

What Makes Audience to Watch Bollywood Films in India: An Empirical Study

Prince Thomas Chemmanoor*, Falguni H. Pandya**

ABSTRACT

Purpose: *The present paper intends to study how to increase box-office collection by producing audience centric film.*

Objectives: *To investigate and examine what audience likes to see in films.*

Design/Methodology/Approach: *The paper has used descriptive research design because it intends to investigate questions and results to discovery of association among variables. Data are collected from 1165 respondents of different age groups, education, income background, and locality through structured questionnaire-cum-interview method. A questionnaire is constructed using Likert scale and administered using convenience sampling approach. Data are analysed using the SPSS 17 software.*

Research Limitations/Implications: *Future research can be conducted for further improvements in measuring techniques and also by adding more volleys of questions. Even in the case of set of influences, they might well differ in other countries or at other times. In sum, all the caveat and limitations just enumerated a need for support in future investigations for using data in multiple time periods, in multiple countries, and with multiple viewers. Future researches can address the generalisability of our findings across time periods and geographical settings.*

Finding: *This paper adds historical and theoretical depth for debate to attitude of audience to watch Bollywood films in respect to characters, likings, and theme of films. The attitude of audience of watching Bollywood films is not an actively researched topic in India. The study therefore contributes toward better understanding to many stakeholders of films while it is being produced.*

Practical Implications: *The findings can provide some relevant scope for directors and marketers of Bollywood movies.*

Keywords: *Film, Attitude, Box-office, Audience*

INTRODUCTION

India, with a population of 1300 million and still counting, has a diversified cinematic culture of Mumbai, Andhra Pradesh, Gujarat, Karnataka, Kerala, Punjab, Tamil Nadu, Uttar Pradesh, West Bengal, and Maharashtra in various languages especially Hindi, Telugu, Gujarati, Kannada, Malayalam, Punjabi, Tamil, Bhojpuri, Bengali, and Marathi. Film industry is a key and perhaps the most vibrant industry of Indian economy. The number of films produced annually in India is higher than in any country of the world including the USA's Hollywood. While Hollywood produces around 550 movies a year, the Indian film (movie) industry produces more than 1000 movies every year (Krishnan & Sakkthivel, 2010). Around four million Indians 'go to movies' on any given day of the year, and this number swells during festivals and holidays (Krishnan & Sakkthivel, 2010). The Hindi film making

industry in India that is based out of Mumbai, referred to as commonly 'Bollywood' by the media and people, is largest film producing centre in India (Sarkar & Nayak, 2009).

The film industry is the part and parcel of Entertainment and Media Industry (E&M) and plays a significant role to economy of any country. It consists of film production houses, film studios, distributors, exhibitors, film directors, music directors, choreographers, script writers, and many more that includes actors, actresses, and other important characters. The major clusters of film making centres in the world are located in United States, Nigeria, Egypt, Hong Kong, Canada, Australia, Indonesia, United Kingdom, and India (European Audiovisual Council, Council of Europe, 2009). Due to low cost production, mainly of place & personnel, and increasing curiosity of viewing the scenes of other countries, many a times shooting of films take place outside of the country where

* Faculty, MBA Department, Center for Management Studies Dharmsinh Desai University Nadiad, Gujarat, India.
Email: frince.mba@ddu.ac.in

** Faculty, MBA Department, Center for Management Studies Dharmsinh Desai University Nadiad, Gujarat, India.
Email: fhpandya@gmail.com

films are produced or originated (Film-reference, 2008). As a result, many times Hollywood and Bollywood movies are produced neither in US nor in India but in some other countries of the World.

The Indian Entertainment and Media (E&M) industry, with revenues of about 805 billion INR (17.2 billion USD) in 2011, is poised to grow strongly over the next few years on the back of strong to moderate economic growth, increasing power of consumer spending, and having positive demographic indicators (India Entertainment and Media Outlook, 2012). The industry revenues are expected to reach 1,764 billion INR (37.6 billion USD) by 2016, with a CAGR of about 17% from 2012 to 2016 (India Entertainment and Media Outlook, 2012). Currently, India ranks 14th largest E&M market in the world, contributing about 1% of GDP.

Hirsch (1972) observed that “films or movies are cultural goods defined as ‘non-material good’ directed at public consumers for whom they generally serve an aesthete or expressive, rather than clearly utilitarian function”. Different from any other product yet it is similar in some aspects like highly customised. William Goldman’s (1983) classic statement is that in the movie business “nobody knows anything”. It is a highly uncertain business because there is no perfect script or winning formula. Everything depends upon the imagination of the creative script, marketing budget and acceptance among the audience. Like any other product, important parameters include financial ratios like ROI, price, promotion, target audience, sales, distribution, and IPRs. Every movie is like starting a new business because same production or same characters may make the film hit as well as flop at different time. A successful movie is not indicator of future successes because every film is unique in own way. Film is a product and also a medium of communication.

REVIEW OF LITERATURE

Drake (2008) observed that in Hollywood, films operate under what economists call “asymmetric information” i.e. incomplete information between producers and consumers and this leads to difficulties in reliably predicting success or failure at box office collection, hence surprises box office with “hits” and “flops”. Market research assesses the opportunities and potential success of the marketing of films that include previewing, market analysis, and tracking survey. On other hand, the movie industry has raised questions concerning the problematic relationship between the success-related aspects of artistic excellence and commercial appeal. The article proposes that when the former is measured by industry recognition (Oscars

and other awards) and the latter by market performance (box office and video rentals) and when the former hinges on the evaluative judgments of reviewers and consumers (ratings of excellence) and the latter on the level of buzz among these audience members (amount of attention, word of mouth, or click of mouse), the two phenomena are essentially separable as independent paths to conceptually distinct and empirically uncorrelated aspects of motion-picture success and analysis shows that reviewer-and-consumer evaluations and buzz respond differently to a film’s marketing clout (production budget, opening screens, and opening box office) and that these audience responses contribute independently to a film’s industry recognition and market performance along two separable paths. These findings suggest various implications for movie marketers, film producers, actors or actresses, and other members of the motion-picture industry (Holbrook & Addis, 2007).

According to a study by Tirumala (2010), Indian cinema has been one of the most dominant and distinguishing features of the subcontinent’s culture for the past 65 years. As Indians continue to seek out jobs and educational opportunities worldwide, they are bringing these cultural artifacts with them. Previous research suggests that Indian diaspora may use these types of media products as a form of cultural maintenance. This article explores the role of Bollywood movies in constructing and maintaining the cultural identity among second generation Indian Americans. In particular, this study seeks to demonstrate how Bollywood film viewing practices of the second generation Indian Americans intersect to create a notion of “Indianness.” Indian Americans depended on various interpersonal and mediated communication activities. Interestingly, the study found that though social factors such as family and peer groups have played a dominant role in constructing the Indian identity.

Sridhar and Mattoo (1997) claimed that India is the largest producer of feature films in the world. It is estimated that between 800 and 1000 films are produced in India annually compared to Hollywood which produces half that number. A study conducted by Srinivas (2000) suggested that Indian cinema has been around since 1913, the year to which most retrospectives trace the first Indian films – the mythological of Phalke. In spite of its early beginnings, phenomenal productivity and continued success, films in India did not receive any industry status from the government till 1998-99. It is therefore more accurate to refer to an Indian film business to describe this heterogeneous body of entrepreneurs and artistes. In Western societies, mainstream film audiences are quiet and highly disciplined in their viewing styles. In

cinema theatres people watch the film for the most part silently, the expectation of silent attentiveness to onscreen happenings provide a guiding etiquette. Viewers rarely talk out loudly and never engage in the overtly interactive and spontaneously expressive style of reception seen in theatres in India and with Indian audiences. Indian audiences are consequently closer to the producers and less alienated from the product compared to their Western counterpart.

Ainslie, Dreze, & Zufryden (2005) observed that the actors have a direct and directors have an indirect effect on consumer's movie choice and releasing a movie contemporaneously with other movies of the same genre adversely affects box-office performance all around, there is a displacement effect, which leads to a less severe sales loss in the long run. In another study it was found that most movie theaters in the multiplex era have adopted a remarkably similar strategy, one which is also very vulnerable to recent trends such as the explosion of home cinema, pay TV, video-on-demand (VOD), discounting by mass merchandisers of DVDs, computer games, and the collapse of video windows. Just as technological convergence has created a challenge for movie theaters, as it has in the past, new technologies and creative use of assets (combined with multiple target marketing) can offer a counter measure for at least some movie theaters; at least, until the next challenge arises. What is unlikely to succeed is the status quo, especially when so many multiplexes offer the same format as their competitors, appear to adopt a narrow definition of what business they are in, and manifest a 'one-size-should-fit all' approach to customers. The industry has employed differentiation and niche marketing much less than other industries. As the extensive variety of necessary strategies cannot be comprehensively explored herein, thus one focuses on two new technologies from the IMAX Corporation, DMX and MPX, as an example of how a theater operator might counter audience declines (Silver & McDonnell, 2007).

Lorenzen and Taube (2008) stated that evolutionary perspectives with social network theory in order to explain the recent growth of a prolific and changing indigenous industry in an emergent economy, namely the Indian film industry in Mumbai, Bollywood. The paper demonstrates how the existence of a well-defined and geographically centred social network among producers, directors and other key roles in filmmaking in Mumbai influences the evolution of a 'Bollywood model' of filmmaking remarkably different from Hollywood's. The paper adds to social network perspectives in evolutionary theory by suggesting that, given certain social network structures, policy regulation and other environmental factors may be instrumental for industry evolution.

According to Watson (2004), risk and uncertainty can make sourcing difficult for firms, particularly when there is an asymmetry of power resources between the buyers and suppliers. The present paper explains how risk and uncertainty can be understood and then outlines the structure of the movie supply chain and the power regime structures within it. The movie industry supply chain is one in which the principal commercial challenge is uncertainty. Given the speculative nature of the business, it is only natural that the studios should seek to insure themselves against the possibility that their investments will not be misplaced. One solution is to understand the variations in market demand and time the release of the product accordingly. A second option is to take into account the competition for this timing decision as well. Studios pursue both strategies. In a study by Eliashberg, Weinberg & Hui (2008), it is stated that decision-making style varies across the different parties involved in the production and distribution of movies. Film makers coming from artistic backgrounds tend to believe in more intuitive styles. In contrast, executives in the home video sector, who interact more closely with retailers and consumers generally, see more value in formal decision models. The rise of a new breed of business-educated executives, who are starting to fill high level positions in the supply chain, also encourages the development and application of new models, specifically addressing the unique need of the movie industry. The models chosen for discussion will take the perspectives of different audiences, including academics, movie distributors, movie exhibitors, and public policy decision makers.

According to De Vany & Walls (1999), movies are complex products and the cascade of information among film goes during the course of a film-run could evolve along many paths. It is impossible to attribute the success of a movie to individual causal factors. Audience makes a movie a hit and no amount of star power or marketing can alter that. So, the real star is the movie. Ravid (1999) presents two alternative explanations for the role of stars in motion pictures. Either informed insiders signal project quality by hiring an expensive star, or stars capture their expected economic rent.

Elberse (2007) stated that the involvement of stars is critical to the success of motion pictures. Film studios, which regularly pay multimillion-dollar fees to stars, seem to be driven by that belief. This article sheds light on the returns on this investment using an event study that considers the impact of more than 1200 casting announcements on trading behaviour in a simulated and real stock market setting. The author found evidence that the involvement of stars affects movies' expected theatrical revenues and provides insight into the magnitude of this effect.

Dhar, Sun & Weinberg (2011) stated a 26-year-long database of nationally distributed movies and estimated the prevalence and effectiveness of sequels over time, while controlling for other factors that might influence demand. In particular, they examined whether the effectiveness of a strategy increases over time (possibly due to managerial learning) or decreases (possibly because its effectiveness is competed away or because of consumer satiation). After taking into account both supply side and demand side effects by using simultaneous equations, they find that sequels have a positive effect indirectly (i.e., supply side effect) through a significantly larger number of theaters showing such movies compared to non-sequel movies. In terms of direct effect (i.e., demand side effect), sequels do better than non-sequels in generating more attendance in the first week and in total. Parent movies, the movies from which sequels originate, also do better than non-sequels in terms of total attendance and first-week attendance. Interestingly, sequel movies generate less total attendance than parent movies. Several studies that include an empirical analysis of box office performance for movies produced in Italy between 1985 and 1996 documented a decrease in the total number of films produced and a sharp reduction in daily revenues and per screen daily admissions during the sample period. In the econometric analysis, various alternative hypotheses on the impact of the ex ante popularity of directors and cast of actors on box office performances are rejected in favour of a quadratic specification with positive externalities between the two factors (Bagella & Becchetti, 1999).

The paper examines the relationship between star power and box office revenues using box office data from nine countries and a continuous measure of star power based on the number of visits to a star's web page on IMDB, the most popular web site for movie-related information. The degree of star power is computed for the top star, top three stars, and the director for the films in our sample (Nelson & Glotfelty, 2012).

OBJECTIVES OF STUDY

To investigate and examine what audience likes to see in films.

LITERATURE GAP

A literature review on studies of Bollywood and its ecosystem, how it works or function leads to following analysis to find the research gaps. From literatures review and related works what we scan and found is that there is lack of research in the Indian context about producing Bollywood films which are audience-centric.

So far, very few studies have carried out about producing films based on audience-centricity even in the foreign context. Further, into depth, there is a dearth of sufficient research to find out segregation of age groups among audience concerning the films. This implies that there is a need to understand the way audience takes films. Given the fact that the Indian marketplace is experiencing a new spurt of films - 100 crore club, the need becomes all more relevant. In such scenario, systematic research aimed at measuring the segregation of respondents in regards to age group about watching the Bollywood films. As a result, it is need of hour to know micro-level analysis of audience in totality for increasing the box office revenues of films. It would leave huge scope for marketers, directors, and production houses to recognise the role of audience and its importance in box office collection, and can enable them to roll out audience-centric films in Indian markets. Moreover, the main focus will be on the segregation of respondents into different age groups in respect to know their views on characters in films. However, it would be interesting to find out whether there is any specific age groups have dominance in respect to liking of films. Last but not the least, the study is to check disparity among age groups in respect to preference of theme offFilms.

METHODOLOGY

Since the objective of the study is to explore audience's interest for watching Bollywood films, authors framed question through structured questionnaire-cum-interview method and took their responses on say characters in film, liking in films and preference of theme. Authors segregated their responses into different age groups of respondents by applying cluster analysis.

DATA COLLECTION AND SAMPLE CHARACTERISTICS

The data have been collected from primary sources. It is collected through online and offline methods. It has been divided into different age groups. Cluster analyses were applied to respond for attaining objective of finding characters that influence to watch films.

DATA ANALYSIS AND DATA INTERPRETATIONS ON CHARACTERS THAT INFLUENCES TO WATCH FILMS

Cluster analysis is carried out to understand how much characters influence a viewer to watch films. For that, respondents were asked to rate the given factors on a five

point likert scale, with respect to degree of importance they attach to them.

Cluster analysis is a means of grouping records based upon attributes that make them similar. Grouping similar respondents/customers and variables are the fundamental purpose of research activity and is used prominently in market segmentation. As it is the objective of the researcher to form market segments based on given attributes/variables, cluster analysis allows segments to be formed that are based on data that are less dependent on subjectivity. Objects in a specific cluster share many characteristics but are very dissimilar to objects not belonging to that cluster. Further, it was decided to conduct cluster analysis as here grouping is based on the responses

to several variables unlike factor where grouping is based in patterns of variation.

Tables 1 represent the last six observations of agglomeration schedule for hierarchical clustering. The first column of the table represents the cluster number, second and third show the cluster combined, and fourth the coefficients. Fifth and sixth columns show stages of the cluster. Eighth column represents the absolute change in the clustering coefficients. This is obtained by subtracting the present stage's coefficient value from the previous stage's coefficient value. Ninth column gives percentage change in the coefficient values. To generate the stopping rule, agglomeration coefficients and percentage changes in these coefficients are used to identify the number of clusters.

Table 1: Average Linkage (Between Groups) – Agglomeration Schedule

Agglomeration Schedule								
	Cluster Combined			Stage Cluster first appears			Difference (Diffusion Coefficient)	Percentage Change
	Cluster 1	Cluster 2	Coefficient	Cluster 1	Cluster 2	Next Stage		
1159	56	1064	12.500	1150	1135	1160	0.778	35.30435
1160	56	963	15.940	1159	1145	1164	3.44	342.1594
1161	1	241	16.985	1157	788	1163	1.045	-69.6221
1162	116	942	17.014	1158	1144	1163	0.029	-97.2249
1163	1	116	19.457	1161	1162	1164	2.443	8324.138
1164	1	56	27.412	1163	1160	0	7.955	225.6242

From the column fusion coefficient it can be seen that there is a difference of 7.955 (27.412 – 19.457) between the cluster 1 solution and cluster 2 solution. The next difference is 2.443 (19.457 – 17.014) between cluster 2 and cluster 3 solutions. After that the difference between cluster 3 and cluster 4 solutions falls which is 0.029 (17.014 – 16.985). Then again the difference between cluster 4 and cluster 5 solutions rises which is 1.045 (16.985 – 15.940). This provides a four cluster solution.

Quick Cluster

K-means clustering is also known as quick clustering. The importance of first stage is just to identify appropriate number of clusters used as an input for the k-means clustering. Objective of carrying out second stage is to consider the practical significance of the clusters derived from hierarchical clustering method to meet out the objectives of classification of respondents based on their level of satisfaction. In this stage, clusters are developed

and interpreted. Table 2 represents initial cluster centres for considering four clusters.

Table 2: Initial Cluster Centres

Initial Cluster Centres				
	Cluster			
	1	2	3	4
Impact_of_Actor	5.00	1.00	4.00	1.00
Impact_of_Actress	5.00	1.00	4.00	1.00
Impact_of_Supporting_of_Actor_Actress	5.00	1.00	1.00	1.00
Impact_of_Director	5.00	3.00	1.00	4.00
Impact_of_Musical_Director	5.00	1.00	5.00	5.00
Impact_of_Production_house	5.00	1.00	2.00	5.00

Iteration History

These initial solutions, presented in Table B, are improved by performing a number of iterations. By these iterations,

initial cluster centres are improved and final cluster centres are developed. The final cluster centres are best solution, and are used to interpret average values of each variable for a cluster, and thereby describe clusters. Table 3 represents iteration history for four cluster solutions.

Table 3: Iteration History^a

Iteration	Change in Cluster Centres			
	1	2	3	4
1	2.030	2.610	3.210	3.156
2	.197	.266	.449	1.007
3	.089	.255	.153	.411
4	.045	.395	.120	.114
5	.038	.406	.141	.035
6	.041	.556	.179	.000
7	.036	.412	.140	.075
8	.055	.376	.148	.171
9	.050	.122	.117	.090
10	.025	.321	.113	.069
11	.016	.048	.076	.069
12	.013	.113	.077	.000
13	.010	.089	.064	.017
14	.000	.000	.010	.022
15	.000	.000	.000	.000

a. Convergence achieved due to no or small change in cluster centres. The maximum absolute coordinate change for any centre is .000. The current iteration is 15. The minimum distance between initial centres is 5.745.

Table 3 shows changes in cluster centres. The iteration process stopped when no more changes in the cluster centres is possible, i.e. the changes in all the centres are 0.000. From Table 3, it is evident that in fifteenth iteration the changes are 0.000 for all four clusters. In other words it can be said that fifteen iterations are required for developing final cluster centres for two cluster solution.

Final Cluster Centre

The final cluster centres describe the mean value of each variable for each of the four clusters, for instance, cluster 1 is described by the mean values of variable 1 = 4.45, variable 2 = 4.45 and so on. After deriving the

final clusters, next stage is to interpret the clusters in terms of original variables.

Table 4: Final Cluster Centre

Final Cluster Centres				
	Cluster			
	1	2	3	4
Impact_of_Actor	4.45	3.63	4.35	3.34
Impact_of_Actress	4.45	3.54	4.31	3.02
Impact_of_Supporting_of_Actor_Actress	4.26	2.44	3.12	3.61
Impact_of_Director	4.29	3.02	3.97	4.77
Impact_of_Musical_Director	4.33	2.50	3.79	4.37
Impact_of_Production_house	4.34	2.93	3.27	4.25

Cluster 1

All the studied variables namely actor, actress, supporting actor/actress, director, musical director, and production house are very important for the respondents falling under cluster 1. In other words, respondents of this cluster consider all the above mentioned variables very important while watching films.

Cluster 2

Respondents of the cluster 2 opined that only actor, actress, and director are acting as important variables to watch a movie; while they are neutral regarding impact of variables namely supporting actor/actress, musical director, and production house.

Cluster 3

As per the respondents of the cluster 3, actor and actress are the only two very important variables to impact the decision to watch a movie, while they consider the remaining three variables to be important.

Cluster 4

A very distinguished opinion was derived from the respondents of cluster 4 as director, music director, and production house act as the most influencing variables for them to impact decision about watching a movie. They did not express any specific views about impact of actor, actress, and supporting actor/actress.

ANOVA

Table 5: ANOVA

	Cluster		Error		F	Sig.
	Mean Square	df	Mean Square	df		
Impact_of_Actor	59.733	3	.403	1161	148.115	.000
Impact_of_Actress	90.575	3	.372	1161	243.443	.000
Impact_of_Supporting_of_Actor_Actress	157.796	3	.482	1161	327.190	.000
Impact_of_Director	71.971	3	.491	1161	146.631	.000
Impact_of_Musical_Director	117.396	3	.427	1161	275.140	.000
Impact_of_Production_house	117.870	3	.490	1161	240.379	.000

The F tests should be used only for descriptive purposes because the clusters have been chosen to maximise the differences among cases in different clusters. The observed significance levels are not corrected for this and thus cannot be interpreted as tests of the hypothesis that the cluster means are equal.

From the ANOVA table (Table 5), it can be inferred that all the variables are significant at 99 percent confidence level.

	Valid	1165.000
	Missing	.000

Number of Cases in each Cluster

Table 6 represents number of cases (respondents) in each cluster. From the table it is found that for four cluster solution, 656 and 113 respondents fall in cluster 1 and 2 respectively; while 266 and 130 respondents fall in cluster 3 and 4 respectively. This means that majority of the respondents (656 out of 1165) consider all the variables important to watch a movie.

Table 6: Number of Cases in each Cluster

Cluster	1	656.000
	2	113.000
	3	266.000
	4	130.000

DATA ANALYSIS AND DATA INTERPRETATIONS ON LIKING IN FILMS

After knowing from respondents about preference of basic characters to watch in films, it is required to probe further from respondents to know what they like in films, by applying cluster analysis. Respondents were asked to rate different variables like dancing, acting, dialogue etc. at a five point likert scale, on the basis of what they consider important to see a film.

Table 7: Agglomeration Schedule

Agglomeration Schedule								
	Cluster Combined		Coefficient	Stage Cluster first appears			Difference (Diffusion Coefficient)	Percentage Change
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	Next Stage		
1161	929	1048	33.8	1149	0	1162	3.219	18.91393
1162	1	929	38.362	1160	1161	1164	4.562	41.72103
1163	1021	1132	44	0	7	1164	5.638	23.58615
1164	1	1021	58.276	1162	1163	0	14.276	153.2104

From the column fusion coefficient it can be seen that there is a difference of 14.276 (58.276 - 44) between the cluster 1 solution and cluster 2 solution. The next difference is 5.638 (44 - 38.362) between cluster 2 and

cluster 3 solutions. After that the difference between cluster 3 and cluster 4 solutions falls which is 4.562 (38.362 - 33.8). Then again the difference between cluster 4 and cluster 5 solutions rises which is 3.219 (33.8 - 30.581). This provides a two cluster solution.

Quick Cluster

Table 8 provides k-means clustering also known as quick clustering to know the appropriate number of clusters which are developed.

Table 8: Initial Cluster Centres

	Centre	
	Cluster	
	1	2
Acting	4.00	4.00
Dialogue	5.00	5.00
Nudity	3.00	2.00
Spicy	5.00	2.00
Story	5.00	5.00
Scenic	2.00	5.00
Script	2.00	4.00
Scene	5.00	2.00
Fight	1.00	2.00
Music	5.00	1.00
Item songs	5.00	1.00
Visual	5.00	4.00

Iteration History

These initial solutions, presented in Table 8 are improved by performing a number of iterations. By these iterations, initial cluster centres are improved and final cluster centres are developed. The final cluster centres are best solution, and are used to interpret average values of each variable for a cluster, and thereby describe clusters. Table represents iteration history for two cluster solutions.

Table 9: Iteration History

Iteration	Change in Cluster Centres	
	1	2
1	4.647	4.293
2	.328	.666
3	.111	.229
4	.043	.095
5	.052	.108
6	.036	.071
7	.025	.048
8	.000	.000

a. Convergence achieved due to no or small change in cluster centres. The maximum absolute coordinate change for any centre is .000. The current iteration is 8. The minimum distance between initial centres is 8.124.

Table 9 shows changes in cluster centres. The iteration process stopped when no more changes in the cluster centres is possible, i.e. the changes in all the centres are 0.000. From Table 9 it is observed that in eighth iteration the changes are 0.000 for all four clusters. In other words, it can be said that eight iterations are required for developing final cluster centres for two cluster solution.

Table 10: Final Cluster Centres

Centre	Cluster	
	1	2
	Acting	4.49
Dialogue	4.54	4.29
Nudity	4.12	2.31
Spicy	4.24	3.00
Story	4.51	4.35
Scenic	4.14	3.74
Script	4.25	4.04
Scene	4.16	3.77
Fight	4.06	3.28
Music	4.32	4.19
Item Songs	4.36	3.36
Visuals	4.26	3.82

Cluster 1

For the respondents of the cluster one, all variables are very important; however, they give most importance to acting and dialogue among the given set of attributes.

Cluster 2

Variables namely acting, dialogue, story, script, and music are very important variables while variables namely *masala* in movie, fight, item songs, and visuals are important for them.

Table 11: Distances between Final Cluster Centres

Cluster	Centre	
	1	2
1		2.665
2	2.665	

Table 12: ANOVA

	Cluster		Error		F	Sig.
	Mean Square	df	Mean Square	df		
Acting	5.063	1	.348	1163	14.564	.000
Dialogue	16.248	1	.321	1163	50.685	.000
Nudity	846.389	1	.749	1163	1130.419	.000
Masala_in_movie	399.519	1	.618	1163	646.448	.000
Story	6.575	1	.348	1163	18.889	.000
Chronology_Sequence	42.410	1	.704	1163	60.259	.000
Script	11.911	1	.504	1163	23.630	.000
Scene	40.612	1	.548	1163	74.107	.000
Fight	159.871	1	.808	1163	197.902	.000
Music	4.895	1	.508	1163	9.641	.002
Item_songs	262.769	1	.715	1163	367.571	.000
Use_of_Technology	51.513	1	.704	1163	73.212	.000

The F tests should be used only for descriptive purposes because the clusters have been chosen to maximise the differences among cases in different clusters. The observed significance levels are not corrected for this and thus cannot be interpreted as tests of the hypothesis that the cluster means are equal.

From the ANOVA table (Table 12), it can be inferred that all the variables are significant at 99 percent confidence level.

Table 13 represents number of cases (respondents) in each cluster. From the table it is found that for two cluster solution 773 and 392 respondents fall under cluster 1 and 2 respectively.

Table 13: Number of Cases in each Cluster

Number of Cases in each Cluster		
Cluster	1	773.000
	2	392.000
	Valid	1165.000
	Missing	.000

DATA ANALYSIS AND DATA INTERPRETATIONS ON SUBJECT OR THEME OF FILMS

Cluster Analysis is conducted to know from respondents about preference for theme or subject of a film.

Hierarchical Clustering

Table 14 represents the last 15 observations of agglomeration schedule for hierarchical clustering. The first column of the table represents the cluster number, second and third columns show the cluster combined and fourth the coefficients. Fifth and sixth columns show stages of the cluster. Eighth column represents the absolute change in the clustering coefficients. This is obtained by subtracting the present stage's coefficient value from the previous stage's coefficient value. Ninth column gives percentage change in the coefficient values. To generate the stopping rule, agglomeration coefficients and percentage changes in these coefficients are used to identify the number of clusters.

Table 14: Average Linkage (Between Groups) – Agglomeration Schedule

Agglomeration Schedule								
	Cluster Combined		Coefficient	Stage Cluster first appears			Difference (Diffusion Coefficient)	Percentage Change
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	Next Stage		
1159	1	5	32.862	1158	1157	1160	2.14	-22.9384
1160	1	44	34.138	1159	1149	1161	1.276	-40.3738
1161	1	138	38.35	1160	271	1162	4.212	230.094
1162	1	83	43.365	1161	487	1163	5.015	19.06458
1163	1	180	47.481	1162	888	1164	4.116	-17.9262
1164	1	58	62.388	1163	1155	0	14.907	262.172

From the column fusion coefficient it can be seen that there is a difference of 14.907 (62.388 - 47.481) between the cluster 1 solution and cluster 2 solution. The next difference is 4.116 (47.481- 43.365) between cluster 2 and cluster 3 solution. After that the difference between cluster 3 and cluster 4 solutions rises which is 5.015 (43.365 - 38.35). Then again the difference between cluster 4 and cluster 5 solutions falls which is 4.212 (38.35– 34.138). This provides a three cluster solution.

	Centre		
	1	2	3
Regarding_Tradition	1.00	1.00	4.00
Politic	5.00	1.00	3.00
Science	5.00	1.00	4.00
Culture_art	2.00	2.00	5.00
Social	5.00	1.00	4.00
War	5.00	1.00	1.00

Quick Cluster

K-means clustering is also known as quick clustering. The importance of first stage is just to identify appropriate number of clusters used as input for the k-means clustering. Objective of carrying out second stage is to consider the practical significance of the clusters derived from hierarchical clustering method in meeting the objectives of classification of respondents based on their level of satisfaction. In this stage, clusters are developed and interpreted. Table 15 represents initial cluster centres for considering three clusters.

Table 15: Initial Cluster Centres

	Cluster		
	1	2	3
Biographies	4.00	1.00	4.00
Contemporary	1.00	2.00	5.00
Regarding_spirit	1.00	1.00	5.00
Crime	5.00	1.00	1.00
Gangster_underworld	5.00	2.00	2.00
History	5.00	1.00	4.00

contd.

Iteration History

These initial solutions, presented in Table 16 are improved by performing a number of iterations. By these iterations, initial cluster centres are improved and final cluster centres are developed. The final cluster centres are best solution, and are used to interpret average values of each variable for a cluster, and thereby describe clusters. Table represents iteration history for three cluster solutions.

Table 16: Iteration History^a

Iteration	Change in Cluster Centres		
	1	2	3
1	4.751	3.762	4.122
2	.461	.695	.210
3	.390	.345	.294
4	.268	.080	.249
5	.239	.065	.193
6	.175	.102	.154
7	.189	.089	.138
8	.157	.060	.099
9	.131	.129	.080

contd.

10	.219	.106	.105
11	.303	.102	.126
12	.114	.094	.052
13	.000	.066	.017
14	.000	.075	.021
15	.000	.072	.021
16	.000	.000	.000

a. Convergence achieved due to no or small change in cluster centres. The maximum absolute coordinate change for any centre is .000. The current iteration is 16. The minimum distance between initial centres is 9.110.

Table 16 shows changes in cluster centres. The iteration process stopped when no more changes in the cluster centres is possible, i.e. the changes in all the centres are 0.000. From Table 16, it is observed that in 16th iteration the changes are 0.000 for all three clusters. In other words, it can be said that 16 iterations are required for developing final cluster centres for three cluster solution.

Table 17: Final Cluster Centre

	Centre		
	Cluster		
	1	2	3
Biopic	4.22	2.10	3.21
Contemporary Issues	3.84	2.94	3.29
Corruption	4.03	3.00	3.32
Crime	3.87	2.25	3.25
Gangster_underworld	4.07	2.52	3.20
Historical	4.45	2.41	3.46
Love	4.09	2.63	3.28
Political Issues	4.41	1.98	3.48
Science Fiction	4.46	2.92	3.64
Sequel	4.25	2.79	3.38
Social Issues	4.46	2.70	3.47
War/Freedom	4.36	2.20	3.50

Table 19: ANOVA

	Cluster		Error		F	Sig.
	Mean Square	df	Mean Square	df		
Biopic	257.165	2	.628	1162	409.426	.000
Contemporary Issues	49.732	2	.698	1162	71.295	.000
Corruption	70.038	2	.897	1162	78.123	.000
Crime	149.589	2	.766	1162	195.328	.000
Gangster_underworld	143.310	2	.769	1162	186.292	.000
Historical	238.582	2	.568	1162	420.110	.000
Love	127.480	2	.637	1162	200.122	.000
Political Issues	336.897	2	.662	1162	508.557	.000
Science Fiction	137.610	2	.682	1162	201.696	.000
Sequel	131.222	2	.592	1162	221.496	.000
Social Issues	184.575	2	.689	1162	268.057	.000
War/Freedom	265.317	2	.723	1162	366.961	.000

The F tests should be used only for descriptive purposes because the clusters have been chosen to maximise the differences among cases in different clusters. The observed significance levels are not corrected for this and thus cannot be interpreted as tests of the hypothesis that the cluster means are equal.

Cluster 1

For the respondents of this cluster all the variables are very important except contemporary issues and crime.

Cluster 2

The results derived for the cluster 2 reveal that the respondents of this cluster are very choosy and have clearly expressed their choices. They do not like movies based on politics, biographies, war and crime and are neutral for movies based on themes like contemporary issues, corruption, gangster, history, love, science, and sequel.

Cluster 3

The respondents of cluster 3 are neutral for the movies based on subjects related to spirit, tradition, culture and art, history, gangster, contemporary, ethics, science, business person, life style, freedom, and social issues. Movies based on this said themes are neither important nor unimportant for them. They do not prefer movies based on crime, biographies, war and politic. They do not have strong liking for a particular type of movie but for movies based on inspiration and successful person.

This shows particular liking towards movie based on inspiration and successful person.

Table 18: Distances between Final Cluster Centres

Cluster	Centre		
	1	2	3
1		5.988	2.938
2	5.988		3.135
3	2.938	3.135	

From the ANOVA table (Table 19), it can be inferred that all the variables are significant at 99 percent confidence level.

Table represents number of cases (respondents) in each cluster. From the table it is found that for three cluster solution, 270 and 196 respondents fall under cluster 1 and 2 while 696 respondents fall under cluster 3.

Table 20: Number of Cases Inin Each Cluster

Cluster	1	270.000
	2	196.000
	3	699.000
	Valid	1165.000
	Missing	.000

IMPLICATIONS

This study adds empirical and theoretical depth to debate on audience’s approach toward Bollywood films with respect to their views on character, liking in films, and preference of theme in films. The role of audience in film-making process is not actively researched topic in India. The study therefore contributes toward better understanding to directors, marketers, and production houses. Every film is a new product. Each film is produced by millions of dollars at stake; even statistic said that only less than 5% of films get break-even (no profit or no loss). So, this study eventually helps to many stakeholders especially production houses to produce audience-centric films. It also helps marketers to use innovative promotional tools for promoting Bollywood films in India and overseas market. Hence, it reflects the paramount importance of audience’s likes and dislikes regarding films.

Most of the studies are conducted with reference only to developed countries. Therefore, there is a need to study the above-mentioned dimensions from the context of developing countries like India where the highest number of films are produced annually in the world. It also helps to get to know the role of audience in box-office revenues in general, and films choice in particular. The research is studied with respect to different multiple dimension together and in comprehensive manner which leads to contribute better decision for producing and promotion of a films in Indian market.

Fundamentally, three major implications of the study could thus be summarised as follows. First, this study supported the existing literatures regarding role of audience approach towards Bollywood films at comprehensive level. Second, the study investigated at micro level (e.g., segmenting

audience into age groups) in respect to audience approach towards films in India. Lastly, the study also acted as an eye-opener for film industry that can take a cue from the findings and use it for future release of films in Indian market.

SCOPE OF FUTURE RESEARCH

Virtually any study in the traditional form which represents findings is subject to various limitations and caveats. The authors have taken consideration of audience approaches toward Bollywood films and look audience in prism of their views, liking, and preference of theme in films. But, the list is not inclusive and even many more variables can be added at will. However, the authors have put some light to produce films keeping them audience-centric. Since each film is a new business and it is universally accepted that customer satisfaction is the core to any business, the authors attempted to address the issue of producing films by taking audience into consideration. The study is carried on Bollywood films which are the one of the major segments of Indian film industry. There are also other major segments in the industry like Tamil film industry, Telugu film industry etc. Hence whether their likings, dislikes etc. would vary or remain same is a matter of debate and altogether another research study. In sum, all the caveats and limitations just enumerated support the need for future investigations by using data from multiple time periods, multiple countries and multiple reviewers. Refreshingly, availability of movie and entertainment related data on the Internet and elsewhere has grown up by leaps and bounds in recent years. This progress augurs well for the potential feasibility of future studies aimed at addressing the issues just raised. In the case of motion pictures, they may well differ in other countries or at other times. Future research should address the generalisability of our findings across time periods and geographical settings. It is proposed the further research needs to be done for the validation of findings of this study.

REFERENCES

Ainslie, A., Dreze, X., & Zurfrayden, F. (2005). Modeling movie life cycles and market share. *Marketing Science*, 24(3), 508-517.

Bagella, M., & Becchetti, L. (1999). The determinants of motion picture box office performance: Evidence from movies produced in Italy. *Journal of Cultural Economics*, 23(4), 237-256.

De Vany, A., & Walls, D. (1999). Uncertainty in the movie industry: Does star power reduce the terror of the box office? *Journal of Cultural Economics*, 23, 285-318.

- Drake, P. (2008). Distribution and marketing in contemporary Hollywood. *The Contemporary Hollywood Film Industry* edited by Wasko, J and McDonald, (pp. 63-82). Oxford and New York: Blackwell.
- Elberse, A. (2007). The power of stars: Do star actors drive the success of movies?. *Journal of Marketing*, 71(4), 102-120.
- Eliashberg, J. A., Elberse, A., Mark, A., & Leenders, M. (2006). The motion picture industry: Critical issues in practice, current research, and new research directions. *Marketing Science*, 25(6), 638-661.
- Eliashberg, J., Hui, S. K., & Zhang, Z. J. (2007). From story line to box office: A new approach for greenlighting movie scripts. *Journal of Management Science*, 53, 881-893.
- Eliashberg, J., Weinberg, C. B., & Hui, S. K. (2008). Decision models for the movie industry. In *Handbook of marketing decision models* (pp. 437-468). Springer US.
- Eliashberg, J., Swami, S., Wierenga, B., & Weinberg, C. B. (2001). Implementing and evaluating Silver Screener: A marketing management support system for movie exhibitors. *Interfaces*, 31(3), S108-S127. May-June.
- European Audiovisual Council, Council of Europe (2009).
- Goldman, W. (1983). *Adventures in the screen trade*. Grand Central Publishing.
- Holbrook, M. B., & Addis, M. (2007). Taste versus the market: An extension of research on the consumption of popular culture. *Journal of Consumer Research*, 34(October), 415-424.
- India Entertainment and Media Outlook (2012). Annual report: Confederation of Indian Industry and PwC India.
- Krishnan, T. V., & Sakkthivel, A. M. (2010). To push for stardom or not: A rookie's dilemma in the Tamil movie industry. *IIMB Management Review*, 22(3), 80-92.
- Lorenzen, M., & Taube, F. A. (2008). Breakout from bollywood? The roles of social networks and regulation in the evolution of Indian film industry. *Journal of International Management*, 14(3), 286-299. doi. 10.1016/j.intman.2008.01.004
- McKenzie, J. (2010). The economics of movies: A literature survey. *Journal of Economic Surveys*, 10(1), 1-37.
- Nelson, R. A., & Glotfelty, R (2012). Movie stars and box office revenues: An empirical analysis. *Journal of Cultural Economics*, 36(2), 141-166.
- Ravid, S. A. (1999). Information, blockbusters, and stars: A study of the film industry. *The Journal of Business*, 72(4), 463-492.
- Sarkar, S., & Nayak, A. (2009). Film Merchandising in India. *GMJ*, 3(1), 1-10.
- Silk, A. J., & Glen, L. (1978). Urban pre-test market evaluation of new packaged goods: A model and measurement methodology. *Journal of Marketing Research* 15(2), 171-91.
- Sridhar, S. N., & Mattoo, N. K. (1997). *Ananya: A portrait of India*. New York: Associations of Indians in America.
- Srinivas, L. (2002). The active audience: Spectatorship, social relations and the experience of cinema in India. *Media, Culture & Society*, 24, 155-173.
- McDonnell, J. J., & Silver, J. D. (2007). Are Movie Theatres doomed? Do exhibitors see the big picture as theatres lose their competitive advantage? *Business Horizons*, 50(6), 491-501.
- Tirumala, L. N. (2009). Bollywood movies and identity construction amongst second generation Indian Americans. *Journal of Academia education*, 1-27.
- Watson, G. (2004). Uncertainty and contractual hazard in the film industry: managing adversarial collaboration with dominant suppliers. *Supply Chain Management: An International Journal*, 9(5), 402-409, doi: 10.1108/13598540410560784.