User Trust in Recommendation Systems: A comparison of Content-Based, Collaborative and Demographic Filtering

Mengqi Liao
Media Effects Research Laboratory, Penn State University, University Park, PA, USA
mul914@psu.edu

S. Shyam Sundar
Media Effects Research Laboratory, Penn State University, University Park, PA, USA
sss12@psu.edu

Joseph B. Walther
Center for Information Technology & Society, University of California, Santa Barbara, CA, USA
jwalther@ucsb.edu

ABSTRACT

Three of the most common approaches used in recommender systems are content-based filtering (matching users' preferences with products' characteristics), collaborative filtering (matching users with similar preferences), and demographic filtering (catering to users based on demographic characteristics). Do users' intuitions lead them to trust one of these approaches over others, independent of the actual operations of these different systems? Does their faith in one type or another depend on the quality of the recommendation, rather than how the recommendation appears to have been derived? We conducted an empirical study with a prototype of a movie recommender system to find out. A 3 (Ostensible Recommender Type: Content vs. Collaborative vs. Demographic Filtering) x 2 (Recommendation Quality: Good vs. Bad) experiment (N=226) investigated how users evaluate systems and attribute responsibility for the recommendations they receive. We found that users trust systems that use collaborative filtering more, regardless of the system's performance. They think that they themselves are responsible for good recommendations but that the system is responsible for bad recommendations (reflecting a self-serving bias). Theoretical insights, design implications and practical solutions for the cold start problem are discussed.

CCS CONCEPTS

• **Human-centered computing**; • **Human computer interaction (HCI)**; • **Empirical studies in HCI**;

KEYWORDS

Personalization, User Experience Design, Empirical study that tells us about how people use a system

ACM Reference Format:

Mengqi Liao, S. Shyam Sundar, and Joseph B. Walther. 2022. User Trust in Recommendation Systems: A comparison of Content-Based, Collaborative and Demographic Filtering. In *CHI Conference on Human Factors in Computing Systems (CHI '22), April 29–May 05, 2022, New Orleans, LA, USA.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3491102.3501936

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

CHI '22, April 29–May 05, 2022, New Orleans, LA, USA
© 2022 Association for Computing Machinery.
ACM ISBN 978-1-4503-9157-3/22/04…$15.00
https://doi.org/10.1145/3491102.3501936

1 INTRODUCTION

With the rapid diffusion of artificial intelligence (AI) in all walks of life, the notion of explainable AI (or XAI) has gained new prominence in an effort to enhance users' understanding of—and trust in—AI [4]. This is particularly the case for recommendation systems that tend to be more user-facing and popular, such as the ones used by Amazon, Netflix, Google, and Facebook. Extant research focuses on the actual performance of recommendation systems, and how to improve algorithms and computational models to predict users' needs and preferences more accurately, under the general assumption that better recommendations will ultimately lead to better user experience [22]. However, recent studies show that ensuring good user experience is far more complicated, with many other factors influencing users' subjective evaluations of recommendation systems [10, 23]. One of the factors to which researchers have dedicated extensive study is the different approaches that the recommender systems use [55]. However, despite the difference in terms of their objective performance, we do not know whether users have a general subjective tendency to trust one type of recommender system more than others [21, 27].

Users usually become aware of the approach that the recommender system uses when a system provides a single-sentence description that hints at the underlying mechanism of the recommender system. For instance, in some e-commerce scenarios, users are told that the product is recommended based on their purchase history or is similar to the products that they have liked before. In other cases, they are told that the products are recommended based on preferences or behaviors of similar users. Other recommender systems will explain to users that the products are recommended based on some demographic factors (e.g., age, gender, etc.). These explanations describe content-based filtering, collaborative filtering, and demographic filtering approaches, respectively. Previous research has found that users tend to evaluate the recommender system using different approaches differently as well, even when the system delivers the same performance [27].

But what happens when a system delivers unsatisfactory performance, which might not be uncommon at the early stage of users' interaction with the system (i.e., the cold start problem, [28])? Will users be less likely to trust the system, and abandon the system if the recommendation comes from any one type of recommender system? Will they blame the system more when a particular type of recommender system delivers the unsatisfactory performance? By focusing on three common approaches that recommender systems use (i.e., collaborative, content-based, and demographic filtering), the current study aims to answer these questions through an experiment that explores the psychological factors underlying users' perceptions of recommendation systems. Unlike previous research that has used scenario-based experiments to gauge responsibility attribution in HCI (e.g. [3, 33]), we designed a prototype of a movie recommender system which allowed users to have real interactions with the system. In order to investigate the psychological effects of different explanations in regard to different approaches used by the systems, we controlled for users' interaction processes and the system's recommendations. Specifically, the recommender system was designed to deliver either good or bad performances so that we can empirically investigate whether users trust one of these systems over another, and if such trust depends on the quality of the recommendations.

We conducted a 3 (Recommender system type: Content vs. Collaborative vs. Demographic Filtering) x 2 (Recommendation Quality: Good vs. Bad) experiment (N=226) to test the psychological effects of explanations that describe different types of recommender systems. We found that users trust collaborative filtering more regardless of the system's performance. If the system delivers bad performances, they think it is the system's responsibility. However, if the system delivers good recommendations, they think they are more responsible, indicating the existence of a self-serving bias. Based on Attribution Theory [17] and Computers Are Social Actors (CASA; [41]), we hope to extend the current literature regarding users' interactions with intelligent systems, as well as to provide design implications for practitioners in improving user experience with recommender systems.

2 RELATED WORK

This section reviews literature pertaining to psychological effects of recommender systems, in order to develop hypotheses and research questions for study.

2.1 Users' trust towards different approaches among recommender systems

The core part of a recommender system is the specific approach that it uses to generate recommendations. Three of the most common approaches used by recommendation systems are content-based filtering, collaborative filtering, and demographic filtering, due to the simplicity of acquiring users' information, preferences, past purchase history and ratings [39]. Specifically, a demographic-based recommender system relies solely upon users' demographic profiles (i.e., age, gender, location, education, etc.) [38], and recommends products that are liked by users who are similar in terms of these demographic factors. Collaborative filtering also relies on user profiles and finding linkage among users' profiles, but it requires more implicit user input compared to demographic filtering because it analyzes the correlations among users based on users' potential common interests in products (expressed in terms of past purchases, ratings, browsing history, etc.) [42]. As a result, collaborative filtering will recommend products that users who share similar preferences and tastes have liked or consumed before. Content based filtering, similar to collaborative filtering, makes inferences about users' interests based on their indicated preferences of products, but instead of analyzing the correlation between similar users, it analyzes the correlation between users' preferred product characteristics with those of new products, and recommends products accordingly [38]. Systems usually explain to users about how they work, in one or two sentences on the interface, e.g., "Based on your past selections, you might be interested in..." for content-based filtering; "Other users with similar interests also enjoyed..." for collaborative filtering; and "Other users in your age group also enjoyed..." for demographic filtering.

Despite the actual computational differences between these approaches, the simple explanations of how the systems work can have significant psychological effects on users' trust towards the system [27]. For instance, Liao and Sundar [27] found that even when systems provide the same results, users tend to differentially evaluate recommender systems that use collaborative filtering and those that use content-based filtering. Hence, in addition to the objective effects of different approaches of recommender systems on recommendation quality [9], how users subjectively evaluate different approaches needs to be investigated further. Users usually get a sense of the approach used by the system by reading the explanations provided on the interface. The reason for perceived differences across different approaches of the recommender system is that each triggers a distinct cognitive heuristic or mental shortcut about the nature and quality of the recommendation.

One of the heuristics that might be applicable in the current study's context is the bandwagon heuristic, which is the rationale that "if others think something is good, then I should think so, too" (p.83) [49]. When interacting with collaborative filtering recommender systems, users might hold greater bandwagon perceptions, since products are recommended based on a large group of users who share similar tastes and preferences with the individual user. The bandwagon heuristic tends to be prominent in the context of e-commerce. For instance, both Liao and Sundar [27] and Sundar et al. [52] found that users tend to trust the collaborative recommendations more than content-based filtering because of the triggering of bandwagon heuristics. An important distinction between demographic filtering and collaborative filtering recommender system needs to be made here. Although they both recommend products based on finding connections among similar users, the degree of similarity is more precise with collaborative filtering because the system analyzes the degree to which an individual is connected with a set of users that shares similar tastes and preferences. Therefore, users' sense of crowd and collective endorsement will be more salient while interacting with collaborative filtering compared to demographic filtering. Based on these principles, we propose the following hypothesis:

H1: Recommendation systems with collaborative filtering will lead to greater bandwagon heuristic perceptions among users, which in turn lead to more positive evaluations and trust of the system, in comparison to systems with demographic filtering or content-based filtering.

Although being part of a social unit is an inherently human tendency and a natural need, we also have a desire to be different from others and to be treated as unique individuals [57]. Therefore, whether or not recommended products reflect one's self-concept may play a role in the degree to which users would accept recommendations either based on other users of a service or themselves solely. That is what the content-based filtering approach entails. Content-based recommendation systems might give individuals a greater sense of personalization by tailoring products to fit one's specific needs, compared to demographic or collaborative filtering. Without mentioning other users, content-based filtering implies that recommendations use one's own preferences and usage history as the primary determinants of recommendations, and are therefore more likely to trigger the identity heuristic, which further leads to more positive evaluative outcomes toward the system. Therefore, the following hypothesis is proposed:

H2: Recommendation systems with content-based filtering will lead to greater identity heuristic perceptions among users, which in turn lead to more positive evaluations of the system, in comparison to systems with demographic filtering or collaborative filtering.

2.2 How users assign responsibility for performance of intelligent systems

While many scholars have called for more responsible and explainable AI (e.g., [2]), especially under circumstances of failure, i.e., when something goes wrong with AI [7], sometimes the operations and outcomes of AI are not veridical with users' lay understandings of how the systems operate. Therefore, it is imperative to understand how users themselves attribute responsibility for systems' failure and success while interacting with AI, both in conjunction with, and independently of, AI's actual performance [13]. Knowing how users understand reliable and trustworthy systems can be very important not only in order to design trustable systems [13], particularly in the initial stage of users' interaction with the systems when systems know relatively little about users and are most likely to make mistakes. It also has critical implications for the design of narrative explanations and illustrations that accompany users' introduction to an intelligent system. However, limited empirical research has been focused on how users attribute responsibility to intelligent systems. Among these few empirical studies, the vast majority focus on either embodied intelligent agents (e.g., robots and autonomous cars; see [19, 26, 29]) or how users assign responsibility in moral scenarios (e.g., [3, 33]).

There is limited research shedding light on how users assign responsibility to specific AI systems such as the three different types of recommender systems. However, we can draw some insights from social psychology research to speculate how users react to AI systems. According to the CASA paradigm [41], humans have a natural tendency to apply social rules while interacting with intelligent technology, because we have not yet developed a set of distinct responses towards technologies. CASA researchers have found that we extend social expectations and human stereotypes towards machines (e.g., gender stereotypes, [35]), being polite while interacting with machines [36], and orienting our responses towards the machines themselves instead of the programmers behind the machines [51]. Therefore, social psychology theories offer fruitful theoretical frameworks to help us explain users' psychological perceptions and behaviors when it comes to assigning responsibility.

In social psychology, one of the most prominent theories to explain how we assign responsibility to others and ourselves is Attribution Theory [17, 18]. According to this theory, we tend to think of our own failures as being due to aspects of the situation (including other people), while we credit our own selves for our successes. Regardless of success or failure, we tend to think that others' behaviors and decisions result from their own individual characteristics [43]. More specifically, we hold a self-serving bias [1], namely, the tendency to make internal attributions for our successes and external attributions for our failures. The reason behind such a tendency is largely due to our needs for self-enhancement in order to maintain our self-esteem [53].

In the context of human-computer interaction, researchers have found that we also tend to hold the self-serving bias while interacting with machines. In a series of experiments exploring CASA, Nass et al. [34] found that when we treat computers as our teammates, we will blame the computer more if a collaborative result is undesirable, while crediting ourselves more when the result is desirable. Similarly, Serenko [46] found that users are more likely to blame a computer's operating systems' user support agent if they fail to achieve successful outcomes, but assign themselves more responsibility when the outcomes are successful. Applying this to an intelligent technology that tends to have more agency over the decision-making process, Hong [15], in a scenario experiment, found that we blame self-driving cars more than the human drivers when accidents happen. Therefore, it is reasonable to hypothesize that users are also likely to show a self-serving bias while interacting with recommender systems. In other words, users might blame the system more if they receive poor recommendations, while crediting themselves more when the recommendations are good. In addition, it is logical to predict that users will generally provide more positive evaluations of the systems that provide good quality recommendations compared to systems that recommend poor quality products. Hence, the following hypotheses are proposed:

H3: Users will evaluate the recommender system more positively and trust the system more if the system provides recommendations of high quality compared to recommendations of low quality.

H4a: Users will assign more responsibility to the recommender system when the system provides low quality recommendations compared to high quality recommendations.

H4b: Users will assign more responsibility to themselves when the system provides high quality recommendations compared to low quality recommendations.

2.3 How users assign responsibility when interacting with different recommender systems

Researchers investigating the influence of self-serving bias during users' interaction with technology have found other factors that affect judgments about humans and technology as well. For instance, the degree of machine agency matters. Psychologists revealed that we blame other people more if we know they have full control over their behaviors [56]. Similarly, we blame robots that are more autonomous [19], and blame robots less if they are only responsible for following directions from human programmers [21].

Applying these insights to users' interaction with recommender systems, we expect users to attribute responsibility differently while interacting with systems that use different approaches. In general, demographic and collaborative filtering all involve third parties, that is, similar users (either in terms of demographic similarity, or similarity in preferences), while content-based filtering relies primarily on users' own preferences. Therefore, when users interact with content-based filtering, they might perceive it as having greater agency and control over the recommendations compared to the other two approaches because content-based filtering is immune to influence from similar users. As a result, individuals are more likely to assign responsibility to the system when interacting with a content-based filtering recommender system. In a similar vein, the collaborative filtering recommendation system has an additional persuasive appeal compared to other recommender systems: the social endorsement from similar users [27]. A recent meta-analysis found that social information is a major source of information that users rely on when interacting with recommender systems [37]. Quijano-Sanchez et al. [40] found that greater tie strength between neighbors (similar users) and the individual user might predict greater satisfaction with recommendations. Sometimes, users will follow similar users' recommendations even if there is an ill fit between the recommended products and the system [6]. Therefore, it is possible that users will rely more on similar users while making decisions and are hence less likely to assign responsibility to themselves or the systems when interacting with systems that involve similar users, leading us to propose the following hypothesis:

H5: Users will assign more responsibility to themselves and the recommendation system when interacting with content-based filtering recommendation systems, compared to collaborative and demographic filtering recommendation systems.

Other research has found that our degree of self-serving bias also depends on our expectations of a machine's performances and intelligence. Malle et al. [31] found that we are more likely to assign greater responsibility to robots when encountering negative performances if that robot has a pre-existing reputation of having higher intelligence. Since demographic filtering is the most basic approach of the recommender system, and might be perceived as being less intelligent compared to the other two approaches, we might see less self-serving bias when users interact with it. Hence, we propose an interaction effect between system type and a system's performance on systems' evaluation in the following hypothesis:

H6: The quality of a recommender system's performance will moderate the effects of system types on the responsibility users attribute to themselves and the recommender system.

Finally, in interpersonal interaction literature, researchers generally found that attributions of responsibility impact the trust we have towards others (see [8, 45]). Whether attribution of responsibility becomes a significant underlying mediator of users' trust towards recommender systems is relatively under-investigated. Therefore, we pose the following research question:

RQ1: Will attribution of responsibility mediate the relationship between system types, system performance, and users' evaluation of their trust in the recommender system?

3 METHOD

To test our hypotheses and answer the research question, we conducted a 2 (Performances: Good vs. Bad) x 3 (Recommender system type: Content-based filtering vs. Collaborative filtering vs. Demographic filtering) between-subjects online experiment.

3.1 Participants

In late February 2021, we recruited 235 participants from Cloud Research, which provides access to high-quality respondents on the Amazon's MTurk platform. We embedded three attention check questions in our questionnaire to filter out low quality data, eventually yielding 226 participants who answered all three attention check questions correctly. The average age of our participants is 41.52 (SD=12.73), and 48.0% of them are female (N=107) and 51.6% are male (N=115). They are relatively highly educated, with 68.7% of them having a bachelor's degree or above (N= 153). Majority of our participants are Caucasian (80.1%, N=169), followed by 9.0% Asian/Pacific Islander (N=19), and 6.2% African American (N=13).

A power analysis was conducted using an alpha of 0.05, a power of 0.8, and a medium effect size (f=0.25) [11]. With 6 conditions and 4 degrees of freedom (2 control variables), the desired sample size is 196, which is smaller than our sample size, indicating sufficient statistical power for our analyses.

3.2 Procedure

Participants were first asked to answer a pre-stimuli survey capturing their demographic information, and the control variables. They were then randomly assigned to one of the six experimental conditions and instructed to interact with the corresponding prototype of one of the ostensibly different recommender systems. That is, unbeknownst to the participants, the recommender systems did not really operate based on any of the three filtering mechanisms. The difference among these conditions was the way the systems were labeled and described as working.

On the home page, we first provided a brief introduction of the recommender system (which we named "Movie Taste," modeled after the mobile app Taste [54] where they received the manipulation of the ostensible system type in the explanation for the first time (see Figure 1 which provides a screenshot of the welcome page for collaborative filtering conditions). After this initial introduction, participants assigned to interact with Content-based filtering and Collaborative filtering recommender systems were asked to provide ratings for 10 popular movies (i.e. *Titanic, Forrest Gump, Clueless, The Lion King, Iron Man, Anchorman, Lady Bird, Inception, Pirates of the Caribbean, Love Actually*) (see Figure 2). We selected these movies from IMDB top 100 movies so that majority of the participants would be familiar with these movies and were hoping to cover a wide range of movie genres. For participants assigned to interact with Demographic filtering recommender systems, the system simply asked them to provide their demographic information (i.e. age, gender, ethnicity and education level).

After the profiles were successfully built, study participants were directed to the loading page, where they received the manipulation for system type for the second time (see Figure 3). The system then generated five different movie recommendations (either good ones or bad ones) with the explanations of how they are generated to explain to participants for the third time (See Figure 4). We manipulated these explanations corresponding to each of the recommender system type (See Stimuli for more details). After browsing through the generated recommendations (5 different movies, with description and trailers), we provided an end code for participants to input in order to access the post-stimuli questionnaire.

Finally, in the post-stimuli questionnaire, we asked questions regarding the mediators and our dependent variables. The study

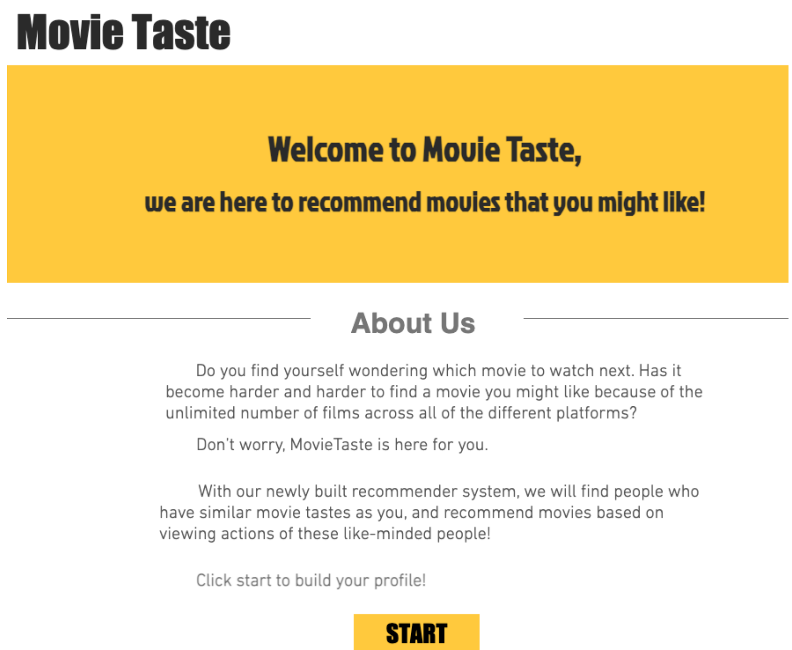

Figure 1: Home page of Movie Taste (Collaborative filtering conditions)

Figure 2: Profile building page of Movie Taste (Collaborative and Content-based filtering conditions)

took an average 16.6 minutes to complete, and participants received $1.2 upon completion of the study.

3.3 Stimuli

3.3.1 Manipulation of systems' performance. To manipulate the recommender systems' performance, we conducted a pilot test using a separate set of participants in order to choose the movie recommendations. In the pilot test, we selected 9 Good movies from the top-rated IMDb movies, and checked their ratings on Rotten Tomatoes to make sure they have a Tomatometer of 95% or above. We then selected bad movies from a pre-existing study, in which Hou et al. [16] systematically calculated the likes vs. dislike ratios of 725 movie trailers on Youtube. We selected 13 bad movies, that scored the lowest on the list.

We then recruited 61 Participants on Cloud Research in the US (average age = 39.70, *SD*=12.24; females = 45.9% and males = 54.1%, and 62.3% Caucasian). They were randomly assigned to watch 7 of the 22 movie trailers and were asked to indicate whether they have watched the movie before. They were also asked to answer 10 different questions assessing perceived quality of the movie and their watching intention on a scale from 1 to 7. Sample questions include "I would like to see this movie sometime in the future", "This movie does not look appealing to me (R)", and "I will probably like this movie". The criteria we used to choose the movies are 1) the good movies should score significantly higher than the bad movies. 2) at least 60% of the participants have not seen the movie to ensure the novelty aspects of movie recommendations in the real study. We finally selected 5 movies for good performance and 5 movies for bad performance manipulation in the real study (See Table 1 below for a summary of the ratings given to the 10 selected movies).

3.3.2 Manipulation of recommender system types. The manipulation of the system type was embedded in the explanations that the respective systems presented to users. As mentioned previously, we provided explanations of the systems three times: on the home page, the loading page, and the results page (See Table 2 for a summary of the manipulation).

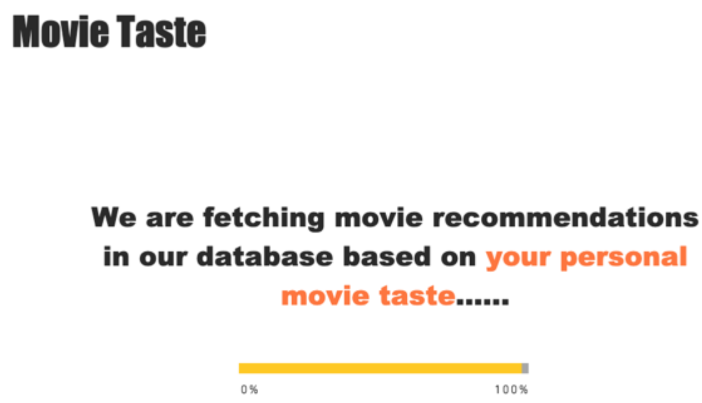

Figure 3: Loading page of Movie Taste (Content-based filtering conditions)

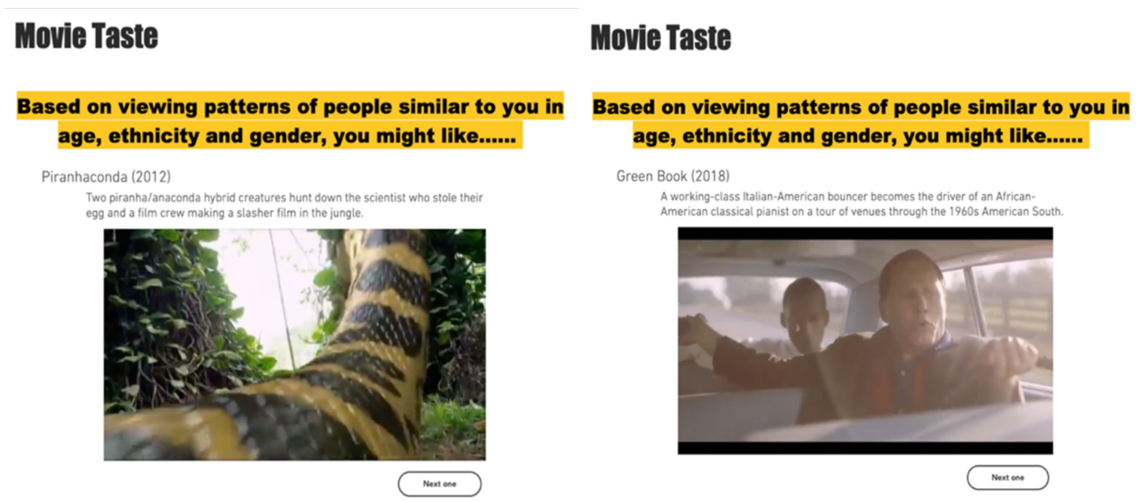

Figure 4: Results page for Demographic filtering conditions (left: bad recommendations; right: good recommendations

Table 1: Summary statistic for 10 selected movies (results from the pilot test, N=61)

Performance (in the main study)	Movie name	Perceived quality	% who have not seen the movie before
Good	Green Book (2018)	5.65 (SD=.95)	75.0%
Good	Parasite (2019)	5.46 (SD=1.19)	94.8%
Good	Spider-Man: Into the Spider-Verse (2018)	5.33 (SD=1.27)	70.8%
Good	The Wolf of Wall Street (2013)	5.30 (SD=1.27)	61.6%
Good	The Godfather (1972)	5.28 (SD=1.24)	66.7%
Bad	Piranhaconda (2012)	2.83(SD=1.90)	100.0%
Bad	Blubberella (2011)	2.89(SD=1.60)	93.3%
Bad	The Hottie and The Nottie (2008)	3.03(SD=1.61)	83.3%
Bad	Sand Sharks (2012)	3.04(SD=1.69)	94.1%
Bad	Chairman of the Board (1998)	3.10(SD=1.75)	81.0%

Table 2: Manipulation of the system types descriptions

Place of manipulation	Content-based filtering	Collaborative filtering	Demographic filtering
Welcome page	With our newly built recommender system, we will build a personal movie taste profile for you and recommend movies you might like based on your unique taste!	With our newly built recommender system, we will find people who have similar movie tastes as you, and recommend movies based on viewing actions of these like-minded people!	With our newly built recommender system, we will find people who are similar to you in age, ethnicity and gender and recommend movies based on their preferences!
Loading page	We are fetching movie recommendations in our database based on your personal movie taste......	We are fetching movie recommendations in our database based on the viewing actions of people who share similar movie tastes with you......	We are fetching movie recommendations in our database based on movie preferences of others like you......
Results page	Based on your preferred movie characteristics and your personal movie taste, you might like......	Based on viewing patterns of people who share similar movie tastes with you, you might like......	Based on viewing patterns of people similar to you in age, ethnicity and gender, you might like. . . . . .

3.4 Measurement

Except for the manipulation check questions, which were measured on a 10-point scale, all items were measured on a 7-point Likert scale (1=strongly disagree; 7=strongly agree).

3.4.1 Control variables. *Power usage* was measured with the 12-item power user scale developed by previous researchers that assesses users' perceived self-efficacy and expertise in using information technology [50]. Sample items included, "I make good use of most of the features available in any technological device", "I feel like information technology is a part of my daily life.," and "I love exploring all the features that any technological gadget has to offer." (α = .81, *M*=4.26, *SD*=.88).

Automation bias was measured by the Automation-induced Complacency Potential Scale [32], which is an update of the Complacency Potential Rating scale developed by Singh et al. [47] previously. Sample items included, "When I have a lot to do, it makes sense to delegate a task to automated systems", "Distractions and interruptions are less of a problem for me when I have an automated system to cover some of the work", and "It's not usually necessary to pay much attention to automated system when it is running" (α = .81, *M*=4.37, *SD*=.82).

3.4.2 Mediating variables. *Bandwagon heuristic* was assessed on the scale used in Sundar et al. [52], modified to fit the context of the movie recommender system. Sample items include, "Many users recommended the movies I just saw", "Many users have liked these suggested movies", and "The system recommended the movies based on how many users have rated them highly" (α = .92, *M*=4.65, *SD*=1.35).

Identity heuristic was measured with 6 items, especially created by the researchers for the current study. Sample items include, "The recommended movies reflect my own preferences" and "The recommended movies reflect my unique characteristics" (α = .97, *M*=3.26, *SD*=1.74).

3.4.3 Manipulation check. The manipulation check questions for content-based filtering were "The system recommended the above movies because the movie's characteristics fit my personal taste in movies" and "The system recommended movies by finding movies that include the characteristics I like in movie content". For collaborative filtering, the questions were "The system recommended the above movies because many users similar in my taste of movies watched them" and "The system recommended movies based on what users similar to me in their taste of movies preferred". Finally, for the demographic filtering, the questions were "The system recommended the above movies because many other users similar in my age, ethnicity, and gender watched them" and "The system recommended movies based on what users similar to me in age, ethnicity and gender preferred."

3.4.4 Dependent variables. *Responsibility attribution* was measured by the three-item scale from Botti & McGill [5]. Participants were asked to indicate perceived responsibility of three different parties: the recommender system, the user themselves, and similar users (for a fair comparison, we asked about their responsibility attribution to the similar users if they are assigned to interact with the content-based recommender system as well, though it is not used in our data analysis). Sample items include "How much responsibility do you assign to yourself/ the recommender system/other similar users for receiving these movie recommendations?" (self responsibility: α = .86, *M*=3.08, *SD*=1.62; system's responsibility: α = .82, *M*=5.15, *SD*=1.36; similar users' responsibility: α = .90, *M*=3.70, *SD*=1.65).

Watching intention was measured with three items asking whether they would like to watch the recommended movies in the near future. Sample items include "I would like to see at least one the recommended movies sometime in the future", "I would like to see a majority of the recommended movies in the future" (α = .93, *M*=3.82, *SD*=2.05)

Perceived quality of the recommender system was measured by modifying an existing scale used in Knijnenburg [22]. Sample items include "The recommender system is useful", "I can find better movies without the help of the recommender system (R)" (α = .94, M=3.44, SD=1.80)

Cognitive trust is measured by the 6-item scale developed by Madsen & Gregor [30]. Sample items include "The recommender system performs reliably", "The recommender system has sound knowledge about how to recommend movies", and "The recommender system correctly uses the information I enter" (α = .96, M=3.96, SD=1.72).

Perceived serendipity was measured by 3 items adopted from the scale developed by Knijnenburg et al. [22]. Sample items include "I discovered new movies through the recommender system" and "The recommender system suggested movies that are pleasantly surprising" (α = .78, M=3.97, SD=1.59).

3.5 Data analysis

To test H1 and H2, we used Model 4 of the Process macro in SPSS, a technique with which to test the effects of one independent variable, and one (or more) mediator(s) on the dependent variables [14]. We conducted a two-way MANCOVA (IV: system type & performance) to test H3, H4a, H4b, H5 and H6. To answer RQ1, we used Model 7 of the Process macro testing the effects of one independent variable, one moderator, and one (or more) mediator(s) on the dependent variables.

4 RESULTS

4.1 Manipulation check results

We found that for users that are assigned to the content-based filtering conditions, they were more likely to indicate that the system recommends products based on their personal taste (M=5.31, SD=2.49) compared to their counterparts in the demographic (M=2.74, SD=2.22) and collaborative filtering (M=4.93, SD=2.24) conditions, F (2, 223) = 26.75, p<.001. Similarly, participants assigned to the collaborative filtering conditions were more likely to indicate that the system recommends products based on what users similar in their taste preferred (M=6.32, SD=1.94) compared to participants in the content-based filtering (M=4.47, SD=2.28) and demographic filtering (M=3.90, SD=2.68) conditions, F (2, 223) = 22.27, p<.001. Finally, participants assigned to the demographic filtering conditions were more likely to indicate that the system recommends products based on what users similar to them in age, ethnicity and gender preferred (M=7.22, SD=2.67) compared to those in content-based (M=3.63, SD=1.97) and collaborative filtering (M=3.56, SD=2.31) conditions, F (2, 223) = 68.16, p<.001. Overall, these results suggest that our manipulation for the three conditions was successful.

4.2 The effects of recommender types and system's performance on user evaluation and trust

H1 proposed that collaborative filtering will lead to higher bandwagon perception which in turn will be associated with higher trust and better evaluative outcomes, compared to the other two systems. Results from mediation analyses support H1. Specifically, collaborative filtering caused higher bandwagon perception (M=5.03_a, SD=1.20) than content-based filtering (M=4.20_b, SD=1.41) and demographic filtering (M=4.61_b, SD=1.32), which in turn was associated with more positive evaluative outcomes (see Table 3 for a summary of the results from the mediation analyses). H2 proposed that content-based filtering will increase identity perception, which in turn would be associated with more positive evaluations of the system. We found partial support for this hypothesis as the identity perception was significantly higher among those who interacted with content-based filtering systems (M=3.57_a, SD=1.81) than those who interacted with demographic filtering system (M=2.71_b, SD=1.66), but not significantly higher than those who interacted with the collaborative filtering systems (M=3.47_a, SD=1.63) (see Table 3). In general, we found that users tended to evaluate the recommender system more positively and trust it more when interacting with the collaborative filtering approach, due to the elicitation of bandwagon heuristic, followed by the content-based filtering due to the elicitation of identity heuristic. Demographic filtering system received poorer evaluation and was associated with lower trust, compared to the other two approaches.

H3 proposed that users will evaluate the systems more negatively and trust the system less if the system provides low-quality recommendations. Results from the MANCOVA supported H3, Wilks' Λ = .57, F (6, 212) = 26.61, p <.001, partial η^2 = .43. Specifically, when receiving good quality recommendations, users had higher watching intention (good performance: M=4.97, SD=1.52; bad performance: M=2.58, SD=1.82), rated the recommender systems to be of higher quality (good performance: M=4.40, SD=1.43; bad performance: M=2.41, SD=1.58), had higher cognitive trust toward the systems (good performance: M=4.71, SD=1.39; bad performance: M=3.06, SD=1.65) and perceived the systems to have higher serendipity (good performance: M=4.32, SD=1.60; bad performance: M=3.58, SD=1.50).

4.3 The effects of system performance and recommender types on responsibility attribution

H4a and H4b proposed that there will be a self-serving bias when users interact with recommender systems. We found support for both H4a and H4b. Specifically, we found that users assigned more responsibility to themselves when they received good quality recommendations from the system than when they received bad quality recommendations, F (1, 224) = 11.10, p<.001, R^2=.05. One the other hand, they assigned more responsibility to the system when it delivered bad quality recommendations compared to good quality recommendations, F (1, 224) = 9.74, p<.01, R^2=.04, supporting the anticipated tendency for self-serving bias (See Figure 5).

H5 proposed that users will assign more responsibility to themselves and the system when interacting with content-based filtering recommender systems. Results from MANCOVA partially supported this hypothesis. Specifically, we found that when users interacted with content-based filtering recommender systems, they assigned more responsibility to themselves compared to collaborative filtering and demographic filtering recommender systems, F (2, 217) = 23.59, p<.001, R^2=.18. At the same time, those who interacted with content-based filtering systems assigned more responsibility to the system compared to those who interacted with collaborative

Table 3: Indirect effects of system type on system's evaluation and trust (summary of the mediation analyses)

Mediation path	B[a]	SE	CI[b]	
			Lower CI	Upper CI
RS type comparison 1[c] → Bandwagon perception → Watching intention	.37	.12	.16	.62
RS type comparison 1[c] → Bandwagon perception → RS quality	.33	.10	.15	.56
RS type comparison 1[c] → Bandwagon perception → Cognitive trust	.40	.12	.18	.64
RS type comparison 1[c] → Bandwagon perception → Serendipity	.33	.10	.14	.55
RS type comparison 2[d] → Identity perception → Watching intention	-.75	.20	-1.17	-.36
RS type comparison 2[d] → Identity perception → RS quality	-.77	.20	-1.17	-.39
RS type comparison 2[d] → Identity perception → Cognitive trust	-.73	.19	-1.12	-.35
RS type comparison 2[d] → Identity perception → Serendipity	-.62	.17	-.96	-.31

[a]Unstandardized path coefficient

[b]Bias-correlated and accelerated 95% confidence interval (CI).

[c]RS type comparison 1 is coded as 0= Content-based filtering, 1= Collaborative filtering

[d]RS type comparison 2 is coded as 0= Content-based filtering, 1= Demographic filtering

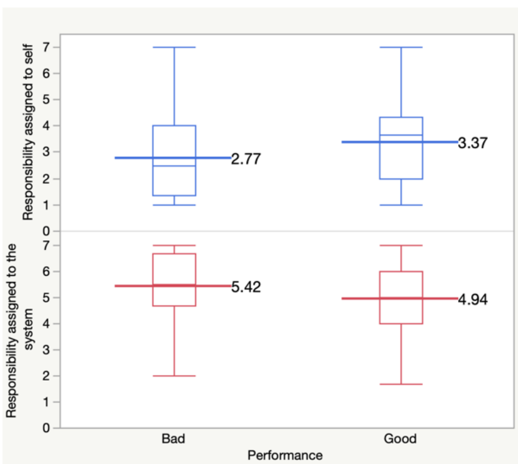

Figure 5: Participants' perceived responsibility of themselves or the recommender system when the system either delivers good or bad performances

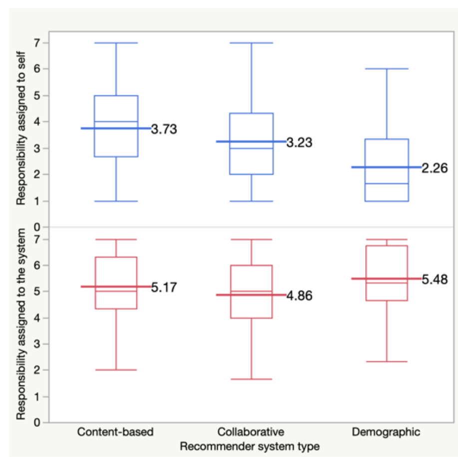

Figure 6: Participants' perceived responsibility of themselves or the recommender system while interacting with different recommender systems

filtering recommender systems, but not significantly more than those who interacted with demographic filtering recommender system, F (2, 217) = 3.83, p=.02, R^2=.03 (see Figure 6).

H6 proposed an interaction effect between system performance and system type on responsibility attribution. We found that the interaction effect on self-responsibility attribution to be non-significant, F (2, 217) = .22, p=.81, R^2=.00, but we found a marginally significant interaction effect of performance and system type on the responsibility users attributed to the system, F (2, 217) = 2.94, p=.06, R^2=.03. Specifically, users attributed more responsibility to the system when the content-based or collaborative systems performed better, but their attribution of responsibility to the demographic filtering system did not appear to be contingent on the quality of its recommendations (see Figure 7).

Here we would also like to report the analyses associated with how users assign responsibility to other similar users, although that was not hypothesized in this study. We ran an exploratory analysis of assignment of responsibility to three different parties (recommender system, the user themselves, and other similar users) in the collaborative filtering condition. Specifically, we ran a 2 (Between-subject factor: Performance: Good vs Bad) x 3 (Within-subject factor: self responsibility vs. recommender systems' responsibility vs. other similar users' responsibility) Mixed Model Repeated Measures ANCOVA, and found no significant difference between users responsibility assigned to the three parties (self responsibility: M=3.23, SD=1.51; recommender system's responsibility: M=4.86, SD=1.45; other similar users' responsibility: M=3.95, SD=1.56), Wilks' Λ = .99, F (2, 70) = .53, p=.59, partial η^2 = .02. However, an interaction effect between responsible parties and performance was observed, Wilks' Λ = .86, F (2, 70) = 5.50, p<.01, partial η^2 = .14. But we only observed the moderation effect of system performance on users'

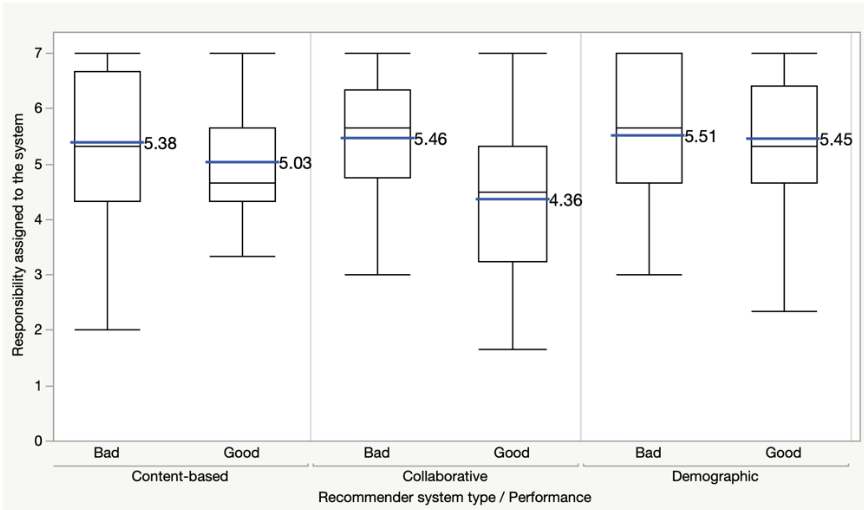

Figure 7: Participants' perceived responsibility of the recommender system while interacting with different recommender systems that deliver good or bad performances

assigned responsibility to themselves and perceived responsibility of the system as we have already reported in the main text of the Results section. There was no difference between receiving good performance (*M*=3.87, *SD*=1.51) or bad performance (*M*=4.03, *SD*=1.62) on how much responsibility participants assigned to other users. We then computed a new variable by subtracting responsibility one assigned to similar other users from self responsibility (i.e. self responsibility − similar other's responsibility), and ran an ANCOVA (controlling for power usage and automation bias), which showed a significant main effect of performance, F (1, 71) = 4.36, p=.04, partial η^2 = .06. Specifically, the difference between self-responsibility and similar other's responsibility was significantly larger when participants received bad recommendations (*M*=-1.27, *SD*=2.05.) compared to when they received good recommendations (*M*=-.17, *SD*=1.79), even though the fundamental premise of collaborative filtering is reliance on others' choices. This result suggests that users show similar self-serving bias when it comes to assigning responsibility to similar others in the case of collaborative filtering. This pattern is consistent with what we have reported in the main text (i.e., self-serving bias when comparing self responsibility with system responsibility).

4.4 The mediating effects of responsibility attribution between system type, system performance and system evaluation & trust

In response to RQ1, results of the mediation analysis revealed that responsibility attribution assigned to the recommender system is a significant mediator in the relationship between the independent variables (recommender system type, system performance) and the dependent variables (system evaluation and trust). Specifically, we found that the indirect effects of recommender system types on system evaluations and trust were moderated by the systems' performance (see Table 4 for a summary of results from the moderated mediation analysis). When the system delivered bad performance, users who interacted with the three different systems assigned relatively similar amount of responsibility to the recommender system. Therefore, subsequently, no difference between users' evaluations and trust was observed among the three systems. When the system delivered good performance, however, users assigned significantly lower responsibility to the recommender system when interacting with the collaborative filtering recommender system compared to demographic filtering recommender system. This in turn was associated with higher evaluative outcomes (with the exception of perceived serendipity) and trust among users who interacted with the collaborative filtering systems.

4.5 Summary of results

In general, as one would expect, we found that users generally trust a system that delivers good performance more than one that delivers bad performance. More interestingly, we found that users tend to perceive different systems differently. Specifically, users trust collaborative filtering more than content-based filtering and demographic filtering due to the elicitation of bandwagon heuristic. Content-based filtering is trusted more than demographic filtering due to the elicitation of identity heuristic.

In addition, we discovered that users tend to hold a self-serving bias when interacting with recommender systems, namely overattributing responsibility to themselves when receiving good recommendations, and overattributing responsibility to the system when receiving bad recommendations. However, this bias is less salient when they interact with demographic recommender systems because they assign more responsibility to the system regardless of performance quality. Furthermore, we found that the responsibility attributed to the system is a significant mediator between system type, performance quality and users' trust in the system. Specifically, we found that users will trust the collaborative filtering recommender system more than content and demographic systems

Table 4: Indirect effects of recommender system type and performance on system's evaluation and trust (summary of the moderated mediation analyses)

Mediation path	B[a]	SE	CI[b]	
			Lower CI	Upper CI
RS type comparison 3[c] * performance → RS responsibility → Watching intention	.34	.18	.05	.73
RS type comparison 3[c] * performance → RS responsibility → RS quality	.29	.15	.05	.61
RS type comparison 3[c] * performance → RS responsibility → Cognitive trust	.25	.14	.02	.59
RS type comparison 3[c] * performance → RS responsibility → Serendipity	.15	.11	-.02	.40

[a]Unstandardized path coefficient

[b]Bias-correlated and accelerated 95% confidence interval (CI).

[c]Recommender system type comparison 3 is coded as 0= Demographic filtering, 1= Collaborative filtering

when it performs well, due to lower attributed responsibility to the system.

5 DISCUSSION

The current study revealed that users tend to trust collaborative filtering recommendation system over the other two recommender systems, because it triggers the bandwagon heuristic. Indeed, compared to content-based filtering, collaborative filtering has an added layer of social information that can serve as an additional persuasive cue for users to trust the system [27]. Our study corroborates previous findings that users rely quite heavily on social endorsement cues in the e-commerce context [37]. But, more importantly, we find that not all social information is trusted. Superficial social information, such as demographics, is not sufficient for building users' trust towards the recommender system; instead, users need social information that reflects others' collective tastes and preferences towards specific products. Designers can address this need for social endorsement by computing bandwagon metrics at a variety of levels (number of views, number of clicks, number of shares, and so on) and displaying them prominently on the interfaces of collaborative filtering systems.

We also found that users' higher trust towards collaborative filtering is independent of the system's performance. Many researchers have acknowledged a practical problem with the design of recommender system, the cold start problem, which happens when a user is at the early stage of using a recommender system, and the system might not make accurate predictions because it knows relatively little about the user [44]. Several researchers have proposed that we could potentially leverage the power of some more basic forms of recommender systems to optimize the accuracy of prediction, such as demographic filtering [12, 28, 44], as it requires less individual-level information to make recommendations. While this solution addresses the cold start problem from a computational standpoint, our study suggests that it might not be useful in facilitating users' subjective trust in the system. In the current study, all our participants interacted with the recommender system for the first time. Yet, we observed that users tend to hold less trust towards demographic filtering system, even when it makes good recommendations. Although collaborative filtering system might face a bigger cold start problem empirically, our study found that it can promote user trust in the system, even when the performance is not good. The mere fact that collaborative filtering reflects the recommendations of similar others seems to go a long way in promoting trust and even helps assuage users when the recommendations are sometimes less than optimal. In addition to research on computational advances, results from our study suggest more user-facing and human-centered empirical studies about the problems of recommender systems in order to promote trust among users and ensure better user experience. For the practitioner, this implies greater inclusion of collaborative elements in the design of recommender systems and prominent display of these elements on the interface, including opportunities for users to explicitly recommend product to similar others. The desirable design goal would be to trigger the bandwagon heuristic by making prominent the aforementioned signals of social approval on the interface of the recommender system, such as number of Likes, Thumbs Up, and so on.

If the system is a content-based one and does not have a collaborative element, then our findings suggest that the key to gaining user trust is by invoking the identity heuristic. This means design should emphasize identity-enhancing elements on the interface, conveying to users that the outcomes of the system are directly related to their preferences and highly reflective of their individual identity. Prior research on customization has shown the powerful role played by products reflecting users at the individual level. Perceived reflection of self is known to explain a major portion of the variance in product loyalty [25]. Content-based recommender systems can leverage the positive effects of customization and individuation by including design elements that allow users to consciously engage in identity-related preference setting (e.g., "are you the kind of person who likes sad movies," "are you a sports fan or sports fanatic?") and accompany system recommendations with explicit appeals to their identity (e.g., "as a sports fanatic, you will love this").

Aside from cognitive heuristics, we also found users' responsibility attribution to be a significant aspect of the theoretical mechanism underlying user trust in recommender systems. Our study extends the current literature regarding responsibility attribution in HCI, from embodied agents to recommender systems. Consistent with CASA [41], we found that users have the same self-serving bias in their interaction with recommender systems as they do in their interaction with other people, as would be predicted by attribution theory [17]. Specifically, we found that we will be more likely to credit ourselves when the recommender system performs well, but blame the system more when it delivers performs poorly. The results from our study have practical implications for system designers who would like to facilitate user collaboration and compliance. Users are more likely to blame the system rather than themselves when the outcomes of their collaboration with the system are less than satisfactory. Translating this to design, recommender systems can be more up front with users about the factors that might influence performance quality, making them aware that withholding certain pieces of information from the system (for privacy and other reasons) could adversely affect the quality of the outcome. Such measures may remind users that they are in part responsible for sub-optimal performance by the system, thereby reducing their tendency to place the entire blame on the system. Psychologists have found that self-enhancement (one's own sense of self-worth) and self-presentation (how one comes across to others) are key motivating factors for self-serving bias. Interface cues that boost the self-esteem of users by explicitly showcasing what they gained or achieved (e.g., protected their privacy) even in the event of a poor collaborative outcome will minimize their need to engage in self enhancement. Reminding users that the system will not identify them to other users, despite its collaborative framework, should help assuage self-presentation concerns. Such interface-based solutions can reduce self-serving bias by promoting a more accurate attribution of responsibility, with obvious implications for building user trust in the system over the long run.

We also found that when users interact with less intelligent systems like the demographic filtering recommender system, they tend to have less self-serving bias. Specifically, we found that users are more likely to credit the system when demographic filtering system delivers good recommendations. This result can be explained by the expectation users hold for different recommender systems, always an important factor in user interactions with algorithms [20, 48]. It is possible that when we are interacting with less intelligent systems, receiving good recommendations means a positive violation of our expectations, which makes us more likely to credit the system. However, it must be mentioned here that more credit to the system means less trust in the system. In fact, only self-responsibility is positively correlated with system evaluation and trust. This could be due to human tendency to trust things that we are responsible for, as opposed to the things that machines are responsible for, in order to maintain our self-esteem and self-control, which is crucial in human computer interaction [24]. An obvious design implication of this finding is that user contribution to the outcome of the system should be explicit at every stage of the interaction. The interface should emphasize collaboration, with the user in the driver's seat, making the key decisions that are related to the outcome. Cues on the interface that visualize user's contribution can serve to trigger the 'ownness heuristic' [49] and thereby increase the perceived credibility of the system.

In summary, our study not only empirically tested how users receive either good or bad recommendations from different approaches, but also uncovered the underlying psychological mechanisms of such effects. These results can help industry practitioners design better messages when explaining the recommendation results to users, with a view to enhancing their trust in the recommender system. Specifically, our studies revealed the importance of computing bandwagon metrics, displaying identity enhancing elements and users' contribution on the interface to promote trust in the system.

6 LIMITATION

Our study is not free from limitations. First of all, in terms of external validity, as stated earlier, our study only investigated one context, which is e-commerce, and the product recommended in our study (i.e. movies) has somewhat lower stakes than many recommender systems. Future studies could investigate how users perceive different recommender systems in other contexts (e.g., health, news, etc.). Furthermore, all our participants had only a one-time interaction with the system, which means that the effects we found in the current study might not hold true when users use the system for a prolonged period of time. Future studies could extend the current experiment by letting users continue to use the system for a longer period, and investigate whether changes in performance of the system over time alter users' trust and evaluations of the system. Moreover, we only included clearly good and clearly bad system performances, while in reality users might receive recommendations that could fall anywhere on a continuum from good to bad. Future studies could potentially incorporate a range of different recommendations that have different levels of quality in an ordinal fashion and assess whether the corresponding variation in users' perception of different recommender systems follow a linear or some other pattern. Secondly, in terms of the internal validity, we acknowledge the fact that we only manipulated the explanation representing each type of recommender system instead of incorporating different algorithms, given that our goal is to understand the psychological effects of different types of recommender systems. Another potential limitation is that all participants in the good condition received the exact same set of recommendations (as did all participants in the bad condition) in order to avoid content confounds, and this lack of personalization might have influenced perceptions of recommendation quality. Future study could incorporate more personalized recommendations for each user, while at the same time, making sure each recommendation is of the similar quality to avoid any confounds that might be induce due to recommending different products to different participants. Finally, we would like to acknowledge one limitation in our measurement validity, particularly the measures for watching intention, which might have low predictive validity in approximating actual watching behaviors. Future studies could unobtrusively track users' actions on the interface (e.g., save/add the item to their favorite list, click to watch, etc.).

7 CONCLUSION

Understanding users' trust in recommender systems is a key issue for researchers as well as developers. Instead of focusing on the advancement of computational complexity and objective performance, the current study aims to understand users' subjective tendency to trust certain types of systems over others. By focusing on three of the most commonly used approaches in recommendation systems (i.e., collaborative filtering, content-based filtering and demographic filtering), our study not only revealed that users tend to have higher overall trust in collaborative filtering regardless of the system's performance, but also uncovered the underlying psychological reason behind this tendency, namely the triggering of the bandwagon heuristic. The current study also helps extend the current literature regarding how users themselves make sense of the recommendation process and the degree to which they assign responsibility to themselves vs. the system, and how this attribution in turn affects their trust in the recommendations. By focusing on such psychological factors and users' subjective evaluations of recommender systems, we offer practical insights for building more human-centered trustworthy systems.

REFERENCES

[1] Robert M. Arkin, Alan J. Appelman, and Jerry M. Burger. 1980. Social anxiety, self-presentation, and the self-serving bias in causal attribution. *Journal of Personality and Social Psychology* 38, 1 (1980), 23–35. https://doi.org/10.1037/0022-3514.38.1.23

[2] Alejandro Barredo Arrieta, Natalia Díaz-Rodríguez, Javier Del Ser, Adrien Bennetot, Siham Tabik, Alberto Barbado, Salvador Garcia, Sergio Gil-Lopez, Daniel Molina, Richard Benjamins, Raja Chatila, and Francisco Herrera. 2020. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion* 58, (June 2020), 82–115. https://doi.org/10.1016/j.inffus.2019.12.012

[3] Edmond Awad, Sydney Levine, Max Kleiman-Weiner, Sohan Dsouza, Joshua B. Tenenbaum, Azim Shariff, Jean-François Bonnefon, and Iyad Rahwan. 2020. Drivers are blamed more than their automated cars when both make mistakes. *Nat Hum Behav* 4, 2 (February 2020), 134–143. https://doi.org/10.1038/s41562-019-0762-8

[4] Krisztian Balog, Filip Radlinski, and Shushan Arakelyan. 2019. Transparent, Scrutable and Explainable User Models for Personalized Recommendation. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '19)*. https://dl.acm.org/citation.cfm?id=3331211

[5] Simona Botti and Ann L. McGill. 2006. When Choosing Is Not Deciding: The Effect of Perceived Responsibility on Satisfaction. *Journal of Consumer Research* 33, 2 (September 2006), 211–219. https://doi.org/10.1086/506302

[6] Chaochao Chen, Xiaolin Zheng, Mengying Zhu, and Litao Xiao. 2016. Recommender System with Composite Social Trust Networks. *IJWSR* 13, 2 (April 2016), 56–73. https://doi.org/10.4018/IJWSR.2016040104

[7] Mark Coeckelbergh. 2020. Artificial Intelligence, Responsibility Attribution, and a Relational Justification of Explainability. *Sci Eng Ethics* 26, 4 (August 2020), 2051–2068. https://doi.org/10.1007/s11948-019-00146-8

[8] A.R. Elangovan, Werner Auer-Rizzi, and Erna Szabo. 2007. Why don't I trust you now? An attributional approach to erosion of trust. *Journal of Managerial Psychology* 22, 1 (January 2007), 4–24. https://doi.org/10.1108/02683940710721910

[9] Ekstrand, M. D., Harper, F. M., Willemsen, M. C., & Konstan, J. A. (2014). User Perception of Differences in Recommender Algorithms. RecSys '14: Proceedings of the 8th ACM Conference on Recommender Systems, 161–168. New York, NY, USA: ACM.

[10] Antoine Falconnet, Wietske Van Osch, Joerg Beringer, Pierre-Majorique Léger, and Constantinos Coursaris. 2021. Improving User Experience through Recommendation Message Design: A Systematic Literature Review of Extant Literature on Recommender Systems and Message Design.

[11] Franz Faul, Edgar Erdfelder, Albert-Georg Lang, and Axel Buchner. 2013. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods* 39, 2, 175–191. https://doi.org/10.3758/BF03193146

[12] Jyotirmoy Gope and Sanjay Kumar Jain. 2017. A survey on solving cold start problem in recommender systems. In *2017 International Conference on Computing, Communication and Automation (ICCCA)*, 133–138. https://doi.org/10.1109/CCAA.2017.8229786

[13] Taehyun Ha, Young June Sah, Yuri Park, and Sangwon Lee. 2020. Examining the effects of power status of an explainable artificial intelligence system on users' perceptions. *Behaviour & Information Technology* 0, 0 (November 2020), 1–13. https://doi.org/10.1080/0144929X.2020.1846789

[14] Andrew F. Hayes. 2017. Introduction to Mediation, Moderation, and Conditional Process Analysis, Second Edition: A Regression-Based Approach. Guilford Publications.

[15] Joo-Wha Hong. 2020. Why Is Artificial Intelligence Blamed More? Analysis of Faulting Artificial Intelligence for Self-Driving Car Accidents in Experimental Settings. *International Journal of Human-Computer Interaction* 36, (June 2020), 1–7. https://doi.org/10.1080/10447318.2020.1785693

[16] Yimin Hou, Ting Xiao, Shu Zhang, Xi Jiang, Xiang Li, Xintao Hu, Junwei Han, Lei Guo, L. Stephen Miller, Richard Neupert, and Tianming Liu. 2016. Predicting Movie Trailer Viewer's "Like/Dislike" via Learned Shot Editing Patterns. *IEEE Transactions on Affective Computing* 7, 1 (January 2016), 29–44. https://doi.org/10.1109/TAFFC.2015.2444371

[17] Harold H. Kelley. 1967. Attribution theory in social psychology. *Nebraska Symposium on Motivation* 15, (1967), 192–238.

[18] Harold H. Kelley. 1973. The processes of causal attribution. *American Psychologist* 28, 2 (1973), 107–128. https://doi.org/10.1037/h0034225

[19] Taemie Kim and P. Hinds. Who Should I Blame? Effects of Autonomy and Transparency on Attributions in Human-Robot Interaction. In *ROMAN 2006 - The 15th IEEE International Symposium on Robot and Human Interactive Communication*. https://doi.org/10.1109/ROMAN.2006.314398

[20] René F. Kizilcec. 2016. How Much Information? Effects of Transparency on Trust in an Algorithmic Interface. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 2390–2395. Retrieved September 8, 2021 from https://doi.org/10.1145/2858036.2858402

[21] G. Klien, D.D. Woods, J.M. Bradshaw, R.R. Hoffman, and P.J. Feltovich. 2004. Ten challenges for making automation a "team player" in joint human-agent activity. *IEEE Intelligent Systems* 19, 6 (November 2004), 91–95. https://doi.org/10.1109/MIS.2004.74

[22] B. P. Knijnenburg, M. C. Willemsen, Z. Gantner, H. Soncu, and C. Newell. 2012. Explaining the user experience of recommender systems. *User Modeling and User-Adapted Interaction* 22, 4–5 (2012), 411–504. https://doi.org/10.1007/s11257-011-9118-4

[23] Joseph A. Konstan and John Riedl. 2012. Recommender systems: from algorithms to user experience. *User Model User-Adap Inter* 22, 1 (April 2012), 101–123. https://doi.org/10.1007/s11257-011-9112-x

[24] John Lee and Neville Moray. 1992. Trust, control strategies and allocation of function in human-machine systems. *Ergonomics* 35, 10 (October 1992), 1243–1270. https://doi.org/10.1080/00140139208967392

[25] Seoyeon Lee and S. Shyam Sundar. 2015. Cosmetic Customization of Mobile Phones: Cultural Antecedents, Psychological Correlates. *Media Psychology* 18, 1 (January 2015), 1–23. DOI:https://doi.org/10.1080/15213269.2013.853618

[26] Xuying Leo and Young Eun Huh. 2020. Who gets the blame for service failures? Attribution of responsibility toward robot versus human service providers and service firms. *Computers in Human Behavior* 113, (December 2020), 106520. https://doi.org/10.1016/j.chb.2020.106520

[27] Mengqi Liao and S. Shyam Sundar. 2021. When E-Commerce Personalization Systems Show and Tell: Investigating the Relative Persuasive Appeal of Content-Based versus Collaborative Filtering. *Journal of Advertising* 0, 0 (March 2021), 1–16. https://doi.org/10.1080/00913367.2021.1887013

[28] Blerina Lika, Kostas Kolomvatsos, and Stathes Hadjiefthymiades. 2014. Facing the cold start problem in recommender systems. *Expert Systems with Applications* 41, 4, Part 2 (March 2014), 2065–2073. https://doi.org/10.1016/j.eswa.2013.09.005

[29] Peng Liu and Yong Du. 2021. Blame Attribution Asymmetry in Human-Automation Cooperation. *Risk Anal* (January 2021). https://doi.org/10.1111/risa.13674

[30] Maria Madsen and Shirley Gregor. 2000. Measuring human-computer trust. In *Proceedings of the 11 th Australasian Conference on Information Systems*, 6–8.

[31] Bertram F. Malle, Matthias Scheutz, Thomas Arnold, John Voiklis, and Corey Cusimano. 2015. Sacrifice One For the Good of Many? People Apply Different Moral Norms to Human and Robot Agents. In *2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 117–124.

[32] Stephanie M. Merritt, Alicia Ako-Brew, William J. Bryant, Amy Staley, Michael McKenna, Austin Leone, and Lei Shirase. 2019. Automation-Induced Complacency Potential: Development and Validation of a New Scale. *Frontiers in Psychology* 10, (2019), 225. https://doi.org/10.3389/fpsyg.2019.00225

[33] Bertalan Mesko. 2017. The role of artificial intelligence in precision medicine. *Expert Review of Precision Medicine and Drug Development* 2, 5 (September 2017), 239–241.https://doi.org/10.1080/23808993.2017.1380516

[34] Clifford Nass, B. J. Fogg, and Youngme Moon. 1996. Can computers be teammates? *International Journal of Human-Computer Studies* 45, 6 (December 1996), 669–678. DOI:https://doi.org/10.1006/ijhc.1996.0073

[35] Clifford Nass and Youngme Moon. 2000. Machines and Mindlessness: Social Responses to Computers. *Journal of Social Issues* 56, 1 (2000), 81–103. https://doi.org/10.1111/0022-4537.00153

[36] Clifford Nass, Youngme Moon, and Paul Carney. 1999. Are people polite to computers? Responses to computer-based interviewing systems. *Journal of Applied Social Psychology* 29, 5 (1999), 1093–1110. https://doi.org/10.1111/j.1559-1816.1999.tb00142.x

[37] Ingrid Nunes and Dietmar Jannach. 2017. A systematic review and taxonomy of explanations in decision support and recommender systems. *User Model User-Adap Inter* 27, 3 (December 2017), 393–444. https://doi.org/10.1007/s11257-017-9195-0

[38] Michael J. Pazzani. 1999. A Framework for Collaborative, Content-Based and Demographic Filtering. *Artificial Intelligence Review* 13, 5 (December 1999), 393–408. https://doi.org/10.1023/A:1006544522159

[39] Pearl Pu, Li Chen, and Rong Hu. 2012. Evaluating recommender systems from the user's perspective: survey of the state of the art. *User Model User-Adap Inter*

22, 4 (October 2012), 317–355. https://doi.org/10.1007/s11257-011-9115-7
[40] Lara Quijano-Sanchez, Christian Sauer, Juan A. Recio-Garcia, and Belen Diaz-Agudo. 2017. Make it personal: A social explanation system applied to group recommendations. *Expert Systems with Applications* 76, (June 2017), 36–48. https://doi.org/10.1016/j.eswa.2017.01.045
[41] Byron Reeves and Clifford Ivar Nass. 1996. The media equation: How people treat computers, television, and new media like real people and places. Cambridge University Press, New York, NY, US.
[42] Paul Resnick, Neophytos Iacovou, Mitesh Suchak, Peter Bergstrom, and John Riedl. 1994. GroupLens: an open architecture for collaborative filtering of netnews. In *Proceedings of the 1994 ACM conference on Computer supported cooperative work* (CSCW '94), Association for Computing Machinery, New York, NY, USA, 175–186. https://doi.org/10.1145/192844.192905
[43] Edward G. Rogoff, Myung-Soo Lee, and Dong-Churl Suh. 2004. "Who Done It?" Attributions by Entrepreneurs and Experts of the Factors that Cause and Impede Small Business Success. *Journal of Small Business Management* 42, 4 (2004), 364–376. https://doi.org/10.1111/j.1540-627X.2004.00117.x
[44] Laila Safoury and Akram Salah. 2013. Exploiting User Demographic Attributes for Solving Cold-Start Problem in Recommender System. *LNSE* (2013), 303–307. https://doi.org/10.7763/LNSE.2013.V1.66
[45] Jens Seiffert-Brockmann, Christopher Ruppel, and Sabine Einwiller. 2019. Gives you wings or not? Exploring the impact of viewers' responsibility attribution and surprise on their attitude, identification and trust. *Corporate Communications: An International Journal* 25, 1 (January 2019), 113–127. https://doi.org/10.1108/CCIJ-07-2019-0087
[46] Alexander Serenko. 2006. The use of interface agents for email notification in critical incidents. *International Journal of Human-Computer Studies* 64, 11 (November 2006), 1084–1098. https://doi.org/10.1016/j.ijhcs.2006.06.004
[47] Indramani L. Singh, Robert Molloy, and Raja Parasuraman. 1993. Automation-Induced "Complacency": Development of the Complacency-Potential Rating Scale. *The International Journal of Aviation Psychology* 3, 2 (April 1993), 111–122. https://doi.org/10.1207/s15327108ijap0302_2
[48] Aaron Springer, Victoria Hollis, and Steve Whittaker. 2017. Dice in the Black Box: User Experiences with an Inscrutable Algorithm. In *2017 AAAI Spring Symposium Series*. Retrieved September 8, 2021 from https://www.aaai.org/ocs/index.php/SSS/SSS17/paper/view/15372
[49] S. Shyam Sundar. 2008. The MAIN model: A heuristic approach to understanding technology effects on credibility. Digital media, youth, and credibility. The MIT Press, Cambridge, MA, 72–100.
[50] S. Shyam Sundar and Sampada S. Marathe. 2010. Personalization versus Customization: the Importance of Agency, Privacy, and Power Usage. *Human Communication Research* 36, 3 (July 2010), 298–322. https://doi.org/10.1111/j.1468-2958.2010.01377.x
[51] S. Shyam Sundar and Clifford Ivar Nass. 2000. Source Orientation in Human-Computer Interaction: Programmer, Networker, or Independent Social Actor. *Communication Research* 27, 6 (December 2000), 683–703. https://doi.org/10.1177/009365000027006001
[52] S. Shyam Sundar, Anne Oeldorf-Hirsch, and Qian Xu. 2008. The bandwagon effect of collaborative filtering technology. https://doi.org/10.1145/1358628.1358873
[53] Melvin L. Snyder, Walter G. Stephan, and David Rosenfield. 1976. Egotism and attribution. *Journal of Personality and Social Psychology* 33, 4 (1976), 435–441. https://doi.org/10.1037/0022-3514.33.4.435
[54] Taste - Movie recommendations and reviews from likeminded people. Retrieved from https://www.taste.io
[55] Poonam B.Thorat, R. M. Goudar, and Sunita Barve. 2015. Survey on Collaborative Filtering, Content-based Filtering and Hybrid Recommendation System. *IJCA* 110, 4 (January 2015), 31–36. https://doi.org/10.5120/19308-0760
[56] Neal Tognazzini and D. Justin Coates. 2021. Blame. In *The Stanford Encyclopedia of Philosophy* (Summer 2021), Edward N. Zalta (ed.). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/sum2021/entries/blame/
[57] Susan M. Weinschenk. 2009. *Neuro Web Design: What Makes Them Click?* (1st ed.). New Riders Publishing, USA.