DePauw University Scholarly and Creative Work from DePauw University

Honor Scholar Theses

Student Work

5-2022

Exploring Methods to Persuade Users to watch Social Problem Films through the Case Study of Race-related Films

Clay Troyer DePauw University

Follow this and additional works at: https://scholarship.depauw.edu/studentresearch

Part of the Computer Sciences Commons

Recommended Citation

Troyer, Clay, "Exploring Methods to Persuade Users to watch Social Problem Films through the Case Study of Race-related Films" (2022). *Honor Scholar Theses*. 188. https://scholarship.depauw.edu/studentresearch/188

This Thesis is brought to you for free and open access by the Student Work at Scholarly and Creative Work from DePauw University. It has been accepted for inclusion in Honor Scholar Theses by an authorized administrator of Scholarly and Creative Work from DePauw University.

Exploring Methods to Persuade Users to watch Social Problem Films through the Case Study of Race-related Films

Clay Troyer

Honor Scholar Senior Thesis, DePauw University

April 11, 2022

Abstract

The present study reports on the success of persuasive techniques and recommendation engine on persuading participants to choose a race-related movie. The study used a technical component as well as an experimental study to see how users respond to certain persuasive techniques. The technical component included a content-based filtering recommendation engine to recommend lists of movies to participants based on films they already enjoy. Persuasive techniques were then applied to those lists and results were measured and analyzed. This project aims to explore the relationship that recommendation engines and persuasive techniques have on users and their willingness to branch out in the content they consume. This study allowed me to get a baseline for how users respond to certain techniques if this study were to be done on a larger scale.

Keywords: recommendation engine, content-based filtering, persuasive technique, racerelated movie

Contents

Introduction5
Technical Background5
Recommender Systems5
Recommendation Engine for Thesis8
Persuasive Technology11
Echo Chambers12
The Wizard of Oz Technique13
Social Problem Films14
Evolution of Idea14
Evaluation15
Purpose of Evaluation15
Participant Demographics17
Study Procedure17
Data Analysis24
Discussion
Future Work
Acknowledgements

eferences

1 Introduction

Since the early 1900s, movies have had an impact on the way we live and the choices we make. Whether it was the decrease in hunting participation after Disney's *Bambi*, or the resurgence of the KKK after *Birth of a Nation* our lives are influenced by what we watch, for better or for worse (O'Toole, 2019). I have always been a movie fan and enjoyed the power of what a movie can do. The feeling I get when the credits start to roll and I am left with a different perspective, a new story, or even a new appreciation for what I watched, is a feeling I can only find in the world of film. While my project was originally conceived as a social media platform for movie viewers, utilizing a recommender algorithm to help its users, I have revised the basis for the project over a few iterations. Throughout this paper I will explain how my idea grew/changed and a detailed description/plan of its final evolution.

2 Technical Background

To properly grasp the purpose and concepts explained in this thesis project, I will provide background and summary for each of the main technical components: recommender systems/engines, persuasive technology, echo chambers, "the Wizard of Oz" technique, and social problem films.

2.1 Recommender Systems

The first and perhaps most important technical component of my thesis project is the use of recommender systems. In simple terms, a recommender system is an algorithm/engine that recommends products or services to a user (Sharma, 2018). We have become accustomed to relying on and using these systems in our daily lives, perhaps without even noticing it. Some of the most common examples include systems such as Netflix's *for you* section in which films that,

based on the user's viewing (as well as like/dislike) history, are catered to films they might enjoy. Other examples include the way Amazon recommends products their user's. Amazon will tell a user what they might like based on their purchasing history, or even based on products that buyers of that product have bought. This is where the reasoning and logic can become a bit more complicated. Recommender systems, while easy to explain on the surface, can range from very basic to very complex algorithms that use multiple data points to try and provide the best experience for their users. On the simple end of the spectrum, some recommender systems include examples such as Netflix's *top 10* section that displays the entire user bases top 10 favorite movies or shows. This follows from the logic that if many other people like it, the user might like it too.

A more complicated form of recommender system offers personalized recommendations based on each specific user and their history/preferences. There are two main types of these system: content-based filtering and collaborative filtering (Sharma, 2018). Each type uses a different method to make recommendations to the user. Content-based filtering recommends products similar to ones the user has liked or used in the past. For example, the item vector for a system that recommends movies might include attributes such as genre, director and cast members, while the item vector for a system that recommends restaurants might include attributes such as food type, cost, and the availability of wine. The algorithm then computes the cosine of the angle between the profile and item vectors, which is referred to as the cosine similarity. The cosine similarity helps the algorithm determine which movies are closer to ones the user might enjoy. The final step of content-based filtering lies in how the recommended items are returned to the user. The two basic ways of displaying the results are top-n approach and the rating scale approach. The top-n approach will show the top n of movies that are closest

to what the user might like, while the rating scale approach can display all movies past a certain similarity threshold. For example, a system that uses the rating scale approach might display all the movies that are a 90% match or higher. A major downside of these filtering methods lies in the echo chambers that they create (Sharma, 2018). Both of these methods are limited to recommending items that are similar to items or products that the user has already consumed. Later in this paper, I will discuss the power of echo chambers and the larger problem they cause.

Collaborative filtering is a bit more complex than content-based filtering. The approach is based on the principle that if person A likes the films: *Inception, Interstellar,* and *Predestination*, while person B likes *Inception, The Prestige,* and *Predestination* then theoretically, person A will also like *The Prestige,* and person B will also like *Interstellar.* This filtering method uses similar behavior to decide what films the user will enjoy, and therefore which films it should recommend. There are a few popular ways to perform this filtering in the form of user-user and item-item (Sharma, 2018). User-user collaborative filtering creates a similarity score for users in relation to other users. This can more easily be explained by inspecting Figure 1 (Sharma, 2018). In Figure 1, because the bicyclist enjoyed a salad and pizza, he also might enjoy the soda. This method of collaborative filtering uses the items the user on the right likes and computes a similarity score for the other user based on what they like in common (i.e., user to user). The second method, item-item filtering can be seen in Figure 2.

Item-item collaborative filtering is very similar; however, it computes the likeness score from item to item.



Figure 1: An example showing user-user collaborative filtering (Sharma, 2018).



Figure 2. An example showing item-item collaborative filtering (Sharma, 2018).

Recommender systems have become an integral part of our lives, by influencing the way we engage with entertainment, social media, shopping and more. They don't show any sign of going away despite some of the negative impacts, namely, echo chambers which will be discussed in a later section. Throughout this thesis, I will explore how to encourage users to branch out with their choices while using a recommendation engine.

2.2 Recommendation Engine for this Thesis

The main technical component of this creative thesis project is the recommendation system I modified. The code, while not extensive was both enjoyable to write as well as rewarding to see come to fruition. Using Jupyter Notebooks and Python, I followed an online tutorial to get a background on a basic recommendation engine and slightly modified it to suit my needs. Much of the initial work was completed during the data cleaning/processing phase as I cleaned up the given dataset, a MovieLens dataset found on Kaggle comprised of all movies released from 1874 to 2016. This phase consisted of lots of cleaning of the data and getting it ready for the code to process. I removed duplicates, unneeded rows and columns, specified necessary values and corrected typographical errors in the dataset. I also had to experiment with options to determine the most appropriate year range to use. I found that the more years of movies, the less films I knew in the output of the code as it was recommending movies based on a much broader selection. I ended up deciding to eliminate all short films, not rated or unrated films, and all films that were released before 1950. This helped to narrow down the recommended results while still allowing for a broad range for the participants to choose from.

Figure 3 shows the code for the recommendation engine after I modified it (Computer Science, 2020). This content-based filtering engine uses the film's description, removes English stop words (commonly used words such as "a", "the", "is", etc.) and creates a matrix using the cosine similarities to determine which movie descriptions are similar to each other. The cosine similarity scores are then used to sort which movies are *closer* in score to the given input movie. The final method prints out *n* movies similar to the movie in the parameter with their corresponding index from the dataset. The video from Computer Science (2020) initially used other attributes (columns) such as actor, director, genre, and year to give a narrower result of recommendations. Given that I wanted to have a race-related movie for each movie submitted

regardless of the movie submitted, I ultimately decided that the slightly broader result by just using the description would suit my intentions better for this experiment. Additionally, the original code from Computer Science (2020), included the CountVectorizer import. I decided to use a different method called TfidVectorizer as it not only takes in the count of the words in a given string but ties an importance to them as well (Saket, 2020). Given that I only used the description attribute, I thought it important to include the importance of each word in the method.

```
import numpy as np
import pandas as pd
from sklearn.feature extraction.text import TfidfVectorizer
from sklearn.metrics.pairwise import linear kernel
df1 = pd.read csv("movieDBFINALCOPY.csv")
df1['plot'].head(10)
tfidf = TfidfVectorizer(stop words='english')
df1['plot'] = df1['plot'].fillna('')
tfidf matrix = tfidf.fit transform(df1['plot'])
tfidf matrix.shape
cosine sim = linear kernel(tfidf matrix, tfidf matrix)
indices = pd.Series(df1.index, index=df1['title']).drop duplicates()
def get recommendations(title, cosine sim=cosine sim):
  index = indices[title]
```

```
sim scores = list(enumerate(cosine sim[index]))
  sim scores = sorted(sim scores, key=lambda x: x[1], reverse= True)
  sim scores = sim scores[1:65]
 movie indices =[i[0] for i in sim scores]
  return df1['title'].iloc[movie indices]
get recommendations ("The Silence of the Lambs")
    5450
                   Trauma
    3190
                 Tightrope
    5589
                Kalifornia
    5142 Shadows and Fog
    4728
             Maniac Cop 2
    4444
               Night Game
    11868
             The Gray Man
    6632
             Escape Clause
    14210
                    Lucky
    11028
                   Saw II
    Name: title, Length: 64, dtype: object
```

Figure 3: The code base for the recommendation engine used during this study.

2.3 Persuasive Technology

Persuasive technology, like recommender systems, is another apparent and important technical aspect to this thesis project. Persuasive technology, according to B.J. Fogg (2003) in *Persuasive Technology: Using Computers to Change What We Think and Do*, is "any interactive computing system designed to change people's attitudes or behaviors." (Fogg, 2003). In other words, persuasive technology is any technology that attempts to influence a user to voluntarily act or

think one way or another. Fogg is careful to define the difference between persuasion in this sense and coercion, as coercion is involuntary while the purpose of these technologies and persuasion is for the users to want to or decide to change on their own (Fogg, 2003). These actions could include clicking a link to the website of some paying advertiser, donating money to a charity, or even buying a certain product or service. This sort of technology influences us in many ways. To continue with our two examples, Netflix and Amazon, we can see the use of recommender systems in both. Netflix often displays trailers at the top of the home page that automatically play. By placing these trailers as the first thing the user sees when using the page, the user might be persuaded to add it to their watch list or even to watch the movie right there and then. Netflix will also, on occasion, allow certain categories to load before the user's, continue watching, section. This technique forces them to scroll by a whole collection of other options that may or may not catch their eye. Amazon positions the Amazon's Choice tag on certain items to help their customers pick that item over another. For example, if they are shopping for a toaster and see two comparably priced items but one with the Amazon's Choice badge, the customers will likely pick that one. In this thesis project I examine how effective some of these techniques really are, and how I can influence users to make certain choices.

2.4 Echo Chambers

Echo chambers, as mentioned in my earlier explanation of recommender systems, is a relatively easy concept to understand, as compared to the two technical concepts presented above. An echo chamber is a narrow-minded view that a person may or may not be conscious of in which can be caused when the user is only fed the same type of content that they have previously been consuming and enjoying. Let's use Netflix and recommender systems as an example. If a user really likes action movies and comedies, Netflix, with a content-based filtering strategy, will keep recommending those same movies because it will keep the user engaged. While this may not be as much of an issue in terms of films, it still takes away from the overall growth as a movie viewer. When looking at social media, oftentimes echo chambers are created on the explore/discover pages of the applications. If a user leans one way politically, their feed will be filled with those leanings' agendas, further enforcing their thoughts, and potentially polarizing them. This becomes problematic in both examples as users are no longer thinking for themselves, nor are they being challenged with new ideas and messages.

2.5 The Wizard of Oz Technique

As seen in the classic 1939 musical, *The Wizard of Oz* (Fleming, 1939), the Wizard was not a wizard at all. In the end, he was just a person, hiding behind tricks and gimmicks to make it seem as though the Wizard's machinery could do things it could not really do.

The Wizard of Oz has inspired human computer interaction researchers to use tricks to test systems that have not been fully built. The basic idea is to use manual processing to fill in for parts of a system that are not functional or do not yet exist (Marge et al., 2016). This technique in context to my study, refers to the way in which each participants survey was created. I used this theory by telling each user to send me four of their favorite movies which will then be put into my system which will then create this individualized survey with four lists catered to them and their preferences. While my recommendation system is indeed outputting the movies that are compiled on the list, due to time and technological constraints, the system is not yet able to generate nicely formatted lists of movies. Instead, I manually created each list in a photo editing software and included them in a survey for each user. The power of *smoke and mirrors* ensured the users thought my system was doing all of the work.

2.6 Social Problem Films

The final background concept to present, related my thesis, involves social problem films. This term refers to films that include a set of concerns or a problem to make social commentary on the problem at hand. These movies are also known as "social-consciousness films" or "message movies" (Doles, 2017). Essentially these films tackle some sort of social or societal problem making commentary in hopes of making a change, drawing new perspectives, or simply increasing awareness.

3 Evolution of Idea

Initially, I wanted to create a social media application that made picking more enjoyable. However, the more I thought about it, there was no problem to solve and no story behind that project other than code and an interesting final product. My favorite part of the Honor Scholar program is finding a way to challenge a status quo and tell a story by doing so and I realized I wanted my project to reflect that. My next thought, with input from my thesis advisor, was looking at how I could influence a user to pick a certain movie given persuasive technology (e.g., Amazon's Choice, primer message). While this idea was better it was still lacking a *problem* to be solved or a story to be told in any meaningful way. Additionally, this version of the project didn't address a key weakness of recommender systems as described above: echo chambers. Despite the shortcomings with this version of the idea, this iteration significantly influenced my thought process as I further changed the project idea into a question asking: how can I influence a user to pick a social problem film? During my poster presentation, I presented this idea. Students and faculty suggested to focus the question at hand even more, i.e., focus on one social problem. So, I then refined my project by seeking to answer this question: how can persuasive technology/recommender systems influence a user to choose a

race-related film that celebrates/honors black culture/experience? For the purpose of this paper, we will call those films "race-related films."

I believe this level of specificity and direction bring more meaning and importance to the story I am telling and problem I am solving. Streaming platforms often recommend films to their users based on movies the users have already watched/enjoyed. These systems also tend to recommend topical movie sections during certain parts of the year; for example, during Black History Month, Netflix and HBO Max have had black history sections containing films related to race that celebrate black culture, i.e., race-related films. But these topical movies are not personalized for an individual user. Thus, the user is forced to sift through those movies rather than being presented with movies in that specific category but catered to their movie watching preferences. For example, a user may only enjoy comedy and sport movies, so their Netflix recommended films are populated with films from those categories. When Black History Month comes around, they may want to watch a film from that category, but they must sort through all those movies rather than having comedy or sport films in this category recommended to me based on my personalized interests. This is the problem I am looking to solve. As mentioned earlier, movies are constantly influencing the way we live, so why not update the current system to help bring awareness, education, or perspectives to the average movie viewer?

4 Evaluation

4.1 **Purpose of Evaluation**

I performed a survey-based study to examine topics discussed in the previous sections. The purpose of the study is to examine how three different persuasive techniques can persuade users to choose one race-related film in a list of five personally recommended movies. The three techniques (complete with one control list) are as described below:

- Actors: The actor technique consists of images of three starring actors in each given film being placed to the right of the race-related film. An example of this technique is shown in Figure 5.
- Poster: The poster technique consists of the indicated movie's poster being placed by the title in attempts of influencing the user to choose that movie based on the poster itself.
 An example of this technique is shown in Figure 6.
- Number One: In this technique, an icon that reads "#1" is placed near the title of the movie that the system is recommending. An example of this technique is shown in Figure 4.

To prepare for the study, I modified a version of a content-based filtering recommendation engine that uses only the films' descriptions to find similar movies and recommend those to the user. The purpose of only using the description as opposed to a more specific combination of different testing attributes was to allow for the broadening of movies recommended to the user, in hopes of broadening their echo chamber.

For the study, I designed an online individualized survey for each participant. This survey consisted of demographic questions, questions regarding their opinions on recommendation engines, and finally, four lists of five movies each that act as the main measurement for the study. Each list, as more thoroughly explained in future sections, is catered to them based on the movies they sent is as their favorites. I hypothesized that the results of this study would show the actors condition being the most influential persuasive technique and would result in the most race-related films chosen, because this condition includes a visually appealing aspect as well as giving helpful information as to who might be starring in the film.

4.2 Participant Demographics

After receiving DePauw Institutional Review Board (IRB) approval for the study, I recruited participants who were at least 18 years in age, current DePauw University students, and capable of completing the survey on an internet connected device.

In total, 30 participants completed the survey. The ages ranged from 18-23 with the mean age being 20.3 years old. The participants' gender included 16 men, 12 women, and 2 non-binary. Additionally, the study included 8 Asian students, 18 White students, 2 Hispanic or Latinx students, 1 Black or African American student, and 1 Multiracial or Biracial student. The class years of the participants included 6 First-years, 7 Sophomores, 7, Juniors, and 10 Seniors. All the participants met all three conditions needed for participation.

4.3 Study Procedure

After agreeing to participate, each participant was prompted to email the researcher four of their favorite movies that were released between 1950-2016. I took one of these four movies and manually inputted it into the recommendation engine. After receiving the results from the system, I compiled a list of five recommended movies from the system results. I took four films from the top ten recommended by the system as the non-race-related films and I took the most highly recommended race-related movie that the system output to act as the fifth movie in the list. I repeated this step for each of the four movies that the user submitted, resulting in four lists of five movies each. One of the movies on each of the four lists was a race-related movie that the system recommended. Depending on the movie the participant submitted, I would have to

get more results from the engine in search for a race-related movie, even if the movie was the 200th most recommended for that case, it needed to still be recommended in some way. After these four lists were generated, the three persuasive techniques described above were randomly assigned to each list. Both the order the lists appeared in the survey as well as the order the movies appeared on each list was randomized to account for testing fatigue and any persuasion that certain orders in the list may have had on the user. This was the part of the study was where the Wizard of Oz technique was employed, as the users believed that the recommendation system was creating the entire survey for them.

When the participant submitted the four movies, they also signed up for a time to come in and complete the study. The thirty participants were distributed over four different session times. The smallest session had one participant and the largest session had six participants. Each survey was administered as a Google Form in a controlled survey room with other participants and me. The Google Form included the four lists recommended to them by the recommendation engine.

After agreeing to the terms of the study by signing an informed consent form, the participants began the survey. First, participants were briefed with some basic background information as follows:

- "What gender do you identify as?" This question was answered with one of the following options: "Male, Female, Prefer not to say, or Other" in which they were free to answer as they saw fit.
- 2. "Which of the following best describes your race/ethnicity?" This question was answered with one of the following options: "Asian, Black or African American, Hispanic or

Latinx, Native American or Alaska Native, Native Hawaiian or Pacific Islander, White, Multiracial or Biracial, or Other" in which they were free to answer as they saw fit.

- "What is your age?" This question was answered with a text box in which each participant answered with a numerical value.
- 4. "What is your year in school?" This question was answered with one of the following options: "First Year, Sophomore, Junior, or Senior."
- 5. "How familiar are you with recommendation engines? (Such as Netflix, Amazon, etc.)" This question was answered on a 5-point scale with 1 indicating "not at all familiar" and 5 indicating "very familiar."
- 6. "Do you use recommendation engines often?" This question was answered on a 5-point scale with 1 indicating "not at all" and 5 indicating "very often."
- "In general, what is your opinion of recommendation engines?" This question was answered on a 5-point scale with 1 indicating "extremely negative" and 5 indicating "extremely positive".

Next, the participants were prompted to move on to the next section of the survey titled: "Recommendation Engine Interaction." This section began by prompting the participant with the following brief message regarding the basis of this section:

"In the following section, you will complete three sets of questions, one for each of the "favorite films" you previously submitted to the researchers. Within each set of questions, you will see a list of 5 personally catered movies for you. Each list is recommended by the system from each of the 4 submitted films. You will select one film from the list of 5 as the one you "want to watch" or would pick if given the chance to actually watch them. For each movie you chose, my recommendation system selects 5 movies that you'd also probably enjoy. With this survey, I'd like to assess how well the recommendation system is doing overall and also how well it does in some finer-grained recommendations. For each set of movies, you'll see films that are presented in different formats. The movies were chosen based on a similarity score my recommendation system generates; the other presentation formats are created based on certain attributes of the movies that my recommendation system thinks you might like. After finishing each list, you will be prompted to answer some follow up questions."

This section of the survey was split into four parts, with each part presenting a list of movies based on one of the four films the participant submitted prior to the study. Each list contained a question regarding the list such as the one below, which assumes that the participant submitted "Captain America: The First Avenger" as one of the four movies they liked:

8. "Because you liked "Captain America: The First Avenger" you might also like the following... Choose one of the following movies that you would watch."

The possible answers to this question were shown in a drop-down list with each of the movies shown in the image above. This example shows the Number One persuasive technique in which *Django the Bastard* (Garrone, 1969) can be seen as the race-related film the system is persuading the user to choose. Figures 4-7 give an example of what each list might look like.



Figure 4: An example of the Number One persuasive technique being used in a recommended list.



Figure 5: An example of the Actors persuasive technique being used in a recommended list.

Cruel Intentions 3	
Beverly Hills Ninja	
Black & White	
Picture This	
The Clique	<u> </u>

Figure 6: An example of the Poster persuasive technique being used in a recommended list.

P.S. I Love You	
Cheech and Chong's The Corsican Brother	S
Tom & Thomas	
Safe Passage	
Alex Cross	
Tom & Thomas Safe Passage Alex Cross	

Figure 7: An example of the control group being used in a recommended list.

Following each list, the participant was prompted with some questions about their experience while choosing the movie in that given condition. These questions were the same for each of the four lists and are as follows:

- 9. "How confident were you in your choice? (i.e., how confident are you that you would enjoy or actually watch this film?)" The question was answered on a 10-point scale, 1 being "not at all confident" and 10 being "extremely confident."
- 10. "Did you feel the recommendation system did well in recommending movies for you based on the movie you submitted?" The question was answered on a 5-point scale, 1 being "did not recommend well at all" and 5 being "recommended extremely well."
- 11. "How was your overall feeling while choosing a movie with the help of this system?" The question was answered on a 5-point scale, 1 being "very bad" and 5 being "very good."

The purpose of these questions was to gauge how the participant felt about the choice they made, regardless of whether they chose the race-related movie or not. The final question of the survey was a follow-up question that came after each of the list sections:

12. "In the previous questions, you may have noticed certain movies were flagged with one of three conditions: the large gold #1, the movie poster, or three significant actors. Did you notice any or none of these techniques influencing you one way or another? If so, which ones worked better and why? If not, how come? What made you choose the movies you chose?" The answer for these questions was a free response text box in which the user was encouraged to give their thoughts on each of the techniques used as well as the study in general.

Once the participant answered this question, they were debriefed and thanked for their time. After all data was collected, 10 participants were selected through a random process to receive a \$25 Amazon gift card as compensation.

4.4 Data Analysis

A chi-square test of independence was completed to first determine a goodness-of-fit (onevariable method) between the frequency of race-related movies chosen in each condition. This initial test was used to determine whether there was significant independence between each condition and the number of race-related films chosen. As shown in Figure 8 and 9, out of thirty total films chosen in each condition (#1, actors, poster, control), the Actors condition had seven race-related films chosen out of thirty participants, the Poster condition had seven race-related films chosen out of thirty participants, the Control condition had three race-related films chosen out of thirty participants, and the Number One category had ten race-related films chosen out of thirty participants. Given a total of 27 race-related films chosen across all conditions, a uniform distribution would predict that each condition would predict that each condition would have 6.75 race-related movies chosen. For each condition, the Actors and Poster had an equal number of race-related films chosen, the Number One had the most and the Control had the least. While this was not have been the predicted result (the original hypothesis suggested the Actor category would result in the most race-related films chosen), it is not entirely unexpected as the control, the only group without a technique, resulted in far fewer race-related films chosen than any other. However, the chi-square test of independence showed there was no significance in the differences between each group. The p-value for the goodness of fit test was 0.2998 which is greater than the alpha value of 0.05, hence confirming the null hypothesis that there is no real dependence between race-related movies chosen and each persuasive technique.



Figure 8: Bar chart displaying the total race-related films chosen for each condition.

	Actors	Poster	#1	Control	Row Total
Total RR Films	7	7	10	3	27
Column Total	7	7	10	3	27

Figure 9: Microsoft Excel table from which Figure 8 was developed.

A chi-square test of independence was also performed to determine if there was a significant dependent relationship between different demographic groups and the number of race-related films chosen for each condition; the demographic groups examined were gender and race/ethnicity. Technically the results seen in figures 10, 11, and 12 involve expected frequencies below five, the minimum needed for chi-square, however I reported the chi-square tests to show the correct analysis approach if I were to have a larger sample size. The test for the gender group was to see if a certain gender (male, female, or non-binary) chose race-related movies in a certain condition more often.



Figure 10: Bar chart showing number of race related films chosen by each gender in each condition.

	Actors	Poster	#1	Control	Row Total
Men	3	2	6	2	13
Women	3	4	3	1	11
Non-binary	1	1	1	0	3

Figure 11: Microsoft Excel table from which Figure 10 was developed.

For each condition with the given genders, Figure 9 shows that Men and Women chose the same in the Actors condition, Women chose the most in the Poster condition, and Men chose the most for the control and #1 conditions. A few factors could have influenced this however, given that there were far fewer non-binary participants than both men and women. This explains why nonbinary participants chose the least in each of the four conditions by a seemingly large amount. The chi-square test of independence for this demographic group found there was once again no statistically significant difference between the genders and the number of race-related films in each condition. The p-value for this test was greater than the alpha value of 0.05 at 0.594, confirming once again the null hypothesis that these two variables were in fact independent of each other. Due to the vastly differing number of participants from each racial/ethnic group, a chart would not appear to give any valuable information as can be seen from Figure 12.

	Actors	Poster	#1	Control
Asian	1	1	2	1
Black or African				
American	0	1	1	0
Hispanic or Latinx	1	1	0	0
Multiracial or Biracial	0	1	0	1
White	5	3	7	1

Figure 12: Microsoft Excel table displaying the race-related films chosen in each racial/ethnic group.

The chi-square test of independence for this group was meant to see if a certain racial/ethnic group would choose race-related films in a certain category more often in one condition, given that the choice is regarding race-related films in the first place. This test resulted in the same result as the previous two in being insignificant. The p-value for this test was greater than the alpha of 0.05, at 0.695. Although this test failed to reject the null hypothesis that these two variables are independent, there were lots of outside variables/downfalls that could contribute to this insignificance. These shortcomings will be further discussed in the discussion section.

Additionally, a one-way repeated measures ANOVA (Analysis of Variance) test was completed to see if any of the four conditions resulted in a more confident choice from the participants. In other words, despite whether a race-related movie was chosen, this ANOVA was conducted to see if one or any of the persuasive techniques used resulted in a more confident answer. Figure 13 shows the participants' confidence levels for each condition on a 10-point scale. As seen in figure 13, the participants were 70% confident when choosing films in the Actors category, 65% confident when choosing films in the Poster category, 67% confident when choosing films in the control category, and 72% confident when choosing films in the Number One category.



Figure 13: A bar chart displaying the average confidence level for each condition.

Figure 15 shows the result of the ANOVA test ran on the table in figure 14. The ANOVA tests did not reveal any statistically significant difference between the four conditions (F(3,87) = 0.3898, p-value > 0.05 (alpha) at 0.761). This result confirms the null hypothesis that there is no significance between the participants confidence and each condition. Figure 14 shows the confidence results for each participant and each condition.

allowed the participants to be more confident?					
Participant	Actors	Poster	#1	Control	
1	8	8	5	8	
2	2	7	10	8	
3	7	8	6	9	
4	3	5	3	9	
5	3	6	6	2	
6	10	8	8	5	
7	10	10	5	10	
8	6	6	6	5	
9	9	10	1	7	
10	9	5	6	8	
11	2	7	7	9	
12	8	1	7	5	
13	10	4	8	6	
14	7	8	9	6	
15	10	9	6	2	
16	10	8	5	8	
17	3	1	8	9	
18	8	9	8	7	
19	9	4	10	6	
20	10	10	10	10	
21	6	9	8	8	
22	9	1	10	6	
23	9	6	8	4	
24	8	9	10	2	
25	3	3	3	10	
26	6	3	6	10	
27	5	5	10	8	
28	10	7	8	3	
29	2	8	10	3	
30	8	10	9	9	

Participant confidence in each condition: Which condition allowed the participants to be more confident?

Figure 14: A Microsoft Excel table showing the user's confidence level of each condition, on a scale of 10.

SUMMARY	Count		Sum	Average	Variance
Actors		30	21	0.7	0.08069
Poster		30	19.5	0.65	0.077069
Control		30	20.2	0.673333	0.066161
#1		30	21.6	0.72	0.056138

Figure 15: A results table from a repeated measures ANOVA test on participants' confidence in each condition.

5 Discussion

Each of the four persuasive techniques was designed with the intention of persuading each user in choosing the race-related film for that condition. These three techniques (one being a control) were chosen intuitively based on how I thought users might most likely choose certain movies. The assumption that certain actors, movies titles, and a large number one were the best techniques to use, could have very well contributed to the lack of significant difference as perhaps the users did not respond to each of these techniques as well as I originally assumed.

The level of insignificance in each of the chi-square tests of independence as well as the repeated measures ANOVA test, can be attributed to lots of factors. Perhaps the recommendation engine did not accurately recommend movies well enough. Additionally, some of the films submitted by participants were animated or children's movies which might have a very small correlation to the race-related film in that list, which could lead to a low chance of being chosen, despite which persuasive technique was being used.

Visually, the results from each of the statistical tests do show some trends. For example, in figure 9 the control was the lowest chosen category by a considerable amount. If the study were performed again on a larger scale (i.e., addressing lots of these shortcomings and using a larger sample size) perhaps these results would show significance and more valuable conclusions could be made. Not only would a larger sample size allow for the chance of more significance, but a more diverse set of participants could help as well, so that certain tests such as the chi-square test regarding race/ethnicity, might be able to be more meaningful.

Although the quantitative data did not demonstrate significant differences, the

participants answered a qualitative question regarding each technique. These responses provide

interesting insight into what their thought process when choosing. All responses can be seen in

Figure 16.

All Participant Responses to final qualitative question

I did not really notice a different, but I ended up choosing the a few with pictures. I liked the gold #1 sign, that kind of drew me in a little bit.

It didn't because it was not close at many times to the movies I would watch as I had chosen some other movies of my closest choice.

The three significant actors really drew my attention to that movie - I hadn't heard anything about it and likely would've overlooked it had I not seen actors I knew I loved. #1 didn't affect my choice at all, and the movie poster didn't make a movie stick out all that much more, although it did help me make the decision on whether or not I'd be interested quicker.

I only noticed the large gold number. I only noticed the other two after going back to them when I was making sure I chose the options I wanted to. I think I didn't notice them because I go more hard on the title or if there's a description available. I'm not one for visuals when it comes to coding a movie or show. It's about the story and that's what I'm signing up for. I'm also less likely to go for a movie just because there was an actor or actress that I liked in another movie.

The reason I chose the movies that I did was because of what I have already watched. I'm someone who is not good at watching new things unless it was recommended by someone. I don't particularly find trailers interesting or follow any recommendations on search engines due to a few prior not good experiences. I would have to be extremely bored to follow those recommendations. However I am very likely to choose the ones around the same story or characters in a sequel or something if I liked the first one. There are obviously certain genres where I feel more confident in watching based on recommendations, if I choose to do so. That would be animated movies or horror movies.

[SIC]

I think that the technique with the movie poster worked best since it gave me the best sense of the movie's tone, whereas the option with the #1 gave me no information about the film.

I picked the movies based on how well the titles seemed to align with the movies I'd

originally chosen -- for example, for the Spirited Away recommendation, I picked the "Wicked Stepmother" title because there is a wicked stepmother-type character in Spirited Away. Likewise, for the Pan's Labyrinth recommendations, I picked the movie title with the Spanish title, knowing that Pan's Labyrinth is originally a Mexican film.

movie poster was best, the names

The number 1 sign definitely drew my attention and i debated picking that one because of it [SIC]

Not really no, I found it a bit distracting for only one of the titles to have an image next to it. I chose movies based on what sounded most like the suggestion that was more closely related to the film I had already seen. I feel as though the movie poster example was the best form of image to have next to the title but It should have it next to every option on the board. [SIC]

I noticed those techniques and I think I was influenced by the large gold #1. I also think this kind of technique works better because people sometimes prefer to choose movies that most people think is good or the movie itself is popular. I think the name triggered me to choose the movies I chose. [SIC]

The #1 and the actors did not influence me because I was more likely to base whether I wanted to see the movie or not off the title. However, if they were actors that I knew, then the pictures of the actors would most likely influence my decision. The movie poster was the one that influenced me the most, as I almost picked it because it gave me a small insight as to what the movie might be about, but another title piqued my interest more.

The image with three significant characters I appreciated because I know most actors and sometimes choose the movie I am going to watch because of the actors/actresses in them. Due to this I did think about picking that movie more, but ultimately I picked the other movie because I have seen it and thoroughly enjoyed it. Also I have seen or heard about most of the movies I picked which I think was good considering the recommendations were more accurate. [SIC]

I did not even notice the large gold #1, however, the two posters in the next two options helped me not choose those movies as I figured they would not be my choice. I chose the movies which I had heard of or had watched before as well, or if their title struck out to me.

I think the image with the actors worked better for me, but that was also the only movie I recognized from the list so I don't know if the images were an actual determining factor for me. I didn't recognize most of the movies so it was somewhat random and just based off what I thought about the title. [SIC]

Those factors did affect my decision. Usually I would choose a movie based on both the synopsis + the image, but because in the previous examples none of the image looked

captivating and I didn't have the synopsis, I opted for the ones with the most interesting titles. The gold #1 did affect my decision. I typically don't care much about actors (except a few actors that I don't like so I will avoid all movies they appeared in). [SIC]

The idea of putting #1 by 12 years a slave made it stand out to me also it was the only movie i knew from the list

I noticed them, but I didn't put much weight in the images because I didn't want to make my decision off of what the computer said, but rather about myself. I judged based off of the titles, but because most of them were unfamiliar, I went with what I believed was to be best and didn't really take the computer prompting into account. I would say the picture of the actors influenced me the most because I wanted to see if I recognized the actors. [SIC]

I Disregarded the actors and number one because I didn't feel they were what I was looking for especially the actors. I chose based on the titles and what meaning I was able to infer from the titles

The large gold #1 made me think about my choice for a minute but not a lot of impact. The movie poster discouraged me from choosing that movie (it looks cheesy lol). I don't really know these actors so they didn't influence me in choosing. I think the poster works the best since it kinda gives you a sense of what is in the movie. [SIC]

The gold number one didn't do much (in fact I felt less like choosing it because of that)— I feel it was unnecessary to just have one number of it's not for a full ranking system of all the recommended movies based on how closely they match. The poster was a good idea too, since those tend to encapsulate important aspects of the film— I liked the poster image better than the three actors, though that was not bad. One having an image with the rest not having one did feel like the image movie had more credibility and background, giving me more information about it before choosing. Overall I chose two movies I knew I liked and two I had never heard of—just because for those, the titles sounded interesting. [SIC]

1. I did notice the techniques and the movie posters did a better job at influencing me than the other two. I think that out of the three way of influencing me the movie poster was the best one because it gave me a sense of the content. I chose mainly off of the title of the movie and what sounded interesting to me.

I don't think these influenced my option very much because I am not really interested in actors as much as I am interested in the content of the movie. I also would've loved to see like emmy award nominee or awards given to a film as something that could potentially influence me. [SIC]

All of the movies that I chose I have previous knowledge about, so the flags did not make much of a difference. [SIC]

The #1 symbol led me to choose that movie because it made me feel like that was the #1 recommended movie based on my pick.

I noticed the #1 sign and felt that it was a little too cliche of a descriptor and that specific condition drew me away. The movie poster and top actors were much more helpful in helping me decide if I'd like that movie. Those pictures gave me a better idea if I'd like the movie, and based on those, I decided not to pick the ones with the pictures and instead gamble that I'd like a different movie based on the name. By the last set of options, I missed not having the visual aid/descriptor to help me pick my movie from the 5 provided.

The number, image and actor names didn't really influence my choice at all. This is since I chose the movies simply based on the name. The flash and big name doesn't really guarantee a movie is gonna be successful; just look at Cats. [SIC]

The large gold #1 definitely influenced me, especially since I did not recognize any of the listed movies. Usually the movie poster would tell you more about the film, so normally I would be intrigued. However, I did recognize another movie listed so I chose that one since I knew what it was about. The three actors did not influence me either since I have seen The Green Mile and I know I like it. It was strange seeing The Green Mile being recommended for Tangled since the were different genres, but I do enjoy them both. [SIC]

The additional features did not influence my decision making at all. For me, choosing movies was based primarily on: either having heard about the movie before, or coming across a very interesting title of the movie. [SIC]

The movie poster influenced me the most as that is usually what I look at when deciding on a movie to watch. The gold 1 interested me the second most, as it drew me to the title of the movie, and that title then intrigued me. Finally, the actors drew me the least as I don't usually pay attention to which actors are in a movie when I decide what to watch.

The one with the three significant actors influenced my decision positively. It definitely made that choice more appealing, knowing there were actors that I liked in the movie. The one with the movie poster influenced my decision negatively. I did not like the poster, so it made me want to watch the movie less. The other movies, I chose based on what I thought they would be about from the name. [SIC]

The different conditions motivated me to consider the other options more closely compared to what I intially predicted that I might enjoy. The large gold #1 didn't have a significant impact compared to the others, which had a more appealing display related to parts of movies that I liked (the poster and actors told me more about my interests than the #1). The movies that I chose were based on mainly sequels or words in the titles that were significant themes from the original movie, so I knew that even without much other context as to what the movies were about there was at least a set of themes that I identified with. [SIC]

Significant Actors

Figure 16: Table displaying all the users' responses to the final question.

While these responses varied widely, they also provided interesting feedback about future work, opinions on the persuasive techniques and even information that could be seen as reasons why the data was not significant. For example, some of the participants said they liked the movie poster technique, but that specific movie poster they found cheesy or uninteresting. This means there wasn't a flaw on the technique itself rather the specific use case. Some responses even gave ideas for techniques that would have influenced them more such as awards won.

Despite the lack of statistical significance in the test results, intuition and participants' qualitative responses gives hope for the success of a future study that addresses the shortcomings of this pilot study for one reason or another. The techniques were successful in persuading certain users and there were favorites as well, but ultimately, a new study would need to be done to further analyze and conclude regarding the success and confirmation of the initial hypothesis.

6 Future Work

As mentioned in the discussion section, additional work based on more robust data would be informative. It is even intuitively promising that a future study could result in a statistically significant conclusion.

Given the nature of this technical project, the first step in improving the study, would be in improving the recommendation engine itself. This study used a content-based filtering approach in which the movies were recommended to the user based on what they liked in the past (i.e., their four submitted movies). The system used the film description as its sole similarity attribute which could result in a broader and less accurate result in the actual movies recommended. In a future study using content-based filtering, this system would be written with more than just one attribute to find the cosine similarity. The one attribute approach becomes an issue, because it leaves room for many outliers. For example, if the movie *Shrek* uses the term monster in its description, that keyword will likely be found in lots of horror movies which wouldn't be an accurate recommendation. Users found this issue apparent during the study. Many of the questions participants had while taking the survey and, in the comments, shown in figure 16, related to not knowing or recognizing any of the movies in the lists they received. This initial variance was supposed to be a good thing as it would allow to break out of the echo chamber ever so slightly, but perhaps it caused too much variance. Reeling this in by including attributes such as actors, director, genre, and rating, could allow for a more accurate engine and perhaps more accurate lists for the users; hence taking away the issue of not knowing or having heard of any of the movies before.

Furthermore, using a collaborative filtering approach where users would have accounts on a social media style platform with their preferences stored in their respective profiles, could allow for more accurate recommendations. In this example, collaborative filtering could be a good addition because it would allow for more data points for each user, resulting in more accurate recommendations. In this example, one could conduct a study over the course of months as they used this system and tracked the movies they watched, liked, and disliked, creating a data base of information for the collaborative approach to recommend movies based on ones that were similar to movies other users liked etc.

Both alternative approaches along with a recommendation engine testing phase where the system itself would be experimentally tested and deemed proficient, would likely add to the

integrity of the study and more weight to the race-related films and persuasive techniques and rather than the recommendation engines themselves.

Additional improvements to future studies could include the subject of the persuasive techniques themselves. Rather than one or two people intuitively choosing which techniques might be more accurate, a study could be run to determine which persuasive techniques could be the best options to test with the recommendation engine. These other techniques could include options such as: film descriptions, year of release, film trailers, awards won/nominations, etc. These research-backed techniques could reduce the concern of whether the persuasive techniques were largely helpful and allow the focus to turn on which of these techniques better persuaded users to choose the race-related movie.

Finally, allowing for a larger and more diverse sample size paired with the automation of the creation of each survey would allow for more accurate and meaningful results. Although the results from the chi-square tests were insignificant, if the trend continued on a larger scale, the theoretical calculations *would* be significant, and a clear dependence could be seen. The sample size used for this project was just too small to make those claims based on the data I had. Each survey was customized for each person, and it became time consuming to create each one. Given the tight schedule of the honor scholar thesis dates, I could not allow for more than the allotted participants. Ways to do this could include an automation system for randomizing lists and orders, complete from a more expansive data base including actors' images, movie posters, trailers, and the information needed for whatever persuasive techniques used. While I could also just allow for a longer timeframe, the automation would also account for any unconscious human error made while randomly selecting and creating each survey.

With the inclusion of one or any combination of the previous future suggestions, it is clear that another study could be run with more confidence and integrity to supplement this pilot study and its results.

7 Acknowledgements

This material and study is based on the work done by principal investigator Clay Troyer as well as thesis committee members Dave Berque, Steve Bogaerts, and Michael Roberts. Additionally, thanks are to be given to each of these persons as the work and discussions found in this material would not be possible without their advisement and leadership. The DePauw Computer Science Department Founder's Fund provided funding for the participant compensation, and I am grateful to the donors who made this support possible.

References

APA Dictionary of Psychology. American Psychological Association. <u>https://dictionary.apa.org/</u>

[Computer Science]. (2020, June 5). Build A Movie Recommendation Engine Using Python [Video]. YouTube.com. https://www.youtube.com/watch?v=ueKXSupHz6Q&t=863s

Doles, S. (2017). Social Problem Films. Oxford Bibliographies. DOI: 10.1093/OBO/9780199791286-0161.

- Fogg, B. J. (2003). Persuasive Technology: Using Computers to Change What We Think and Do. Morgan Kaufmann Publishers.
- Leroy, M. (Producer), & Fleming, V. (Director). 1939, August 25. *The Wizard of Oz*. United States: Metro-Goldwyn-Mayer.
- Marge, M., Bonial, C., Byrne, B., Cassidy, T., Evans, A. W., Hill, S. G., & Voss, C. (2016).
 Applying the Wizard-of-Oz Technique to Multimodal Human-Robot Dialogue.
 International Symposium on Robot and Human Interactive Communication, 930.
 https://doi.org/10.48550/arXiv.1703.03714

Martino, P. (Producer), & Garrone, S. (Director). 1969. Django the Bastard. Italy: S.E.P.A.C.

- O'Toole, K. (2019, January 29). *The Power of Cinema: 10 Films That Changed the World*. Raindance. <u>https://raindance.org/the-power-of-cinema-10-films-that-changed-the-world/</u>.
- Saket, S. (2020, January 12). Count Vectorizer vs TFIDF Vectorizer | Natural Language Processing. LinkedIn. https://www.linkedin.com/pulse/count-vectorizers-vs-tfidf-natural-

language-processing-sheel-

saket#:~:text=TF%2DIDF%20is%20better%20than,by%20reducing%20the%20input%20d imensions.

Sharma, P. (2018, June 21). Comprehensive Guide to Build a Recommendation Engine from Scratch. Analytics Vidhya.
<u>https://www.analyticsvidhya.com/blog/2018/06/comprehensive-guide-recommendationengine-python/</u>