

Tidy Tuesday - Week 11 (2021)

Julia Tache

3/9/2021

Bechdel Test

Happy Women's History Month!

The Bechdel Test for film is defined as followed: "there are at least two named women in the picture, they have a conversation with each other at some point, and that conversation isn't about a male character." One of the big questions that arises in debates about female representation in film is whether or not movies about women will be as successful as those about men? Let's put the data to the test- the Bechdel test:

```
library(tidyverse)

tuesdata <- tidyTuesdayR::tt_load('2021-03-09')

##
## Downloading file 1 of 2: `raw_bechdel.csv`
## Downloading file 2 of 2: `movies.csv`

movies <- tuesdata$movies
head(movies)

## # A tibble: 6 x 34
##   year imdb title test clean_test binary budget domgross intgross code
##   <dbl> <chr> <chr> <chr> <chr>      <chr>   <dbl> <chr>    <chr>    <chr>
## 1  2013 tt17~ 21 &~ nota~ notalk   FAIL    1.30e7 25682380 42195766 2013~
## 2  2012 tt13~ Dred~ ok-d~ ok      PASS    4.50e7 13414714 40868994 2012~
## 3  2013 tt20~ 12 Y~ nota~ notalk   FAIL    2.00e7 53107035 1586070~ 2013~
## 4  2013 tt12~ 2 Gu~ nota~ notalk   FAIL    6.10e7 75612460 1324930~ 2013~
## 5  2013 tt04~ 42   men  men     FAIL    4.00e7 95020213 95020213 2013~
## 6  2013 tt13~ 47 R~  men  men     FAIL    2.25e8 38362475 1458038~ 2013~
## # ... with 24 more variables: budget_2013 <dbl>, domgross_2013 <chr>,
## #   intgross_2013 <chr>, period_code <dbl>, decade_code <dbl>,
## #   imdb_id <chr>, plot <chr>, rated <chr>, response <lgl>,
## #   language <chr>, country <chr>, writer <chr>, metascore <dbl>,
## #   imdb_rating <dbl>, director <chr>, released <chr>, actors <chr>,
## #   genre <chr>, awards <chr>, runtime <chr>, type <chr>, poster <chr>,
## #   imdb_votes <dbl>, error <lgl>
```

Having financial backing for a film is very important: capital, resources, and investment into the film can all aid in its eventual release, promotion, and reception.

(Side note: Wes Anderson films probably won't pass the Bechdel test, but at the very least the palettes are very pretty and fit with today's cinematic theme.)

```
library("wesanderson")

movies$decade <- str_replace(movies$year, "\\d{1}$", "0s")
```

```

totals <- movies %>%
  group_by(decade, binary) %>%
  summarize(total = sum(budget_2013))

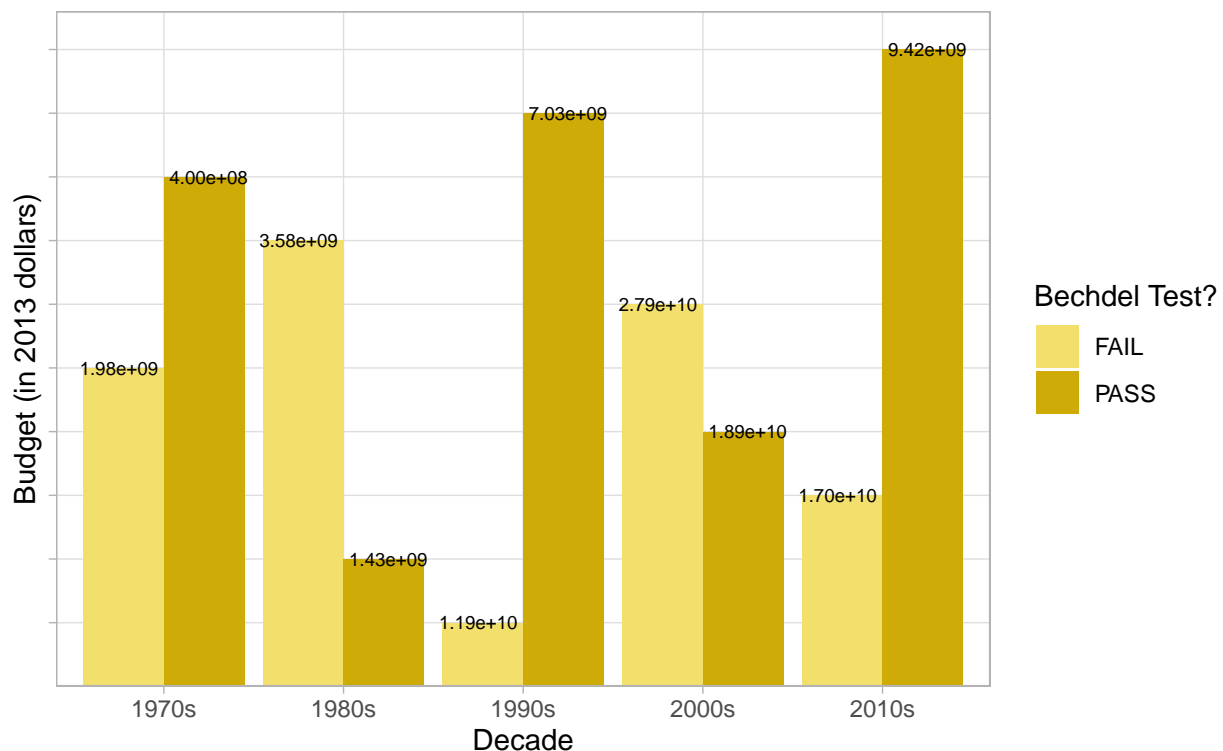
totals$total <- formatC(totals$total, format = "e", digits = 2)

ggplot(totals, aes(x = decade, y = total, group = binary)) +
  geom_bar(aes(fill = binary), stat = "identity", position = "dodge") +
  geom_text(aes(decade, total, label = total, fill = NULL), data = totals, size = 2.5, position = position_dodge()) +
  xlab("Decade") +
  ylab("Budget (in 2013 dollars)") +
  labs(title = "Film Budgets over the decades (1970 - 2013)",
       subtitle = "Films that pass the Bechdel test hit budget lows in the 80s and 2000s",
       fill = "Bechdel Test?") +
  scale_fill_manual(values = wes_palette("Moonrise1", n = 2)) +
  theme_light() +
  theme(axis.text.y = element_blank())

```

Film Budgets over the decades (1970 – 2013)

Films that pass the Bechdel test hit budget lows in the 80s and 2000s



Interestingly, despite lower budgets, films that passed the test had higher gross domestic profits in those same decades (80s and 90s).

```

movies$domgross_2013 <- as.numeric(movies$domgross_2013)

totals <- movies %>%
  group_by(decade, binary) %>%
  summarize(total = sum(domgross_2013, na.rm = TRUE))

```

```

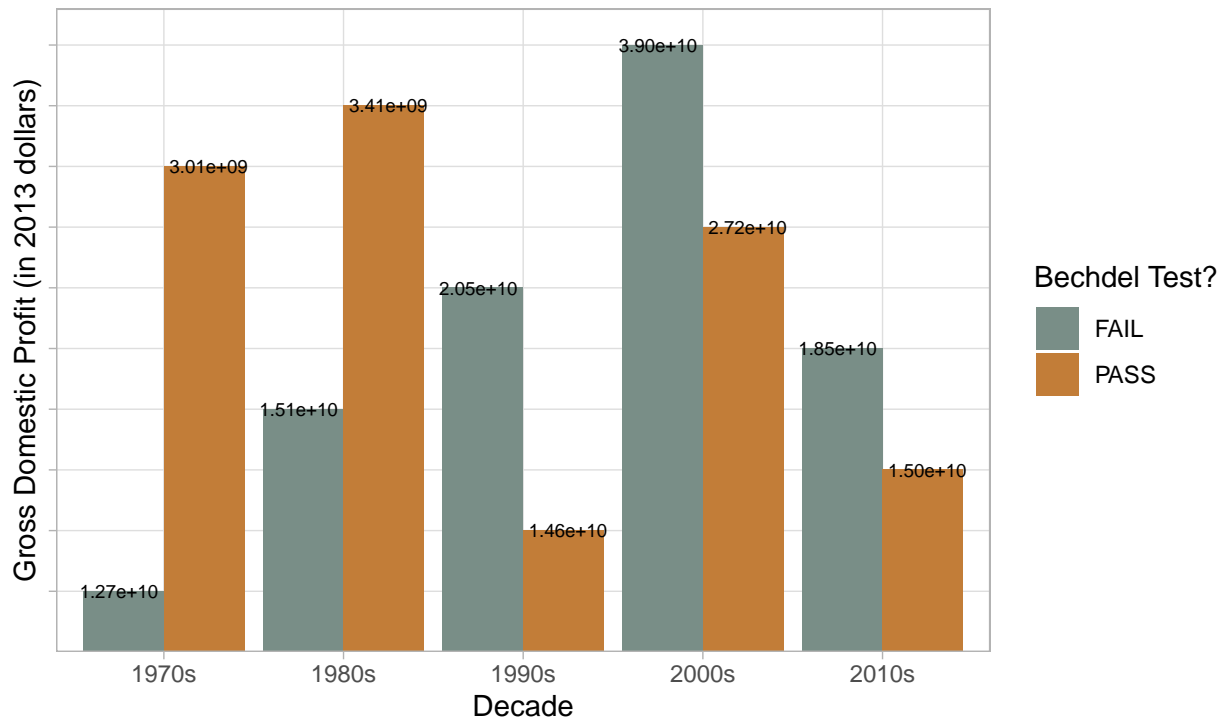
totals$total <- formatC(totals$total, format = "e", digits = 2)

ggplot(totals, aes(x = decade, y = total, group = binary)) +
  geom_bar(aes(fill = binary), stat = "identity", position = "dodge") +
  geom_text(aes(decade, total, label = total, fill = NULL), data = totals, size = 2.5, position = position_dodge()) +
  xlab("Decade") +
  ylab("Gross Domestic Profit (in 2013 dollars)") +
  labs(title = "Gross Domestic Profit over the decades (1970 - 2013)",
       subtitle = "Looks like films that pass the test did pretty well in the 70s and 80s, \nbut their success dropped in the 90s",
       fill = "Bechdel Test?") +
  scale_fill_manual(values = wes_palette("Moonrise2", n = 2)) +
  theme_light() +
  theme(axis.text.y = element_blank())

```

Gross Domestic Profit over the decades (1970 – 2013)

Looks like films that pass the test did pretty well in the 70s and 80s, but their success dropped in the 90s



User ratings can tell us both how a movie is perceived and, for older pictures, the film's legacy as new people watch older movies and rate them on new platforms.

```

totals <- movies %>%
  group_by(decade, binary) %>%
  summarize(total = mean(imdb_rating, na.rm = TRUE))

totals$total <- round(totals$total, digits = 2)

ggplot(totals, aes(x = decade, y = total, group = binary)) +
  geom_bar(aes(fill = binary), stat = "identity", position = "dodge") +
  geom_text(aes(decade, total, label = total, fill = NULL), data = totals, size = 2.5, position = position_dodge()) +
  xlab("Decade") +
  ylab("Gross Domestic Profit (in 2013 dollars)") +
  labs(title = "Gross Domestic Profit over the decades (1970 - 2013)",
       subtitle = "Looks like films that pass the test did pretty well in the 70s and 80s, \nbut their success dropped in the 90s",
       fill = "Bechdel Test?") +
  scale_fill_manual(values = wes_palette("Moonrise2", n = 2)) +
  theme_light() +
  theme(axis.text.y = element_blank())

```

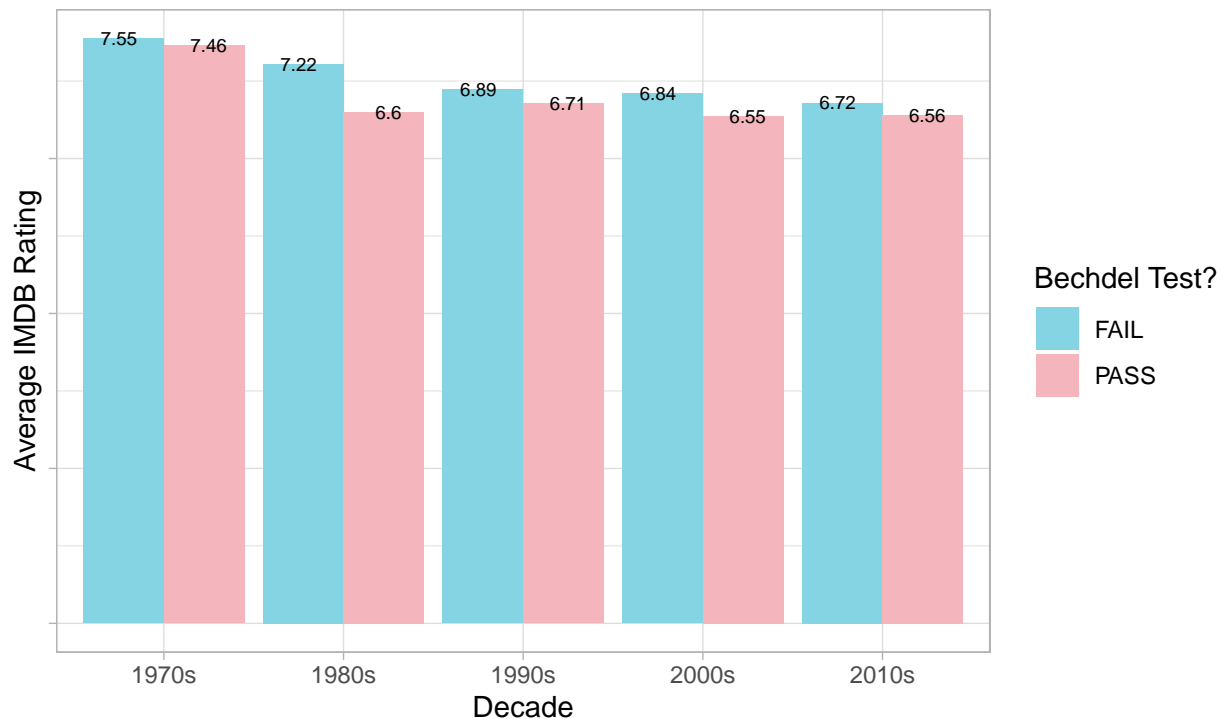
```

xlab("Decade") +
ylab("Average IMDB Rating") +
labs(title = "Film Ratings (1970 - 2013)",
      subtitle = "It LOOKS like movies that pass the Bechdel test are more poorly rated, \nbut looks c",
      fill = "Bechdel Test?") +
scale_fill_manual(values = wes_palette("Moonrise3", n = 2)) +
theme_light() +
theme(axis.text.y = element_blank())

```

Film Ratings (1970 – 2013)

It LOOKS like movies that pass the Bechdel test are more poorly rated,
but looks can be deceiving!



```

totals <- movies %>%
  group_by(decade, binary) %>%
  summarize(weighted_imdb_rating = weighted.mean(imdb_rating, imdb_votes, na.rm = TRUE))

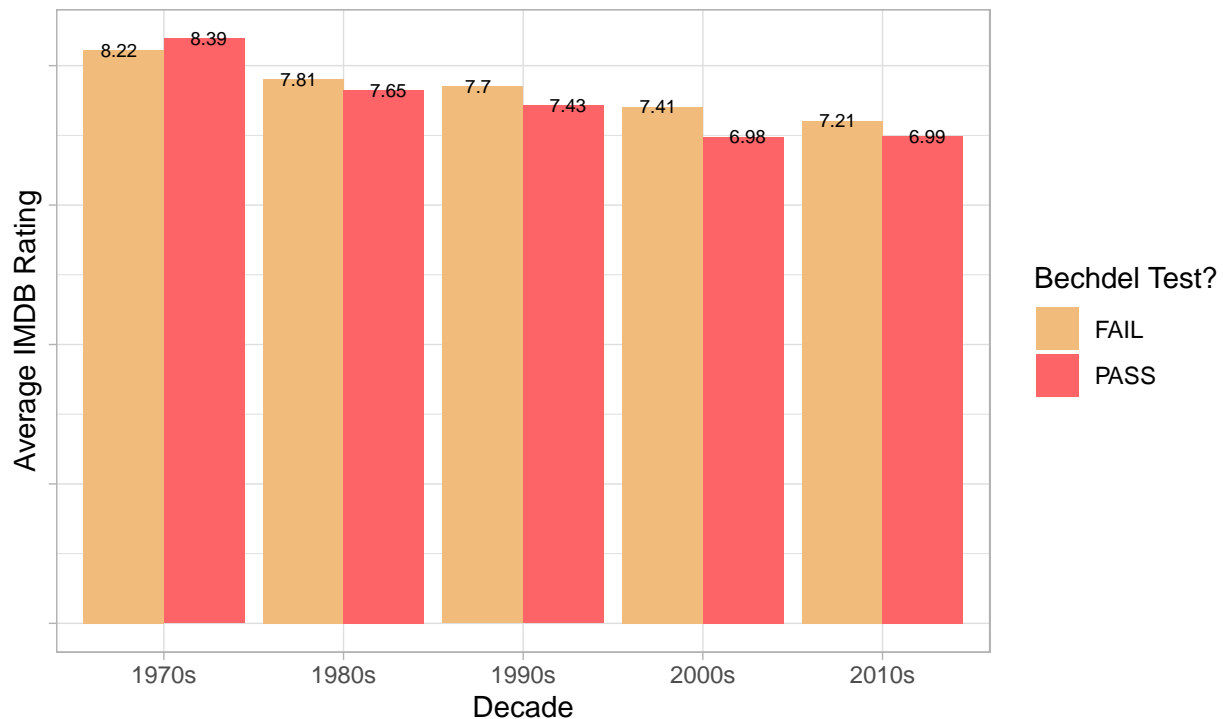
totals$weighted_imdb_rating <- round(totals$weighted_imdb_rating, digits = 2)

ggplot(totals, aes(x = decade, y = weighted_imdb_rating, group = binary)) +
  geom_bar(aes(fill = binary), stat = "identity", position = "dodge") +
  geom_text(aes(decade, weighted_imdb_rating, label = weighted_imdb_rating, fill = NULL), data = totals)
xlab("Decade") +
ylab("Average IMDB Rating") +
labs(title = "Film Ratings (1970 - 2013)",
      subtitle = "With the weighted mean, we see that scores do not vary that much \nbetween the two c",
      fill = "Bechdel Test?") +
scale_fill_manual(values = wes_palette("GrandBudapest1", n = 2)) +
theme_light() +
theme(axis.text.y = element_blank())

```

Film Ratings (1970 – 2013)

With the weighted mean, we see that scores do not vary that much between the two categories of movies



```
library(plotly)

movies <- separate_rows(movies , genre, sep = ",")

movies$domgross_2013 <- as.numeric(movies$domgross_2013)

movies <- movies %>%
  filter(!is.na(rated)) %>%
  filter(rated != "N/A")

movies$rated <- factor(movies$rated,
  levels=c("G", "PG", "TV-PG", "PG-13", "TV-14", "R", "NC-17", "X", "Unrated", "Not Rated"))

pal <- wes_palette("Zissou1", 100, type = "continuous")

p <- ggplot(movies, aes(x = rated,
  y = clean_test,
  fill = domgross_2013),
  text = domgross_2013) +
  geom_tile() +
  scale_fill_gradientn(colours = pal) +
  scale_x_discrete(expand = c(0, 0)) +
  scale_y_discrete(expand = c(0, 0)) +
  labs(title = "Bechdel Test Results vs. Rating",
    fill = "Domestic Gross") +
  theme_light() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1),
```

```
axis.title.x = element_blank(),  
axis.title.y = element_blank()  
  
ggplotly(p, tooltip = "text")
```

It's interesting to note that the highest grossing films tend to be those targeted to older kids (PG, TV-PG, and PG-13 rated films). The highest grossing of this subsection also tend to be “dubious” with regards to their test results. This is troubling because it shows that lack of representation extends beyond mainstream adult films and into children's content. Creators of media for younger audiences should be mindful of the portrayals and representations of marginalized groups in their films because of the impressionable nature of these demographics. Having diverse writers, directors, actors, and producers means a wide-range of ideas will be shared and more film consumers will be able to see themselves through the stories of people like them. This can be extremely empowering for not only those in marginalized groups, but for everyone to see positive or at the very least varied depictions of women, people of color, and queer folx in media.