AMERICAN PSYCHOLOGICAL ASSOCIATION

Psychology of Popular Media

ISSN: 2689-6567

2023, Vol. 12, No. 4, 499–504
https://doi.org/10.1037/ppm0000436

BRIEF REPORT

Women in Fiction: Bechdel-Wallace Test Results for the Highest-Grossing Movies of the Last Four Decades

Markus Appel[1] and Timo Gnambs[2]

[1] Psychology of Communication and New Media, Julius-Maximilians University of Würzburg

[2] Leibniz Institute for Educational Trajectories, Bamberg, Germany

The representation of women in cultural products is a main concern to the social sciences and humanities and a topic that many citizens care about. We analyzed female characters in the 30 worldwide highest grossing movies per year, for the past 40 years, amounting to 1,200 movies. Our study was based on the Bechdel-Wallace test (BWT, Bechdel test): Are two or more named women present in a movie, and do they talk to each other about something besides a man? We used data from an impromptu online citizen science project (bechdeltest.com) and complemented the data with analyses of movies that were not covered by this database. Our results show that only half of the most popular movies (49.58%) pass the test. Almost all movies (95.31%) pass the reverse BWT (movies featuring named men who talk to each other about something besides a woman), as indicated by an analysis of a subset of 341 movies. Time trends indicate that the percentage of movies passing the BWT is on the rise for the past 10 years. The probability of a movie passing the BWT and respective changes over time are related to the movies' setting (alternate vs. contemporary world), audience evaluations (IMDb ratings), production budget, and revenue. Implications and limitations are discussed.

Public Policy Relevance Statement

The representation of women in movies was analyzed based on the 30 most popular movies each year, spanning the years from 1980 to 2019 (1,200 movies in total). Bechdel-Wallace test results show that less than 50% of all movies include two or more named women who talk to each other about something besides a man. Even if the share of movies passing the Bechdel-Wallace test increased in the 2010s, the depiction of women in fictional worlds need to be acknowledged when trying to understand the origins of gender stereotypes and the underrepresentation of women in real life.

Keywords: fiction, movies, gender, women, Bechdel test

Supplemental materials: https://doi.org/10.1037/ppm0000436.supp

This article was published Online First September 1, 2022.

Markus Appel https://orcid.org/0000-0003-4111-1308

The authors are indebted to everyone who provided test scores for movies to Bechdeltest.com. The authors are unaffiliated with this website. Data, including a list of all movies, materials, computer code, and analyses outputs are available at https://osf.io/jb6cf.

The data are available at https://osf.io/jb6cf

The materials are available at https://osf.io/jb6cf

Correspondence concerning this article should be addressed to Markus Appel, Psychology of Communication and New Media, Julius-Maximilians University of Würzburg, Oswald-Külpe-Weg 82, Würzburg 97074, Germany. Email: markus.appel@uni-wuerzburg.de

Mediated stories are considered to shape how humans perceive the world and themselves (Green et al., 2020). Thus, stereotypes and the underrepresentation of women and minorities are a matter of concern (Aubrey & Roberts, 2020; Gestos et al., 2018). The representation of women in motion pictures has attracted a particularly strong attention by journalists and citizens. In 1986, comic author Alison Bechdel (Bechdel, 2005) created a comic strip on the topic and suggested a test to assess the representation of women in a movie. The test has three requirements to be passed:

(i) The movie has to have at least two [named] women in it, who

(ii) talk to each other about

(iii) something besides a man.

The Bechdel-Wallace test (BWT, Bechdel test) has become a mainstay in the public discourse about women in movies (Garber, 2015; Racic, 2018; Seth, 2020). Importantly, citizens have started to conduct the BWT and to post their results online. This impromptu citizen science project (Kullenberg & Kasperowski, 2016; Rughiniş et al., 2016) has yielded a database of Bechdel test scores for over 8,500 movies (bechdeltest.com). In contrast to the relevance of the BWT in public discourse, a scientific analysis of BWT results for a large corpus of movies is missing. Thus, our aim was to build upon this database to analyze the representation of women in movies over the past 40 years. Our target were the 30 highest grossing movies per year, produced in the years 1980 to 2019 (1,200 movies in total; worldwide revenue as of March 15, 2020). Target movies that were not available in the bechdeltest.com database were coded as part of this project.

The Bechdel-Wallace Test in Context

The analysis of characters and stories depicted in mainstream media is widely considered a relevant endeavor, not least due to the potential impact of media content on individuals' thoughts, emotions, and behavior. According to *social–cognitive theory* (Bandura, 1978, 2001), much of human psychological functioning is acquired through observational, vicarious learning. Often, learning is unintentional. Importantly, a large part of knowledge about the world, about norms, social roles, and behavioral patterns is learnt from the symbolic environments that media provide. Regarding gender stereotypes and gender roles, mainstream media have been described as particularly influential symbolic environments, due to the abundance of potential role models, the attention they receive (by the individual as well as parents and peers), and the substantial time spent with media by adolescents and young adults (Ward & Grower, 2020). Relatedly, narrative theory and research has emphasized the pivotal role of stories in the development of individuals' perceptions of themselves and the world around them (Green et al., 2020; Mar, 2018). The power of stories to build and change beliefs and self-knowledge is not restricted to news and other nonfictional fare, it similarly applies to fictional stories in which authors are free to invent the characters and events taking place (Appel & Malečkar, 2012). This perspective is shared by *cultivation theory* which further highlights the need for understanding media worlds as an important step in analyzing how everyday media use shapes real-world knowledge and beliefs (Gerbner, 1998).

An analysis of BWT test scores adds to previous work that used different content-analytic methods to analyze the representation of women in movies, movie subgenres, or other screen media (see Ward & Grower, 2020, for an overview). Content-analytic research conducted in the past suggests that women are underrepresented in screen media and portrayed in stereotypical roles. However, some evidence suggests that the depiction of women could have changed over the years. On U.S. primetime TV, for example, female characters are portrayed as similarly competent (articulate, motivated, and intelligent) as men (Sink & Mastro, 2017; primetime TV from 2013 was analyzed). A content-analysis of movies popular between 2007 and 2018 (Smith et al., 2019), however, shows that the percentage of female characters among all speaking or named characters is still low (33.1% in 2018) with little change from the first year that was considered (29.9% in 2007). An analysis of BWT data of the most popular movies of the past 40 years could add to the debate by providing a longer time perspective, and by using the BWT as a metric that is frequently referred to in the public debate.

The Present Project

Our starting point was the data available on bechdeltest.com, a BWT impromptu citizen science database of movies. This database, however, is incomplete: values for 261 movies belonging to our target pool of the 1,200 most popular movies across the past four decades were missing (particularly for older production years). These missing data had been an obstacle to valid scientific analyses of BWT results so far. The missing movies were analyzed as part of the present work. A supplementary reverse BWT (male character focus) was conducted for a subset of movies. BWT results (failed or passed) and the percentage of movies that passed the test over the years were of key interest.

In addition to providing an estimate of the percentage of popular movies of the past 40 years that pass the BWT (and the reverse BWT), we were interested in variables that predict the percentage of passing the BWT for the total sample of movies and the trajectory of movies passing the BWT over time. We considered the setting of the movie—in terms of contemporary world versus alternate world—as a focal variable in this regard. We argue that movies set in the here and now (contemporary world) are often based upon prevalent social roles and related beliefs and expectations, including gender roles (Eagly, 1987). According to prevalent gender role beliefs and gender stereotypes women are less agentic than men (Sczesny et al., 2019), which should lead to fewer active female characters in popular movies. This tendency should be reduced for movies set in an alternate world (in the future, or in a fantasy world) in which prevalent social roles may not apply. We tested the assumption that movies set in an alternate world (in which, e.g., dogs talk or protagonists have superpowers) have a higher likelihood of passing the BWT than movies set in a contemporary world. Reflecting changes in gender roles and stereotypes over time (Bhatia & Bhatia, 2021; Eagly et al., 2020) and initial evidence on changing gender representations in the media (Ward & Grower, 2020) we further expected a higher likelihood to pass the BWT for more recent movies. On a more explorative note, we further examined whether audience evaluations (IMDb ratings), the relative budget, or the relative revenue were associated with BWT results.

Method

BWT Scoring

In accordance with the specifications and norms expressed on bechdeltest.com, the first part of the BWT was operationalized as the presence of two named women. Moreover, female characters of nonhuman species, such as Grace (voiced by Sigourney Weaver) in James Cameron's Avatar were filed as women, and male characters of nonhuman species were filed as men. Please

note that the terms *female* and *male* characters are used here to indicate gender rather than biological sex. Agender, gender-fluid, or nonbinary characters were neither categorized as female characters nor as male characters. The communicative sequence required to pass subtests two and three consists of a conversation in which a verbal communicative act by a named female character is responded to with a verbal communicative act by a second named female character. The Bechdeltest database (bechdeltest.com) includes scores representing the sum of subtests passed (min = 0, max = 3). For example, a film that has two or more named female characters, but no two named female characters ever talk to each other during the film, would receive a score of 1. If all three subtests were passed, the Bechdel test as a whole would be passed.

Data and Procedure

Movies released in the years 1980 to 2019 were examined (N = 1,200). The 30 highest grossing movies per year were identified on the-numbers.com (worldwide revenue as of March 15, 2020). The production year 2020 was not considered because of the limited movie releases in this year due to the COVID-19 pandemic. Additional information retrieved on the-numbers.com were the production budget (in $) and worldwide revenue (in $). Based on the-numbers.com and movie summaries on IMDb, the movie setting (contemporary world vs. alternate worlds) was categorized. If the Bechdeltest database (bechdeltest.com) provided a test score for a movie, this test score was integrated into the analyses. Among the 1,200 target movies, 261 movies were missing in the Bechdeltest database.

These missing movies as well as a random sample of 82 movies with data in the Bechdeltest database (343 movies in total) were coded by 153 undergraduate media communication students (87% women) for partial course credit. A total of 152 students coded two movies, one student coded one movie. Another 35 movies were coded by three student research assistants (all women), three movies were coded by the first author. Ten movies were not available to us and, therefore, replaced by a movie with the highest popularity that did not meet our original threshold (i.e., the 31st most popular movie in a given year). The students were introduced to the task in an in-person training session in which the BWT and the coding task were introduced in detail. The coding itself was accomplished with the help of a coding sheet and a handout with coding instructions (see data and materials on our OSF page). The coding sheet included the three questions of the BWT and the three complementary questions of the reverse BWT. Coders were further required to specify details on the characters and conversations that made a movie pass the BWT. The coding was conducted on the basis of a DVD that was provided.

The large majority of the 261 newly coded movies that were missing in the database were produced in the United States (218) or were U.S.-involved coproductions (16). The remaining movies were produced internationally, for example, in China (seven) or the United Kingdom (six) or coproduced with several countries involved. The coders were advised to conduct the coding with the English-language version of the movie, if possible. Of the 261 movies, 245 were coded with the English-language version, 15 were coded with the German-language version (by German-language native speakers), one was coded based on the Mandarin version with English subtitles.

Based on a random sample of 82 movies included in the Bechdeltest database that were coded a second time by us, we obtained an interrater reliability of Cohen's κ of .65. Although this reliability score is far from perfect, it can be interpreted as acceptable based on widely used guidelines (Landis & Koch, 1977; LeBreton & Senter, 2008).

For a subset of 341 movies, we applied the reverse Bechdel test to examine the prevalence of men who talk to men about something else than a woman, as a point of comparison.

Analyses

Variables measured in U.S. $ were detrended to account for inflation across years. Moreover, logistic regression analyses provided marginal effects of time trends for these correlates. Inference tests are not reported because rather than sampling a subset of movies we aimed at examining the full population of highest grossing movies.

Data Archive

Data, including a list of all movies, materials, computer code, and analyses outputs are available at https://osf.io/jb6cf. Bechdeltest.com data are CC BY-NC- licensed.

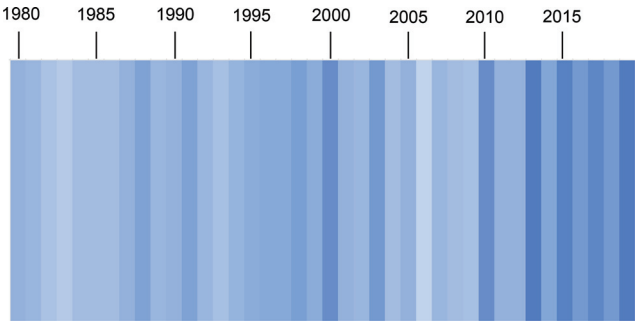
Results and Discussion

The BWT was passed by 49.58% of the 1,200 most popular movies of the past 40 years. As a point of comparison, 95.31% of coded movies passed the reverse BWT (based on a subset of 341 movies), speaking to a much stronger representation of men than women.

The percentage of movies that passed the BWT for each year is illustrated in Figures 1 and 2. The color scheme (Figure 1) and the smoothed regression line (Figure 2) show the share of movies passing the BWT each year. Particularly during the past decade, the respective share of movies gradually increased. While in the 1980s (42.28%), 1990s (48.80%), and 2000s (48.80%) less than half of the movies passed the BWT, the share in the last decade

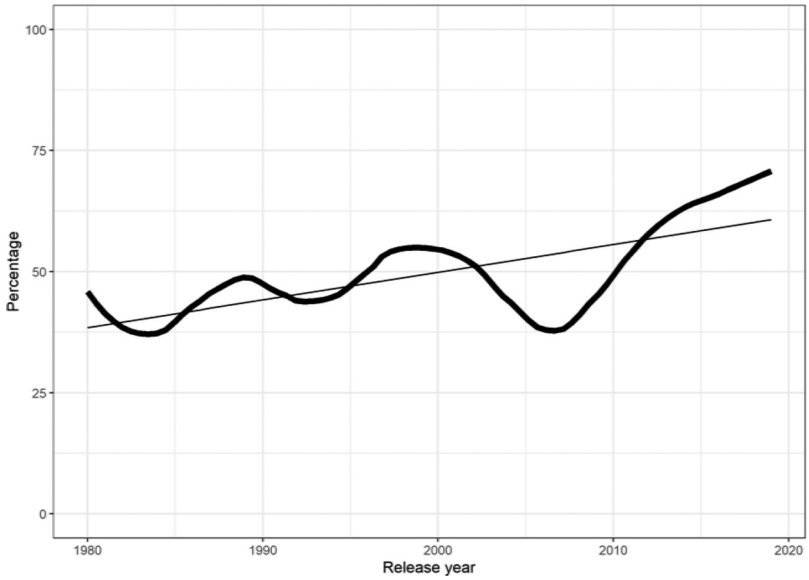
Figure 1

Stripe Chart Displaying the Percentage of Popular Movies Per Year That Passed the Bechdel-Wallace Test (BWT)



Note. Two-color scaling ranging from 0% (white) to 100% (deep blue). The lowest percentage of movies passing the BWT was observed in 2006 (26.67%), the highest percentage of movies passing the BWT was observed in 2013 and 2019 (76.67%). See the online article for the color version of this figure.

Figure 2
Percentage of the 30 Most Popular Movies Per Year That Pass the Bechdel-Wallace Test as Smoothed (Bold Solid Line) and Linear Regression Lines (Pale Solid Line)



(2010s) rose to 62.59%. A linear association between movie release year and percentage of movies that passed the BWT per year amounts to $r = .556$.

Next, a logistic regression analysis based on the 1,200 movies as the unit of analysis and the BWT score (pass = 1; fail = 0) as the criterion was conducted. The probability of passing the BWT changed across time ($r = .13$) and was associated with the movie setting, audience evaluations, production budgets, and revenues (Table 1). The estimated average marginal mean (AME) showed that, on average, the probability of passing the BWT increased by 1.94% each year. The probability of passing the BWT was higher (AME = .025) and increased more strongly (AME_{year}) for movies set in alternative worlds as compared with settings in contemporary worlds. It appears that screenwriters and directors use the freedom of future or fantasy settings to reserve a more prominent spot for female characters.

While production budget was negatively associated with the probability of passing the BWT overall (AME = −.037), increases of movies passing the BWT over the years were stronger for movies with higher production budgets, higher revenues, or higher audience evaluations.

We provided the first large-scale scientific analysis of the prevalence of women in movies based on the BWT. Given its easily accessible metric, our results can inform not only the scientific but also the public discourse on mainstream storytelling and the representation of women in culture. Across the 30 most popular movies per year during the past 40 years (1,200 movies in total) only around half of the movies passed the BWT (whereas around 95% of a subset of movies analyzed passed the reverse BWT). Until the 2000s, little systematic change across release years was observed, however, since 2010 a trend toward a stronger representation of women based on the BWT emerged. If this trend continued, it would still take many years until BWT numbers reached the representation of men in movies (as indicated by the reverse BWT).

Despite the contribution of our research, limitations need to be noted. A main limitation is the BWT itself. We believe that the BWT has strong face validity, but its representation of story content is rather indirect. Not all movies that pass the BWT necessarily feature prominent female characters or a feminist plotline. Likewise, a movie with a valuable emancipatory storyline that

Table 1
Correlations and Average Marginal Means for Passing the BWT

Variable	*r*	AME	AME_{year}
Release year	0.132	0.019	
Movie setting	0.088	0.025	
Contemporary world[a]			0.014
Alternate world			0.031
IMDb rating	−0.039	−0.028	
Low (4)			0.002
Medium (6)			0.011
High (8)			0.032
Relative budget[b]	−0.069	−0.037	
Small ($M - 1\ SD$)			0.011
Average (M)			0.017
Large ($M + 1\ SD$)			0.024
Relative revenue[b]	0.055	0.011	
Small ($M - 1\ SD$)			0.003
Average (M)			0.021
Large ($M + 1\ SD$)			0.035

Note. Unit of analysis: movie ($N = 1,200$). r = correlation between passing the BWT and a given variable. BWT = Bechdel-Wallace test; AME = average marginal mean for the respective variable; AME_{year} = average marginal mean for release year based on a logistic regression of passing the BWT on all given variables.
[a] Reference category. [b] z-standardized.

features one female protagonist only—a possible, but unlikely scenario—would fail the BWT. Such cases are arguably rare; still, the possibility that the role of women is misrepresented by the BWT score needs to be acknowledged when evaluating singular movies. As a second limitation we need to acknowledge that the BWT in its original form and its application in this article is based on a binary concept of gender. Agender, gender-fluid, or nonbinary characters were virtually absent in the movies we analyzed, and we acknowledge that our analyses do not contribute to questions of related representations in movies. The analyses further do not provide insights regarding the representation of race and ethnicity in movies (for ideas on adapting the BWT in this regard, see, e.g., Latif & Latif, 2016).

Our work has implications beyond a mere assessment of aesthetic choices in fictional worlds. Humans acquire gender-related knowledge, norms, social roles, and behavioral patterns from the symbolic environments that media provide (Bandura, 1978, 2001; Ward & Grower, 2020). Decades of media effects research on group representations and stereotypes show that media can have manifest consequences in real life (Appel & Weber, 2021; Aubrey & Roberts, 2020; Banas et al., 2020). Stories are considered to be particularly influential, as they can absorb audience members and transport them into narrative worlds, reducing their willingness and ability to critically reflect on settings, characters or storylines (Appel & Richter, 2007; Green et al., 2020; Mar, 2018). The research presented here did not examine the influence of media on cognitions, feelings, or behavior. However, these complementary lines of research suggest that the limited agency of women in many popular movies that was identified in our BWT analyses likely translates to social roles, expectations, and aspirations.

Achieving gender equality and empowering women and girls is one of the goals that societies worldwide have subscribed to (United Nations, 2015). Women in media is a topic that many people care about, as indicated by the popularity of the BWT. The research presented here shows that progress has been made in the last decade, but still a substantial part of the worldwide most popular movies does not incorporate agentic female characters, as indicated by a failed BWT. Changing the popular symbolic environments of feature films could be one means to approach gender equality in real life.

References

Appel, M., & Malečkar, B. (2012). The influence of paratext on narrative persuasion: Fact, fiction, or fake? *Human Communication Research*, *38*(4), 459–484. https://doi.org/10.1111/j.1468-2958.2012.01432.x

Appel, M., & Richter, T. (2007). Persuasive effects of fictional narratives increase over time. *Media Psychology*, *10*(1), 113–134. https://doi.org/10.1080/15213260701301194

Appel, M., & Weber, S. (2021). Do mass mediated stereotypes harm members of negatively stereotyped groups? A meta-analytical review on media-generated stereotype threat and stereotype lift. *Communication Research*, *48*(2), 151–179. https://doi.org/10.1177/0093650217715543

Aubrey, J. S., & Roberts, L. (2020). Effects of media use on development of gender role beliefs. In J. van den Bulck (Ed.), *The international encyclopedia of media psychology* (1st ed., pp. 1–12). Wiley. https://doi.org/10.1002/9781119011071.iemp0081

Banas, J. A., Bessarabova, E., & Massey, Z. B. (2020). Meta-analysis on mediated contact and prejudice. *Human Communication Research*, *46*(2–3), 120–160. https://doi.org/10.1093/hcr/hqaa004

Bandura, A. (1978). Social learning theory of aggression. *Journal of Communication*, *28*(3), 12–29. https://doi.org/10.1111/j.1460-2466.1978.tb01621.x

Bandura, A. (2001). Social cognitive theory of mass communication. *Media Psychology*, *3*(3), 265–299. https://doi.org/10.1207/S1532785XMEP0303_03

Bechdel, A. (2005, August 16). *The rule*. Dykes to watch out for. https://dykestowatchoutfor.com/the-rule

Bhatia, N., & Bhatia, S. (2021). Changes in gender stereotypes over time: A computational analysis. *Psychology of Women Quarterly*, *45*(1), 106–125. https://doi.org/10.1177/0361684320977178

Eagly, A. H. (1987). *Sex differences in social behavior: A social-role interpretation*. Erlbaum.

Eagly, A. H., Nater, C., Miller, D. I., Kaufmann, M., & Sczesny, S. (2020). Gender stereotypes have changed: A cross-temporal meta-analysis of U.S. public opinion polls from 1946 to 2018. *American Psychologist*, *75*(3), 301–315. https://doi.org/10.1037/amp0000494

Garber, M. (2015, August 25). Call it the 'Bechdel-Wallace-Test'. *The Atlantic*. https://www.theatlantic.com/entertainment/archive/2015/08/call-it-the-bechdel-wallace-test/402259/

Gerbner, G. (1998). Cultivation analysis: An overview. *Mass Communication and Society*, *1*(3–4), 175–194. https://doi.org/10.1080/15205436.1998.9677855

Gestos, M., Smith-Merry, J., & Campbell, A. (2018). Representation of women in video games: A systematic review of literature in consideration of adult female wellbeing. *Cyberpsychology, Behavior, and Social Networking*, *21*(9), 535–541. https://doi.org/10.1089/cyber.2017.0376

Green, M., Bilandzic, H., Fitzgerald, K., & Paravati, E. (2020). Narrative effects. In M. B. Oliver, A. A. Raney, & J. Bryant (Eds.), *Media effects: Advances in theory and research* (3rd ed., pp. 130–145). Routledge.

Kullenberg, C., & Kasperowski, D. (2016). What is citizen science? – A scientometric meta-analysis. *PLoS ONE*, *11*(1), Article e0147152. https://doi.org/10.1371/journal.pone.0147152

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174. https://doi.org/10.2307/2529310

Latif, N., & Latif, L. (2016). How to fix Hollywood's race problem. *The Guardian*. https://www.theguardian.com/film/2016/jan/18/hollywoods-race-problem-film-industry-actors-of-colour

LeBreton, J. M., & Senter, J. L. (2008). Answers to 20 questions about interrater reliability and interrater agreement. *Organizational Research Methods*, *11*(4), 815–852. https://doi.org/10.1177/1094428106296642

Mar, R. A. (2018). Stories and the promotion of social cognition. *Current Directions in Psychological Science*, *27*(4), 257–262. https://doi.org/10.1177/0963721417749654

Racic, M. (2018, March 3). Do this year's best picture Oscar nominees pass the Bechdel test? *The New Yorker*. https://www.newyorker.com/culture/culture-desk/do-this-years-best-picture-oscar-nominees-pass-the-bechdel-test

Rughiniş, C., Rughiniş, R., & Humă, B. (2016). Impromptu crowd science and the mystery of the Bechdel-Wallace test movement. In J. Kaye & A. Druin (Eds.), *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems,* Association for Computing Machinery (pp. 487–500). Association for Computing Machinery. https://doi.org/10.1145/2851581.2892580

Sczesny, S., Nater, C., & Eagly, A. H. (2019). Agency and communion: Their implications for gender stereotypes and gender identities. In A. Abele & B. Wojciszke (Eds.), *Agency and communion in social psychology* (pp. 103–116). Routledge.

Seth, R. (2020, July 15). Why are films failing the Bechdel test when TV has progressed? *Vogue*. https://www.vogue.co.uk/arts-and-lifestyle/article/the-bechdel-test

Sink, A., & Mastro, D. (2017). Depictions of gender on primetime television: A quantitative content analysis. *Mass Communication and Society*, *20*(1), 3–22. https://doi.org/10.1080/15205436.2016.1212243

Smith, S. L., Choueity, M., Pieper, K., Yao, K., Case, A., & Choi, A. (2019). *Inequality in 1,200 popular films: Examining portrayals of gender, race/ethnicity, LGBTQ & disability from 2007 to 2018*. Annenberg Foundation.

United Nations. (2015). *Transforming our world: The 2030 agenda for sustainable development*. https://www.unfpa.org/resources/transforming-our-world-2030-agenda-sustainable-development

Ward, L. M., & Grower, P. (2020). Media and the development of gender role stereotypes. *Annual Review of Developmental Psychology*, *2*(1), 177–199. https://doi.org/10.1146/annurev-devpsych-051120-010630

Received March 16, 2022
Revision received August 2, 2022
Accepted August 2, 2022 ■