives. We conclude our report with a description of methodological innovations provided by MNL and AII that extended beyond the contractual agreements, and with recommendations for future DEI research at Snap. A detailed report of all methodologies and findings has been submitted to Snap on November 1st, 2022, and is attached to this summary.

Overall Key Findings: Snapchat Partnered Content

We performed an analysis of Snapchat's 2021 Partnered and Original Content. Snapchat's Partnered Content shows were analyzed in a hybrid fashion by AII's human coders and MNL's computational pipelines to capture prevalence metrics on gender, race/ethnicity, and LGBTQ identity. We initially sampled 16,254 unique Snapchat stories randomly selected from datafiles provided by the Snap team. From this population of videos, there were 1,944 Snapchat Partnered Content units with at least 1 speaking or named character that were extracted from across 636 channels and included within the hybrid evaluation procedure. These 1,944 videos comprised our final sample.

Each individual character in the Snapchat Partnered Content shows was assessed for gender, race/ethnicity, and LGBTQ identity. The hybrid approach provides more conservative estimates on prevalence measures as compared to our computational pipelines. For the sake of ensuring reliability in reporting statistical estimates in our key findings, we default to the hybrid approach. Detailed comparisons between these two approaches are provided in the attached full report.

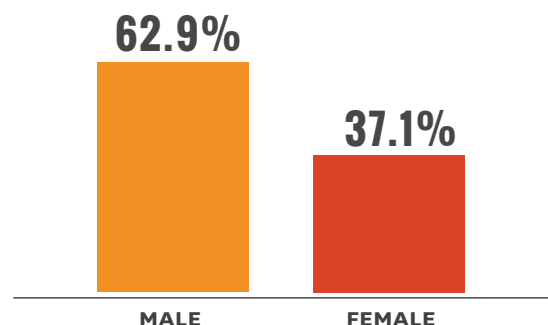
Gender Prevalence

Of the 7,456 human characters identified, 62.9% were male, 37.1% were female, and < 1% (n=9) identified as non binary. In comparison to the percentage of girls/women in the U.S. population (50.5%), Snapchat Partnered Content fell below proportional representation. Our findings on gender prevalence

patterns suggest that approximately 1.7 male characters appeared for every female character featured in Snapchat Partnered Content.

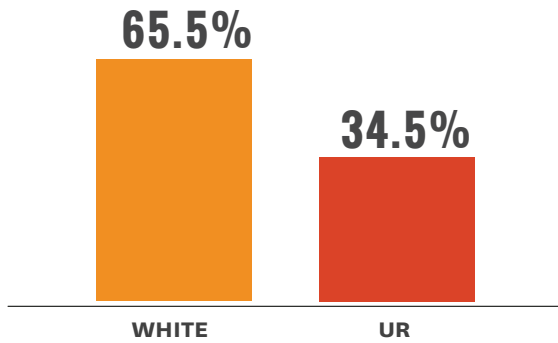
Using genre tags provided by Snap, each video was assigned to a specific genre. Then, gender prevalence was evaluated across each genre in the sample. Female characters were more likely to appear in specific genre categories such as Beauty, Fashion & Style, Animals, DIY & Crafts, Animation, and Parenting. Female characters were least likely to appear in video content published across the majority of other genres. The identification of specific genre categories where girls/women were least likely to appear provides Snap with targeted areas for possible intervention.

GENDER OF CHARACTERS ACROSS SNAPCHAT PARTNERED CONTENT
Prevalence of men and women in Snapchat Partnered Content evaluated with the hybrid approach



UNDERREPRESENTED STATUS IN SNAPCHAT PARTNERED CONTENT

Percentage of characters from underrepresented racial/ethnic groups



Race/Ethnicity Prevalence

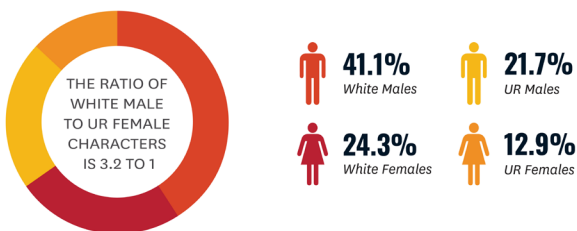
A total of 7,245 individual characters were evaluated for racial/ethnic identification. Based on U.S. Census categories, 65.5% of characters were White, 17.7% were Black/African-American, 5.2% were Hispanic/Latino, 5.7% were Asian, 4.3% were Multiracial/Multiethnic, 1.4% were Middle Eastern/North African, <1% were Native Hawaiian/Pacific Islander, and <1% were American Indian/Alaska Native.

Overall, 34.5% of all characters were from underrepresented racial/ethnic groups, which approached proportional representation to the U.S. population (39.9%). In other words, approximately 1.9 White characters appeared for every underrepresented character featured in Snapchat 2021 Partnered Content.

Underrepresented characters were more likely to appear in video content published from channels associated with the genres of Sports, Beauty, Parenting, Travel, and General Satisfying. Underrepresented characters were least likely to appear in video content published across the majority of other genres.

INTERSECTIONAL REPRESENTATION IN SNAPCHAT CONTENT

Percentage of characters by underrepresented status & gender in Snapchat partnered content



Taking an intersectional perspective by examining race/ethnicity and gender, 41.1% of speaking characters were White males, 24.3% were White females, 21.7% were underrepresented males, and 12.9% were underrepresented females.

As a point of comparison, across the U.S. population, 30% are White males, 30% are White females, 20% are underrepresented males, and 20% are underrepresented females. Thus, underrepresented girls/women fell below proportional representation in Snap Partnered Content. Put differently, White male characters outnumbered underrepresented female characters in 2021 Partnered content at a ratio of 3.2 to 1.

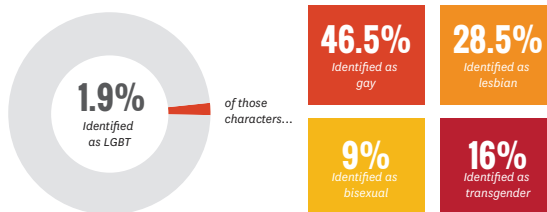
LGBTQ Prevalence

Of the 7,456 characters evaluated, 1.9% were identified as LGBTQ. Snapchat Partnered Content underrepresents LGBTQ characters in comparison to the percentage of U.S. adults who self-identify as lesbian, gay, bisexual, and transgender (7.1%), according to recent research (Gallup, 2021).

Of the LGBTQ-identified characters, 46.5% identified as gay, 28.5% as lesbian, 9% as bisexual, and 16% were transgender. More than half of the LGBTQ characters (56.4%) were male while 43.6% were female.

LGBT STATUS OF CHARACTERS ACROSS SNAPCHAT PARTNERED CONTENT

Out of 7,456 speaking characters...



Visual Portrayals and Structural Representations of Underrepresented Females

In addition to examining overall prevalence, we conducted advanced statistical modeling to evaluate how visual portrayal (on-screen appearance) and structural representation (networks within narratives) might be associated with different demographic groups. This was accomplished using computational procedures that relied upon formal features of each video in addition to other metrics.

We found an interesting and consistent pattern related to the visual portrayal of underrepresented females. On average and as compared to White males, girls/women of color had more screen time, were more often featured individually, and received more substantial visual attention overall. Across the structural metrics, underrepresented females were significantly more likely than White males to be in a central or prominent role when they appear in Snap content. These results suggest that although girls/women of color are less prevalent, their roles have greater importance to the narrative.

Despite these intriguing findings, we are not able to make a conclusive inference about whether a greater focus on underrepresented females is a positive or negative outcome. There was no qualitative review of a selection of these Snapchat stories that could provide insight into whether these metrics reflected meaningful narrative portrayals (e.g., journalistic contexts) or negative stereotypes (e.g., hypersexualization). Nonetheless, these findings, and the computational method that produced them, provide an important perspective on how narrative features may differ across demographic groups. Additionally, these findings speak to what may be

unconsciously salient within the creative process for publishers of Snapchat Partnered Content.

There were also compelling findings related to the structural representation of groups. Underrepresented and female characters were more frequently associated with other members of gender and racial/ethnic minority groups. Put differently, when girls/women were on screen they were more likely to appear in narratives with other girls/women. The same was true for underrepresented characters.

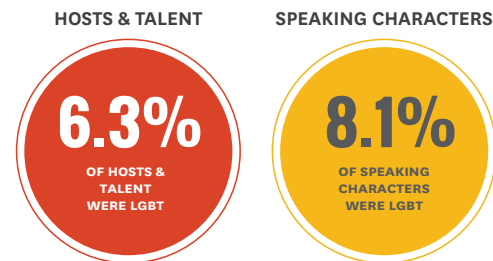
One inference from these findings is that publishers of Snapchat Partnered Content create content that is frequently homogeneous in terms of cast members' racial/ethnic and gender affiliations. While we believe that these findings may simply be a reflection of the type of narratives that are generally created on social media, we provide in this study a computational framework and a diagnostic toolkit for creators of Snapchat Partnered Content to use in their own work, push for increased awareness, and help chart the way towards greater on-screen and structural equity in representation across the entire Snapchat Partnered Content ecosystem.

Key Findings: Snapchat Originals

A total of 300 Snapchat Original episodes from 26 Snapchat series were evaluated by human coders at the Annenberg Inclusion Initiative to identify hosts and talent as well as every speaking character. Of the 63 hosts and talent, 52.4% were girls/women and 47.6% were boys/men. More than half (57.1%) of hosts/talent were from an underrepresented racial/ethnic group while 42.9% were white. Less than 10% of hosts/talent were LGBTQ-identified, including one transgender host. Only 1 host/talent was a person with a disability.

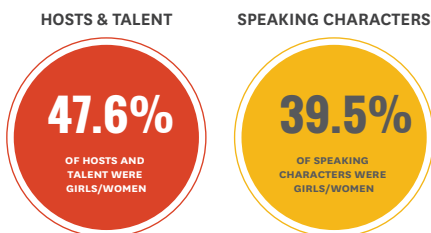
LGBT STATUS OF CHARACTERS ACROSS SNAP ORIGINALS

Out of 63 Hosts & Talent and 726 speaking characters...



LESS THAN HALF OF SNAP ORIGINAL SERIES STAR GIRLS & WOMEN

Prevalence of girls and women in Snap Original series, 2021



Among the 726 speaking or named characters identified across Snapchat Originals, 39.5% were female-identified and 60.5% were male-identified. In terms of race/ethnicity,

61% of speaking characters were from an underrepresented racial/ethnic group and 39% were white. Nearly 70% of Snapchat Original series achieved proportional representation, or featured underrepresented speaking characters in at least 39.5% of all speaking roles. Crossing race/ethnicity and gender, 22.7% of all speaking characters were white males, 16.2% were white females, 37.5% were underrepresented males, and 23.5% were underrepresented females. This suggests that Snapchat Originals feature girls/women of color at proportional representation to the U.S. population. Finally, 8.1% of speaking characters in Snapchat Originals were LGBTQ-identified and 1.4% were shown with a disability.

Methodological Innovations

Supported by Snap, MNL and AII have generated novel and innovative solutions for pushing the envelope on classical and computational inclusion research. Central for the methodological innovations is the development of the Measuring and Tracking Inclusion online platform (MTI; <https://inclusion.mnl.ucsb.edu>), a full workflow platform for extracting and sampling content, preparing content for coding, administering content coding across diverse coder groups, computational analyses of content, and the interactive reporting of content inclusion profiles. Among a number of new algorithms, MTI includes an innovative and previously unknown computational approach for the detection of LGBTQ identities, computer vision pipelines for the extraction of visual and structural representations of characters (e.g., are female characters central or peripheral in stories; beyond prevalence, how much “screen time” falls to UR characters), and new algorithms for the reporting (and improvement) of content coding reliability. We recommend future DEI research initiatives that take advantage of the extensively validated and purposefully designed infrastructure that MNL and AII have worked towards in collaboration with Snap. These may include API provisioning for evaluating DEI metrics within Snap Partnered Content on an almost real-time basis and at scale as well as providing an interactive inclusion reporting platform (MTI-Snap) specifically tailored for content-producers, Snap decision-makers, the public, and other DEI stakeholders.

Conclusion

The purpose of this study was to investigate inclusion in Snapchat 2021 Partnered and Original Content. Using hybrid procedures, we assessed gender, race/ethnicity, and LGBTQ identification across more than 7,000 characters in Snapchat Partnered Content. The results provide an indication of where Snapchat content represents its users and where greater inclusion is needed.

Gender Parity is Elusive in Snap Content

The portrayal of female-identified characters lags behind proportional representation. Snapchat Partnered Content features 37.1% female-identified characters overall. While girls/women may have prominence in some Snapchat content, there is room to grow when it comes to overall inclusion.

Racial/Ethnic Representation Varies by Content Type

Slightly more than a third (34.5%) of characters in Snapchat Partnered Content were from underrepresented racial/ethnic groups, which is below the 39.9% of the U.S. population that identifies as underrepresented. In line with the findings for the prevalence of underrepresented characters overall, there were discrepancies for characters from specific racial/ethnic groups. In particular, Hispanic/Latino, Middle Eastern/North African, and Native/Indigenous representation are places where Snapchat under- indexes compared to population figures.

Additionally, increasing the representation of girls and women from these groups is an important way to bolster inclusion across content. Women of color were outnumbered by White male characters by over 3 to 1, and thus are clearly underrepresented compared to their prevalence in the U.S. population. While the overall results for racial/ethnic representation are encouraging, disaggregating the data in this manner reveals places for improvement.

Depictions of the LGBTQ Community are Rare

Snapchat Partnered Content featured relatively few LGBTQ characters. While this underrepresentation is reflective of broader trends in the entertainment industries, for this inclusion metric, Snapchat can make clear gains and should consider ways to introduce creators and channels that feature comparatively higher rates of the LGBTQ Community into its programming.

One additional note regarding the analysis presented in this report regards the presentation of characters with disabilities in Snap content. At this time, the hybrid methodology used in this report did not include a computational analysis of characters with disabilities. Human coders did evaluate portrayals of disability across content related to Snap Original series. That analysis is not reported here. Results on this community were presented to Snap as a reflection of the company and research

team's joint commitment to understanding how Snap content represents individuals with disabilities. Computational analyses of individuals' with disabilities are in development and results will be included in future reports.

Final Notes

This report recognizes the areas in which Snap is making progress in diversity and inclusion measures. Using the techniques outlined in the full report and with the empirical, theoretical, and methodological strength of the research teams, the groups examined the inclusion profile of social media content that reaches millions of users each day. As a result, the report reveals current inclusion gaps and areas of improvement for future Snapchat Partnered content. Like other media and technology companies, Snap Partnered Content reflects trends for the representation of women, people of color, and the LGBTQ+ community that continue to persist across entertainment. There are pathways for improvement, however. From the findings of this comprehensive analysis, Snap can continue to evaluate its content to determine how it aligns with external population data, and thus its user base. From there, Snap can focus its efforts on encouraging, creating, and distributing content that represents the diversity of its users, leveraging its relationships with its partners to ensure its values for inclusion are expressed through its content.

Acknowledgements

This project brought together two institutions in a collaborative effort to examine content. We are grateful to the team at Snap for their support and insights, including Michael DiBenedetto, Ben Reininga, Oona King, Jess Harvie, Rebecca Lee, Farrin Jay, Abigail Lucas, Paul Gillan, David Hornsby, and others who facilitated this project and provided necessary assistance along the way.

At the Media Neuroscience Lab, we would like to thank Yibei Chen, Paula Wang, Brittany Wheeler, Kylie Woodman, and Sungbin Youk for all conceptual and motivational input. A special thanks goes to our terrific staff and IT support, Andreas Boschke and Jeff Oaks.

At the Annenberg Inclusion Initiative, thank you to Ariana Case, Katherine Neff, Sam Wheeler, Bryan Davis, Terrell Shaffer, Karla Hernandez, Zoe Moore, Zoily Mercado, Sarah Neff, and Matthew Davis, our fantastic staff team, for their work to support this project. We had a dedicated student team as well who worked to examine Snap content. Thank you to our student research assistants!

Annenberg Inclusion Initiative Student Coders

Aimee Christopher	Jessica LeBron
Allison Walsh	Julia Zara
Amanda Fassett	Julianna Birlin
Amy Andrade	Lauren Milstein
Anish Lahorani	Lindsey Huguet
Camryn Robinson	Louis Wong
Carmela Van der Meulen	Neeharika Goyal
Cat (Yuetong) Chen	Ni Huang
Daphne Zhu	Olivia Mossler
Deeptavarna Mummineni	Ravi Chandra Basireddy
Devanshi Krishna Shah	Samanyu Raju Kakarlapudi
Elena (Yifei) Zhao	Sarah Kim
Emily Yeung	Sky (Zihang) E
Gabriella Rudy	Sofia Heuchert
Garima Danidhariya	Sophia Pelaez
Gavin Murillo	Swapnil Gyanprakash Arya
Gretchen Rudolph	Varsha Pramod Kini
Hannah Ritz	Vivienne Yang
Isabelle Lim	Xinyi Fan
Jenisty Colon	Yifan Jiang