Exploring the Relationship Between Viewer Experience and Movie Genre – A Study Based on Text Mining of Online Movie Reviews

Urszula Świerczyńska-Kaczor*

This paper aims to provide insight about viewer experience with movies of different genres: the suspense movie, the western and the comedy.

The qualitative and quantitative analyses based on text mining were conducted for online reviews of the three selected movies. The results of text mining were complemented by a qualitative manual analysis of reviews.

The results of text mining indicated the spectrum of factors shaping viewer experience at different stages of the consumer journey. These factors relate to, for example, the product features (e.g. the recognized movie director, movie stars, soundtrack, and the historical context of the plot), and different touch points (e.g. movie distribution channel). The categories of viewer experience captured in the study can be assessed as adequate to the genre.

The present study is preliminary in nature and is therefore exploratory.

The results indicate the potential usefulness of text mining of online reviews as a method constituting the background for studies based on interviewing subjects. The study also points to the importance of looking for interdisciplinary frameworks in the research field of viewer experience.

Keywords: text mining of online reviews, viewer experience.

Relacja pomiędzy doświadczeniami widza a gatunkiem filmowym – badanie oparte na *text mining* recenzji filmowych

Niniejsza publikacja oferuje wgląd w doświadczenia widzów związane z konsumpcją filmów należących do różnych gatunków, takich jak suspens, western oraz komedia.

Badanie zostało przeprowadzone z wykorzystaniem jakościowej i ilościowej analizy *text mining* recenzji trzech wybranych tytułów filmowych. Wyniki analizy opartej na *text mining* uzupełnia badanie jakościowe recenzji filmowych.

Wyniki badania pozwoliły na uchwycenie spektrum czynników oddziałujących na doświadczenia widzów na różnych etapach konsumpcji produktu, jakim jest film – podróży zakupowej konsumenta. Czynniki te związane są m.in. z cechami produktu (np. marka reżysera, gwiazdorska obsada, ścieżka dźwiękowa,

---

*   **Urszula Świerczyńska-Kaczor** – prof. ucz., dr hab. Faculty of Film Art Organization, The Polish National Film, Television and Theatre School in Lodz, Poland. https://orcid.org/0000-0001-5368-0247.

Correspondence address: The Polish National Film, Television and Theatre School in Lodz, Targowa 61/63, 90-323 Lodz, Poland; e-mail: uswierczynska@filmschool.lodz.pl.

historyczny kontekst fabuły filmu) oraz różnymi punktami „kontaktu" kreującymi doświadczenia widza (np. kanał dystrybucji filmu). Wyodrębnione kategorie doświadczeń widza odpowiadają analizowanym gatunkom filmowym.

Prezentowane w niniejszej publikacji badanie ma charakter badań wstępnych, eksploracyjnych.

Słowa kluczowe: text mining recenzji filmowych, doświadczenia widza.

JEL: M30, M31

1. Introduction

This paper aims to provide insight about the viewer experience with movies of different genres: the suspense movie, the western and the comedy. However, apart from investigating various dimensions of viewer experience, this paper also aims to point out that the 'traditional' field of research of viewer experience can significantly benefit from other disciplines, particularly from consumer neuroscience. This 'call' for a more interdisciplinary approach was reflected in the selection of the research topic and the study design – the present study was inspired by a particular study within the field of neurocinematics.

The paper is structured as follows. The next two sections – 'Literature Background' and 'Research Question' – present the investigated topic and its literature overview. These sections are followed by the description of how the present study was conducted ('Study Design'), what the key findings are ('Findings') and what limitations of the study are ('Limitations of the Study'). The paper ends with the section pointing to the conclusions and future research.

2. Literature Background

2.1. The Approach to Viewer Experience in the Study Presented in This Paper

In this paper, the viewer experience is analyzed through the perspective of customer experience – the construct which is widely defined in literature as *"[…] a multidimensional construct focusing on a customer's cognitive, emotional, behavioral, sensorial, and social responses to a firm's offerings during the customer's entire purchase journey"* (Lemon & Verhoef, 2016, p. 71). So far, there is no single approach to the measurement of consumer experience – the methodologies vary significantly as they draw from different theoretical backgrounds, e.g. consumer-based or organization-based perspective, static or dynamic analysis, the features or outcomes of the consumer journey (see, e.g., Lemon & Verhoef, 2016; Kranzbühler, Kleijnen, Morgan, & Teerling, 2018). As to the methods of gathering data,

the majority of studies conducted in the field of consumer experience are based on the implementation of 'standard' techniques such as gathering data by interviewing subjects (e.g. open-structured individual interviews) or observation (e.g. consumer behavior during service consumption, the interaction between an employee and a consumer, the customer's interaction with such touch points as corporate websites). Moreover, the process of examining the dynamics of viewer experience and viewer journey requires the selection of methods which are adequate to the nature of each stage at which the experience is created. Text mining of online reviews seems to be a useful method of exploring viewer experience which is expressed after or before the actual movie watching; in contrast, consumer neuroscience opens up new possibilities to analyze viewer experience at the stage of actual product consumption. Both methods of analysis – text mining of online reviews and neurocinematics – are briefly presented below.

2.2. Text Mining of Online Movie Reviews – What Aspects of Viewer Experience Can Be Captured by Text Mining?

Online movie reviews written by customers constitute the word-of-mouth (WOM) or rather the word-of-click about the movie. It can be argued that the WOM about a particular movie can reflect many aspects of viewer experience, such as:
– the viewer's expectation before watching the film, reflected in such reviews as, e.g., 'I can't wait to see the next work of this director' (WOM referring to the pre-consumption stage);
– the viewer's evaluation of the experienced movie, reflected in such reviews as, e.g., 'Amazing', 'The music is awesome' (WOM referring to the post-consumption stage);
– the viewer's evaluation of the spectrum of factors or touch points which frame the viewer experience reflected in reviews referring to, e.g., the distribution channel (e.g. 'I can't play the DVD on my player'), the social aspect of film watching ('I watched the movie with my grandson').

So far, the majority of studies presented in the literature – with reference to text mining and online movie reviews –indicate other aims than the contribution to the field of research of consumer (viewer) experience. For example, many studies which include text mining of movie online reviews refer to problems lying in the field of IT studies or mathematics as, for example, 'how to conduct the sentiment and opinion analysis', 'how to predict movie popularity', 'how to predict movie rating by reviewers' (e.g. Shaziya, Kavitha, & Zaheer, 2015; Singh, Saxena, Singh, & Rajendran, 2017; Thorat & Priya, 2018). The next interesting scope of research points to the relationships between the movie reviews and the movie sale or movie attendance (e.g. Lee, Jung, & Park, 2017; Niraj & Singh, 2015; Rui, Liu, & Whinston, 2013). So far, few studies have focused on viewer experience and viewer journey as the focal point of text mining of online reviews (e.g.

Świerczyńska-Kaczor & Kotlińska, 2017; Świerczyńska-Kaczor & Wachowicz, 2017).

2.3. Consumer Neuroscience as an Approach to Research of Viewer Experience

The methods applied in consumer neuroscience – for example, functional Magnetic Resonance Imaging (fMRI) or electroencephalography (EEG) – aim to capture the consumer's neural response to a stimulus at the moment when experience to the stimulus is created (a film, a video or printed advertisement, information about prices, a product packaging are exemplary stimuli used in studies). 'Neurocinematics' – with its reference to the movie and cinema market – is a relatively new area within consumer neuroscience contributing to the better understanding of viewer experience (e.g. Świerczyńska-Kaczor, 2019).

To present the nature of neurocinematics more clearly, it is worth pointing to research conducted by U. Hasson, O. Landesman, B. Knappmeyer, I. Vallines, R. Rubin, and D. Heeger (2008) presented in the seminal paper "*Neurocinematics: The Neuroscience of Film*". The outline of the study design and findings are presented below.

In one of the fMRI (functional Magnetic Resonance Imaging) studies conducted by U. Hasson et. al. (2008), the researchers evaluated and compared the neural responses measured as the level of inter-subject correlation (ISC) while viewers were watching the following films:

1. The episode of "Alfred Hitchcock Presents": "Bang! You're Dead" (1961), directed by Alfred Hitchcock;
2. The western "The Good, The Bad, and The Ugly" (1961), directed by Sergio Leone;
3. The episode of the TV series "Curb Your Enthusiasm" (2000), created by and starring Larry David;
4. An 'unstructured segment of reality video' – the video presented shots from Washington Square Park, intentionally without 'artistic' features.

The main results of the experiment indicated above were as follows: *"The extent of ISC differed for the four movies […]. The percentage of cortex exhibiting high ISC provided a measure of the overall effectiveness, or collective engagement power, of each movie to induce similar responses across viewers […]. The Hitchcock episode […] evoked similar responses across all viewers in over 65 percent of the cortex, indicating a high level of control of this particular episode on viewers' minds. The high ISC was also extensive (45%) for the Good, the Bad and the Ugly […], but much less so (18%) for Curb Your Enthusiasm […]. Finally, as noted above, the unstructured segment of reality […] induced high ISC only in a small fraction of the cortex (less than 5%)."* (Hasson et al., 2008, p. 14, the ellipses in this quotation refer to the graphical figures which are not presented in this paper).

Hasson et al. (2008) stressed that the high level of ISC does not imply artistic merits of the film, but it means that some cinematic styles make viewers' brains 'tick together' – viewers' neural responses measured across the viewers' group are similar while watching a film.

The study reference above is a salient example of studies within the field of neurocinematics. However, nowadays the movie is often used as a stimulus in fMRI studies conducted in other areas than consumer neuroscience, e.g. in the field of medicine, psychology – these studies also deliver results pointing to different aspects of viewer experience. For example, in the case of a comedy or a suspense movie, the literature relating to the psychological and/or neural processes of experiencing suspense or humor offers the scientific background for understanding viewer experience related to these genres. Examples of the investigated topics are: the development of the general, psychological model of tension and suspense (Lehne & Koelsch, 2015), the investigation of the neural correlates of emotional experience of suspense in the case when the stimulus is a suspenseful novel (Lehne, Engel, Rohrmeier, Menninghaus, Jacobs, & Koelsch, 2015), the investigation of neural correlates between the suspense and the viewer's attention focus (Bezdek, Gerrig, Wenzel, Shin, Pirog Revill, & Schumacher, 2015), the investigation of the neural correlates for humor processing taking into account different variables, e.g. non-verbal, different humor styles (e.g. Chan, Hsu, Liao, Chen, Tu, & Wu, 2018; Samson, Zysset, & Huber, 2008; Vrticka, Black, & Reiss, 2013; Watson, Matthews, & Allman, 2007; Wild, Rodden, Grodd, & Ruch, 2003).

3. Research Question

This study aimed to capture features of consumer (viewer) experience with films of different genres and the main research question was formulated as:

– What kind of viewer experience – relevant for movies of different genres – can be revealed by text mining of online reviews?

However, there is also the 'second layer' of the conducted analysis which refers to neurocinematics. The analyzed aspect of viewers' brains 'synchronization' or 'ticking together' investigated in the abovementioned study of U. Hasson et al. (2008) was reflected in this study in the question: "Does text mining of online reviews reflect the fact that viewers' experiences are somehow more 'synchronized' for some films than others?"

This 'second layer' of analysis was reflected in the selection of movie titles. Therefore, in this study the following movies were analyzed (Table 1):

– The film "Psycho" (1960), directed by Alfred Hitchcock;
– The western "The Good, The Bad, and The Ugly" (1961), directed by Sergio Leone;

– The first season of the TV series "Curb Your Enthusiasm" (2000), created by and starring Larry David.

These movies did not exactly match the movies analyzed by U. Hasson et. al. (2008) for the following reasons:

– There were too few reviews for conducting text mining of reviews of "Alfred Hitchcock Presents" and the episode of "Bang, You're Dead" (1961) (an episode within season 7) – the reviews which were available on the Rotten Tomatoes website, IMDb (Internet Movie Database) or Amazon.com. The movie selected for analysis – "Psycho" (1960) – has been acknowledged by professional film critics as one of the finest Hitchcock's masterpieces (see, for example, the professional critics' reviews on Rotten Tomatoes);

– in the analysis of "Curb Your Enthusiasm", the analyzed reviews related to the whole first season aired in 2000 (including such episodes as: "The Pants Tent" and "Ted and Mary"), without reference to a particular episode.

This study was exploratory in nature, and I looked for the various indicators – findings of qualitative and quantitative analyses based on text mining of online reviews – which may point to viewers' experiences with the suspense movie, the western movie and the comedy movie.

4. Study Design

The process of sampling the data was designed as follows:

1. In August 2019, reviews were randomly sampled from the pool of audience reviews available on Rotten Tomatoes and Amazon.com (Table 2). The analysis included 624 online reviews of "Psycho" (1960), 806 reviews of "The Good, The Bad, and The Ugly" (1966) and 600 reviews of the first season of "Curb Your Enthusiasm" (2000). The distributions of 'stars'– scores given by a reviewer as the quantified metric of movie evaluation – in the sample seem to reflect the general evaluation of the movie on the basis of the Tomatoemeter (Rotten Tomatoes website) or by Amazon rankings (Table 2).

2. Only English language reviews were included.

3. Reviews were analyzed without considering the channel on which the film was watched: DVD, television, streaming services, or – even in some cases – in cinemas.

The qualitative and quantitative analyses were conducted with the implementation of Statistica software and KH Coder software. Additionally, viewers' experiences were also analyzed by conducting a manual text analysis of 50 reviews of each movie – the main topics and sub-topics were extracted for each review.

The movie	Director	Genre	Selected movie awards received
"Psycho" (1960)	Alfred Hitchcock	horror, mystery, thriller	– Nominee for Oscars in 1961 in the following four categories: Best Director; Best Actress in a Supporting Role; Best Cinematography, Black-and-White; Best Art Direction-Set Decoration, Black-and-White; – Winner of Golden Globes in 1961 in the category Best Supporting Actress
"The Good, the Bad and The Ugly" (1966)	Sergio Leone	western	– Laurel Awards in 1968 in category Action Performance (Clint Eastwood)
"Curb Your Enthusiasm". Creator: Larry David	First season – 2000; the premiere episode "The Pants Tent" (2000) directed by Robert B. Weide, staring Larry David	TV comedy series	Awards over all seasons: – Golden Globes in 2006: Nominee for Best Television Series – Comedy or Musical, Best Performance by an Actor in a Television Series – Comedy or Musical (Larry David); – Golden Globes in 2003: Winner in Best Television Series – Comedy or Musical, Best Performance by an Actor in a Television Series – Comedy or Musical (Larry David) – AFI Awards, USA in 2002: Winner – Comedy Series of the Year

Tab. 1. Attributes of analyzed movies – the director, the genre and selected awards received. Source: Based on information on the website: IMDb website (https://www.imdb.com/). (01.08.2019).

	"Psycho" (1960)		"The Good, the Bad and The Ugly" (1966)		"Curb Your Enthusiasm" (2000)	
	Number of reviews	%	Number of reviews	%	Number of reviews	%
1 star	9	1.44%	24	2.98%	23	3.83%
2 stars	7	1.12%	14	1.74%	32	5.33%
3 stars	18	2.88%	31	3.85%	46	7.67%
4 stars	80	12.82%	88	10.92%	106	17.67%
5 stars	510	81.73%	649	80.52%	393	65.50%
	624	100.00%	806	100.00%	600	100.00%

continued Table 2

	"Psycho" (1960)		"The Good, the Bad and The Ugly" (1966)		"Curb Your Enthusiasm" (2000)	
	Number of reviews	%	Number of reviews	%	Number of reviews	%
Rotten Tomatoes and Amazon overall evaluation	– Rotten Tomatoes rating: Certified Fresh – 96% Tomatometer, 95% – audience score – Amazon reviews score from 1598 reviews: 4.7 out of 5 points; 81% of reviews scored 5 points.		Rotten Tomatoes rating: Certified Fresh – 97% Tomatometer, 97% – audience score Amazon reviews score from 2115 reviews: 4.7 out of 5 points; 81% of reviews scored 5 points.		Rotten Tomatoes rating for the first season: – 89% Tomatometer, 95% – audience score Amazon reviews score from 4784 reviews of the first season: 4.4 out of 5 points; 71% of reviews scored 5 points.	

Note: The scores of reviews sampled from Rotten Tomatoes were adjusted to the next integer: from 0.5 point to 1 point – 1 point, from 1.5 to 2 points – 2 points, from 2.5 to 3 – 3 points, from 3.5 to 4 points – 4 points, from 4.5 points to 5 points – 5 points.

Tab. 2. Distribution of review scores within the sample. Source: The author's own analysis based on online reviews sampled from 01.08.2019 to 20.08.2019 from the following websites: Rotten Tomatoes (https://www.rottentomatoes.com) and Amazon (https://www.amazon.com).

5. Findings

The findings – presented below in the form of tables and figures – were selected from the overall results, considering their relevance to the further discussion of the research problem (text mining provides a very large volume of data).

The summary of the main findings is as follows:

1. The overall results indicated a very broad spectrum of factors which impact on the viewer experience with the western, the suspense movie or the comedy series (summarized in manual analyses – Table 5, Table 6, Table 7).
2. The analysis pointed to 'general factors' relevant for all genres such as viewers' expectations for the genre, viewers' expectations for the main characters in the movie, viewers' expectations based on the previous work of the movie director/creator (Alfred Hickock, Sergio Leone, Larry David), viewers' characteristics including their individual sense of humor.
3. Apart from 'obvious' categories of viewer experience – strongly related to the genre – the analysis pointed to 'surprising aspects' such as the impact of a historical event ('The Bad, The Good, The Ugly").
4. The analysis of the horror movie "Psycho" highlighted the difference between 'suspense' perceived and 'suspense' felt (reviews such as 'not so scary as I expected') – which relate to the contemporary viewer's

expectations for the genre (e.g. Lehne and Koelsch (2015) indicated the differences between the suspense perceived and the subjectively felt suspense).

5. The analysis of the western movie "The Good, the Bad, and the Ugly" highlighted the significance of the soundtrack. It may be presumed that in the case of this movie the music also influences (enhances) the level of suspense perceived by the viewer.

6. The analysis of the TV comedy series "Curb Your Enthusiasm" indicated a new trend in consumer behavior: 'binge-watching'.

7. The 'similarities' of viewer experience – for "Psycho" and "The Good, the Bad, The Ugly" – may be explained by that both films include suspense and that both movies are 'classics' of their genres. Although "Psycho" (1960) is 'the perfect example' of a suspense, horror, thriller movie, it may be argued that the movie "The Good, The Bad, and The Ugly" (1966) classified as a western is also a suspense movie (the main plot creates viewer uncertainty over how it will develop – 'Will Blondi, Tuco and Angel Eyes find the money?' – and the movie includes many suspenseful scenes, each following one another, e.g. 'Will Tuco be hanged?', 'Will Tuco and Blondi blow up the bridge?').

8. The visual analysis indicates that the comedy "Curb Your Enthusiasm" may be perceived differently in comparison with "Psycho" and "The Good, the Bad, the Ugly" (Figure 5, Figure 6, Figure 7).

"Psycho" (1960)		"The Good, The Bad, The Ugly" (1966)		"Curb Your Enthusiasm" (2000)	
Concept number	31	Concept number	27	Concept number	33
Concept 1	73.77	Concept 1	64.99	Concept 1	43.54
Concept 2	25.38	Concept 2	26.58	Concept 2	23.7
Concept 3	21.44	Concept 3	24.42	Concept 3	20.12
Concept 4	20.45	Concept 4	22.09	Concept 4	19.28
Concept 5	19.83	Concept 5	20.6	Concept 5	18.33
Concept 6	18.55	Concept 6	20.3	Concept 6	16.92
Concept 7	18.10	Concept 7	19.74	Concept 7	16.07
Concept 8	16.92	Concept 8	17.52	Concept 8	15.07
Concept 9	16.48	Concept 9	17.26	Concept 9	14.71
Concept 10	16.06	Concept 10	15.89	Concept 10	14.00

Tab. 3. The number of SVD concepts and the singular value of the top ten concepts. Source: The author's own analysis (Statistica software).

"Psycho" (1960)		"The Good, The Bad, The Ugly"(1966)		"Curb Your Enthusiasm" (2000)	
movi	100.00	movi	100.00	show	100.00
psycho	39.88	western	54.78	larri	97.38
hitchcock	37.66	one	42.20	funni	84.40
classic	34.12	great	41.23	watch	78.73
watch	32.00	best	40.29	like	75.20
one	31.55	time	38.27	season	72.75
time	30.85	good	37.58	seinfeld	71.85
great	30.48	classic	37.07	love	69.39
scene	27.68	star	36.76	episod	66.27
horror	27.03	watch	33.43	larry david	62.24
love	24.03	eastwood	32.76	great	52.07
best	22.86	five	30.19	seri	49.53
good	22.56	scene	30.04	pretti	48.85
norman	22.38	clint	26.45	humor	45.46
star	22.09	version	26.17	star	45.35

Tab. 4. The top 15 words according to their SVD word importance. Source: The author's own analysis (Statistica software).
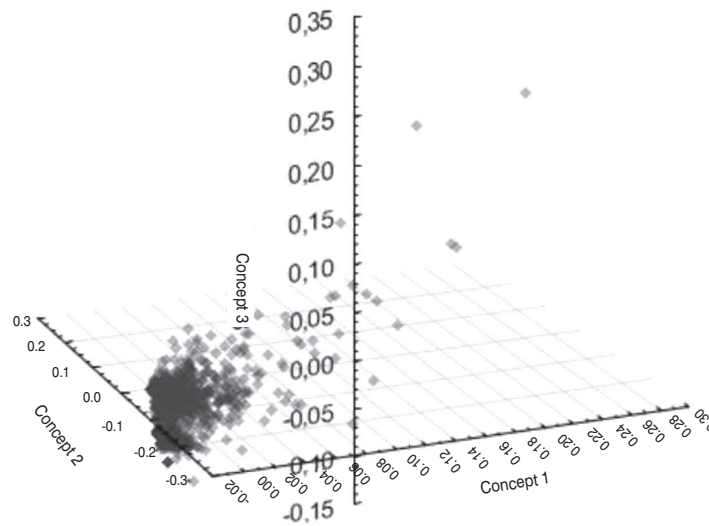

Fig. 1. The distribution of document scores (reviews) for "The Good, the Bad, and the Ugly" (1966) in relation to Concept 1, Concept 2, Concept 3. Source: The author's own analysis – Statistica software.
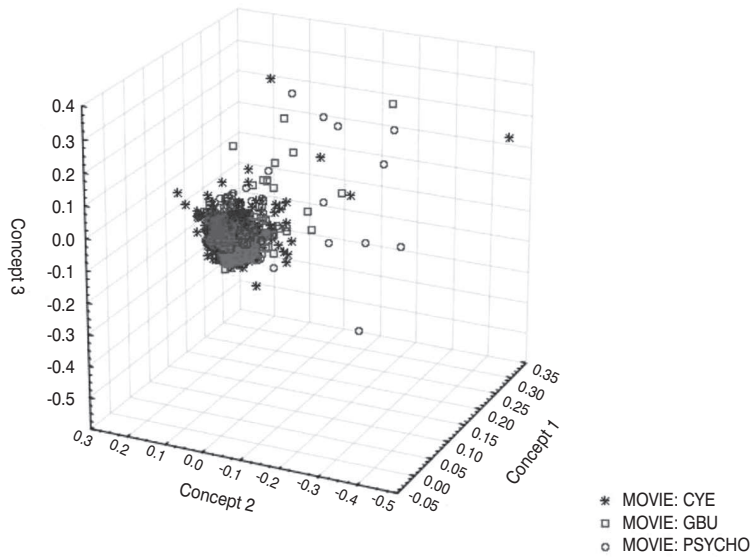
Fig. 2. The distribution of document scores (reviews) for "Psycho" (1960), "The Good, the Bad, and the Ugly" (1966): GBU, "Curb Your Enthusiasm" (2000): CYE in relation to Concept 1, Concept 2, Concept 3 (concepts refer to one movie). Source: The author's own analysis – Statistica software.
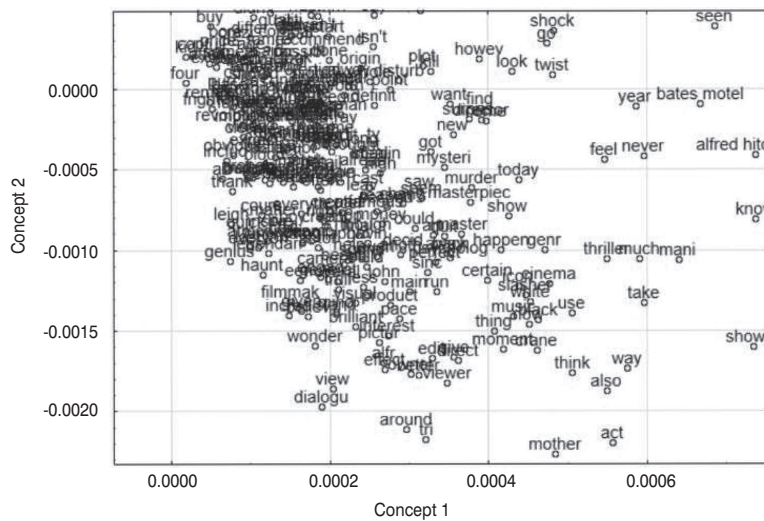

Fig. 3. The distribution of words – SVD word coefficients – for "Psycho" (1960) in relation to Concept 1 and Concept 2 – an excerpt of scatterplot. Source: The author's own analysis – Statistica software.
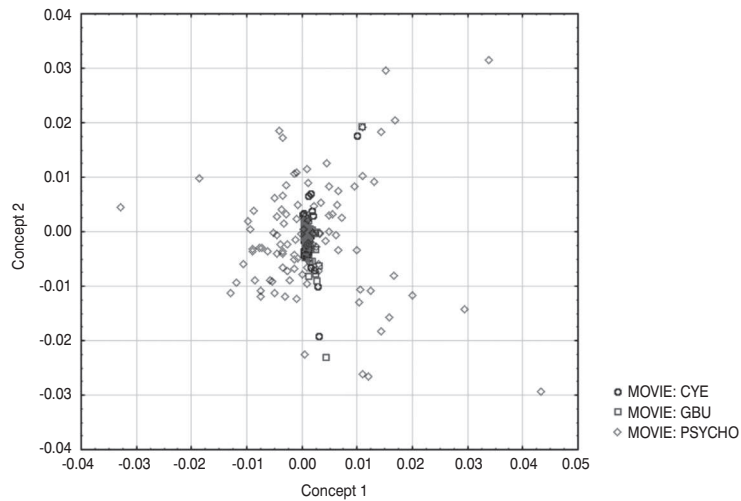
Fig. 4. The distribution of the top 100 words (according to their SVD word importance) − SVD word coefficients for "Psycho" (1960), "The Good, the Bad, and the Ugly" (1966): GBU, "Curb Your Enthusiasm" (2000): CYE − in relation to Concept 1 and Concept 2, and Concept 3. Source: The author's own analysis − Statistica software.
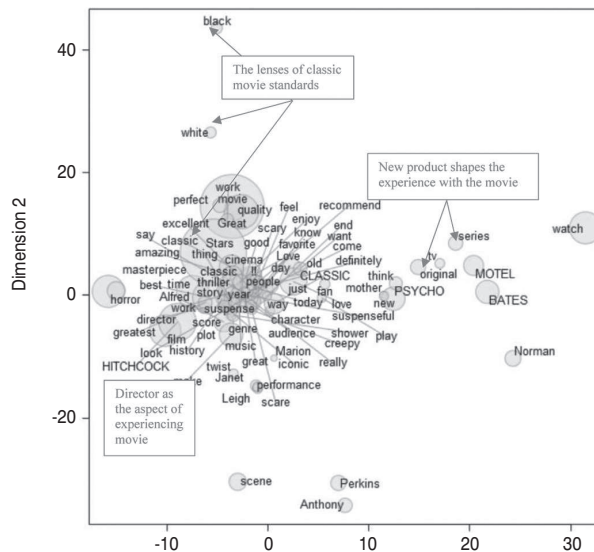

Fig. 5. The selected aspects of viewer experience indicated in a visual analysis of the movie "Psycho" (1960) − Multi-Dimensional Scaling of Words. Source: The author's own analysis − KH Coder software with the following settings: 'Filter words by Term Frequency': 20; 'Method': Sammon; 'Distance': Euclid.
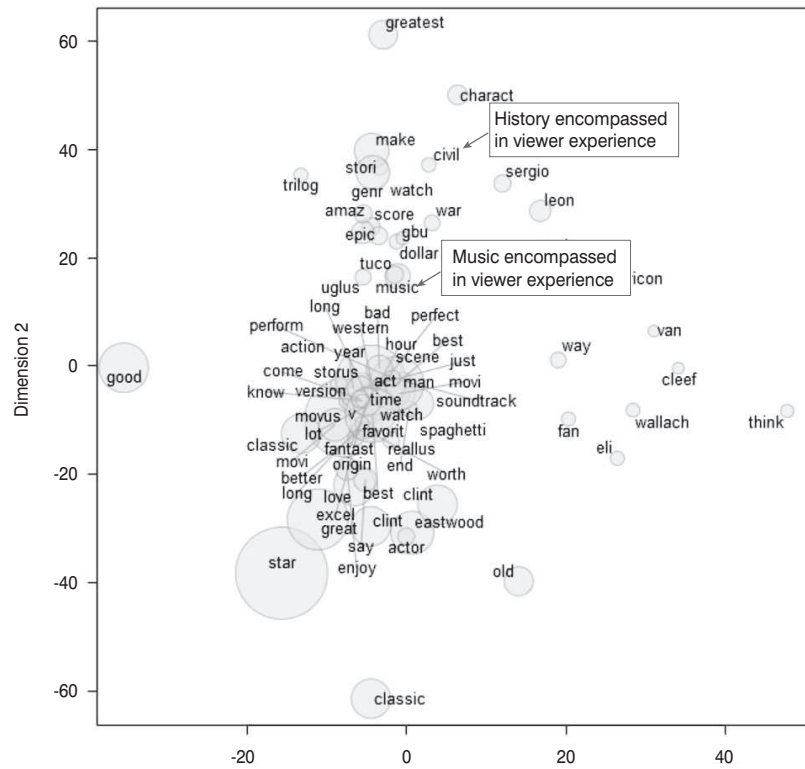
Fig. 6. The selected aspects of viewer experience indicated in a visual analysis of "The Good, The Bad, and The Ugly" (1966) – Multi-Dimensional Scaling of Words. Source: The author's own analysis – KH Coder software with the following settings: 'Filter words by Term Frequency': 20; 'Method': Sammon; 'Distance': Euclid.
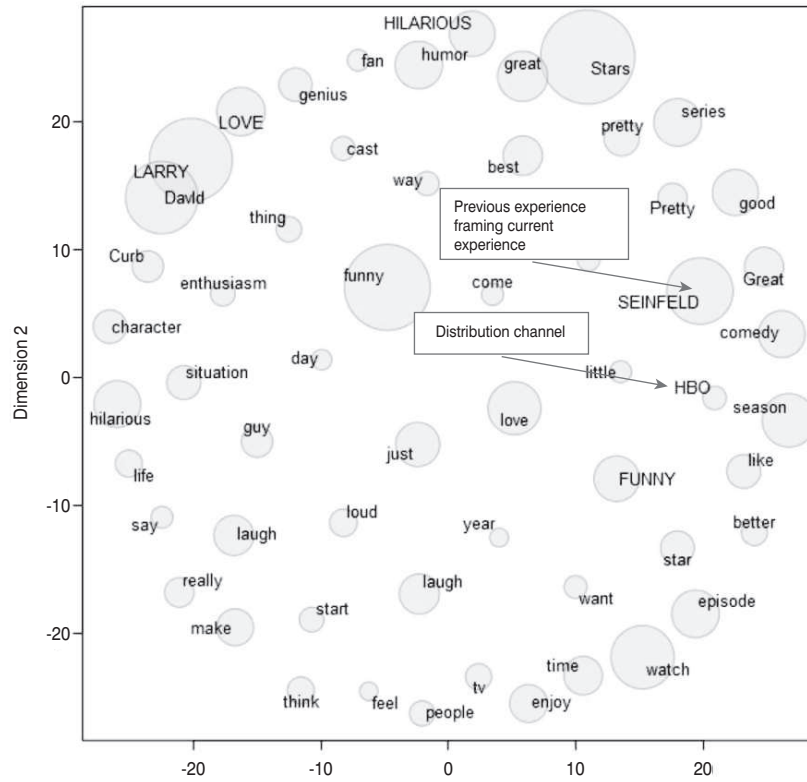
Fig. 7. The selected aspects of viewer experience indicated in a visual analysis of "Curb Your Enthusiasm" (2000) – Multi-Dimensional Scaling of Words. Source: The author's own analysis – KH Coder software with the following settings: 'Filter words by Term Frequency': 20; 'Method': Sammon; 'Distance': Euclid.
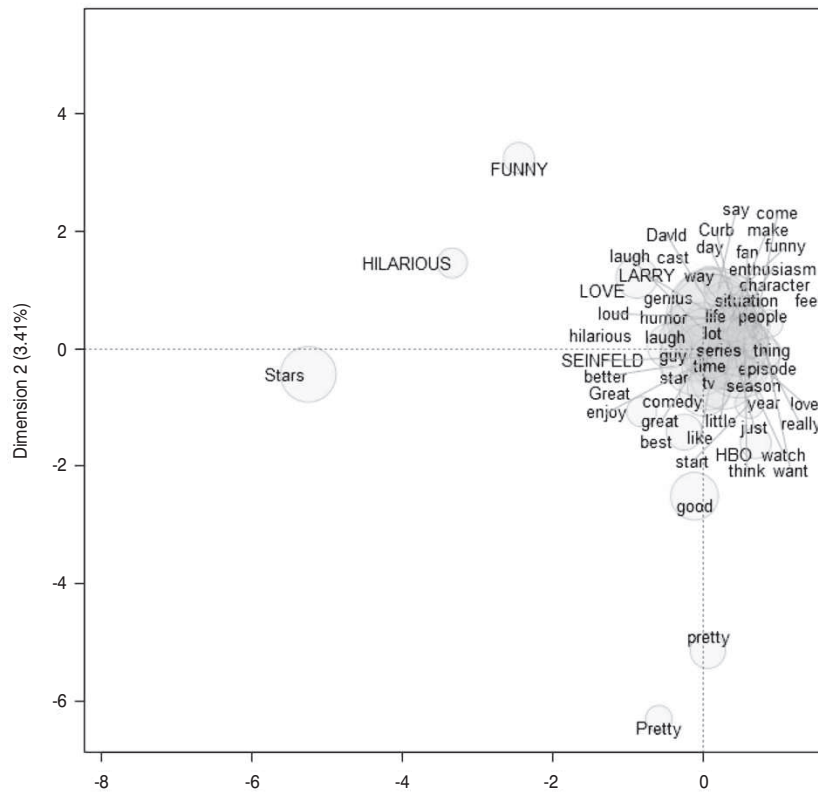
Fig. 8. "Curb Your Enthusiasm" (1966) – Correspondence Analysis of Words (KH Coder software). Source: The author's own analysis – KH Coder software.

The main topic	The sub-topics
The milestone in growth of film industry	Movie sets new standards for genre – thriller, horror, suspense
	'Classic scene' – contributes to the cinema development, e.g. 'classic shower scene'
	Masterpiece in many dimensions – suspense movies, style of directing
Alfred Hitchcock – director as the reference point in the viewer's film experience	Experiencing the film through the lens of other productions, particularly "Vertigo", "Rear Window", "North by Northwest"
	"The best of Hitchcock", the movie contributes to Alfred Hitchcock's legacy
'Classic production'	'Classic' film with 'classic' style of directing
	Unique experiences – different than with contemporary film production – as the movie is a 'black and white' film
The change of perception of 'how scary is this film' over the time'	'The old scary movies are not so scary nowadays', 'the slow plot in comparison with the contemporary pace of films'
	'Still masterpiece' – still frightening, still scary
Nostalgia	Film watched many years ago – the comparison with the viewer's experience of that time
The actors shaping the viewer's experience	The appraisal of acting: Anthony Perkins
Viewers' experience shaped by their liking of the genre	'If you like horror movies'
	'If you like the old classic horror movies'
The comparison the movie with the "Bates Motel" – the prequel of "Psycho", TV series 2013–2017	Original "Psycho" delivers better experience
	New production delivers better experience
	"Bates Motel" complements viewer experience, e.g. enhancing the viewer's plot understanding
'Never ending watching'	Watching the film frequently
Viewers' experience through the 'stars'	Rating available on websites – 'in stars' as the main way to express the viewer's overall experience (e.g. five stars comment)
The technical problems with film viewing	Problems referring to different channels of viewing (damaged DVD, problem with streaming)

Tab. 5. "Psycho" (1960) – a manual analysis of 50 reviews: the main topics and sub-topics. Source: The author's own analysis.

The main topic	The sub-topics
The milestone in growth of film industry	Movie sets new standards for westerns
	'Masterpiece' in many dimensions – acting, music, style of directing
Sergio Leone – the director as the reference point in the viewer's film experience	Experiencing the film through the lens of other productions, such as "For a Few Dollars More" and "Fistful of Dollars"
	The masterpiece of film directing
The actors shaping the viewer's experience	The appraisal of acting: Clint Eastwood, Eli Wallach, Lee Van Cleef
The soundtrack as the factor shaping viewing experience	Appreciation of Ennio Morricone music
The length of movie	'Too long a production' for a contemporary viewer
	'Even though the movie is so long, it is worth watching'
	'Spaghetti western' as the sub-genre
The memorable movie quotes	Enjoyment linked with remembering some movie quotes by heart
'Classic production'	'Classic' film
	'Classic Clint Eastwood'
The change how audience perceives 'the pace of the story'	The 'classic spaghetti-western' – too slow a pace of plot compared to the pace of contemporary films
	'Still a masterpiece'
Nostalgia	Comparison with the viewer's experience of watching the film many years ago, especially with family members
Movie as a gift for member of the family	The purchase of the film as a memorable gift (especially for father, husband, grandson)
Viewers' experience shaped by genre preferences	The movie perception strongly depends 'if you like spaghetti-western movies'
Viewers' experience shaped by the perception of violence portrayed in the movie	'Violence as a part of genre'
	'Unacceptable violence in the movie'
USA history as a 'layer' shaping viewer's experience	Viewing the movie with 'the Civil War in mind'
'Never ending watching'	Watching the film frequently
Viewers' experience through the 'stars'	Rating available on websites – 'in stars' as the main way to express the viewer's overall experience (e.g. five stars comment)
The technical problems with film viewing	Problems related to the different channels of viewing (e.g. DVD, problem with streaming)
Viewers' experience shaped by how the movie was restored using modern techniques	Viewer's attitude to the film version watched

Tab. 6. "The Good, The Bad, and The Ugly" (1966) – a manual analysis of 50 reviews: the main topics and sub-topics. Source: The author's own analysis.

The main topic	The sub-topics
Larry David – the creator and the actor as the reference point in viewer experience	Experiencing the series through the lens of the TV series "Seinfeld"
	Larry David as the main character in the series: the perception of acting, the features of character
The whole season or seasons creates the viewer's experience	The viewer's experience expressed with reference to other seasons or to other episodes within the season
	'Binge-watching' – watching many episodes in one time span
Humor – 'what you consider as funny'	'I do not like/enjoy such humor'
	'I like/enjoy such humor'
	'I do not understand such humor'
	'Perception of the humor through the series' (not as an experience of one episode)
Viewers' perception of the pace of the plot	'I like the plot'
	'Too slow action'
Viewers' experience of watching the episodes in different times and technology	Watching the episodes on HBO in 2000 compared to the contemporary experience
Sharing the viewer's experience with others	Watching with another person (e.g. spouse) and sharing opinions about the movie
Viewers' experience expressed through the 'stars'	Rating available on websites – 'in stars' as the main way to express the viewer's overall experience (e.g. five stars comment)
The technical problems with film viewing	Problems related to the different channels of viewing (DVD, problem with streaming)

Tab. 7. "Curb Your Enthusiasm" (2000) – a manual analysis of 50 reviews: the main topics and sub-topics. Source: The author's own analysis.

6. Limitations of the Study

In comparison with other methods of obtaining data (e.g. interviews with a viewer, eye-tracking), text mining seems to have many advantages such as the relative ease of obtaining a large volume of data and the unobtrusive way of gathering data. On the other hand, text mining of online reviews has its limitation such as:

1. Text mining does not capture the customer experience at the moment of viewing or – more generally – at the moment when the experience is created (contrary to neuroscientific studies or some more 'traditional' surveys, e.g. when a viewer expresses their feelings during the navigation on the website with streaming services).

2. Results represent a 'general picture' of viewers – and, therefore, they do not offer insight into the experience of a particular individual person.

3. It can be said that online reviews are somehow 'fragmented' pieces of viewer experience – as one review often covers only a small fragment of the individual experience.

4. Due to the nature of online reviews, in some cases, the analysis includes false data – the problem of 'fake reviews' which are difficult to extract from the general pool.

Apart from the abovementioned, further limitations of the study presented in this paper are:

– A visual analysis of text mining results is not always easy for interpretation through the lens of viewer experience. Moreover, some aspects of visual analysis can only be captured with specific settings. For example, the differences in the pattern of distribution of words captured in the qualitative analysis of Multi-Dimensional Scaling of Words (Figures 7–9) is noticeable with the selection of a particular method and a measurement of distance – the method: Sammon, with Euclid distance. When different methods of analysis, Kruskal or Classical, and different measurements of distance, Jaccard or Cosine, were applied, the differences are less apparent or negligible.

– Each analysis included different volumes of word-of-mouth generated for a particular movie: the analyses included the different numbers of reviews for each movie, and the reviews also differed in the length.

7. Conclusions

7.1. Broad Spectrum of Viewer Experience Obtained by the Implementation of Text Mining of Online Reviews

This study emphasizes that text mining of online movie reviews reveals a very broad spectrum of factors included in or affecting the viewer experience construct or consumer (viewer) journey. Looking at the results through the process model of customer journey and experience (Lemon & Verhoef, 2016) as the reference model, the following aspects are visible:

– The significance of previous experiences for the viewer journey. These experiences can relate both to features of the current product (e.g. the director, the awards received), features of the product related to current products (e.g. sequels, prequels) or other aspects, for example, distribution channels, social interactions, and the 'standards set for a genre' at a specific time.

– The significance of touch points for the pre-purchase, purchase and post-purchase stages. These touch points include: brand-owned touch points (e.g. movie website, trailer, the features of the movie), partner-owned

touch points (e.g. the cinema, the website of streaming services or the retailer with a DVD offer), customer-owned touch points (particularly the customer's own account on a social media platform where they share their own experience with the movie) and social touch points (e.g. movie online forums).
– The significance of current experience as the baseline for the creation of viewer experience in the future.

The results of this study are consistent with previous studies focusing on consumer experience with movies (Świerczyńska-Kaczor & Kotlińska, 2017; Świerczyńska-Kaczor & Wachowicz, 2017).

7.2. Lack of 'General' Criteria for Differentiating Viewer Experience With Different Genres

Although this study indicates different categories of understanding viewer experience, it is difficult to point to more general criteria which can be used to distinguish viewer experiences created by different genres: a western, suspense or comedy movie (apart from 'obvious' – experience of fear when watching a horror or humor in a comedy).

7.3. Text Mining of Online Reviews as a Useful Background for Further Research

The results of this study prompt the question to what extent text mining can constitute the base for further research – particularly for studies based on interviews. It seems that text mining of online reviews:
– can be helpful in highlighting the aspects which otherwise could be omitted at the stage of designing the questionnaire for interviews. For example, the historical frame – linked to the perception of the Civil War – can be seen as a 'not obvious' element incorporated in viewer experience but captured in this study;
– can be helpful in the selection of 'adequate words' for framing the questions for the interviews. The knowledge about words which are used by viewers in expressing their experience can contribute to a 'better' questionnaire design.

7.4. Viewer Experience Needs to be Examined With an Interdisciplinary Framework

Although a seminal paper of U. Hasson et al. (2008) inspired the design of this study, it seems that the presumption that the results obtained in text mining could somehow be compared with the results from the neuroscientific study turned out to be a far-fetched idea. On the other hand, this study emphasizes the necessity of interdisciplinary research. So far, the research of viewer experience has still been mostly conducted within the borders of a particular discipline.

8. Future Research

An interesting directions of future research – linking text mining and online movie reviews – lie in the following fields:
– Product planning: To what extent could text mining be useful in 'designing the completed movie' when the analysis refers to online comments on the movie trailer?
– Understanding and noticing new consumer trends: How to improve and extend text mining in order to capture a wider spectrum of viewer experience, and especially to be able to notice new trends in consumer behavior?

References

Bezdek, M.A., Gerrig, R.J., Wenzel, W.G., Shin, J., Pirog Revill, K., & Schumacher, E.H. (2015). Neural evidence that suspense narrows attentional focus. *Neuroscience*, *303*, 338–345. https://doi.org/10.1016/j.neuroscience.2015.06.055.

Chan, Y., Hsu, W., Liao, Y., Chen, H., Tu, C., & Wu, C. (2018). Appreciation of different styles of humor: An fMRI study. *Scientific Reports*, *8:15649*, 1–12. https://doi.org/10.1038/s41598-018-33715-1.

Hasson, U., Landesman O., Knappmeyer B., Vallines I., Rubin N., & Heeger D. (2008). Neurocinematics: The neuroscience of film. *Projections*, *2*(1), 1–26. https://doi.org/10.3167/proj.2008.020102.

Kranzbühler, A., Kleijnen, M.H., Morgan, R.E., & Teerling, M. (2018). The multilevel nature of customer experience research: An integrative review and research agenda. *International Journal of Management Reviews*, *20*, 433–456. https://doi.org/10.1111/ijmr.12140.

Lee, J.H., Jung, S.H., & Park, J. (2017). The role of entropy of review text sentiments on online WOM and movie box office sales. *Electronic Commerce Research and Applications*, *22,* 42–52. https://doi.org/10.1016/j.elerap.2017.03.001.

Lehne, M., Engel, P., Rohrmeier, M., Menninghaus, W., Jacobs, A.M., & Koelsch, S. (2015). Reading a suspenseful literary text activates brain areas related to social cognition and predictive inference. *PLoS ONE*, *10*(5): e0124550. https://doi.org/10.1371/journal.pone.0124550.

Lehne, M., & Koelsch, S. (2015). Toward a general psychological model of tension and suspense. *Frontiers in Psychology*, *6*, 1–11. https://doi.org/10.3389/fpsyg.2015.00079.

Lemon, K.N., & Verhoef, P.C. (2016). Understanding customer experience throughout the customer journey. *Journal of Marketing: AMA/MSI Special Issue*, *80*, 69–96. https://doi.org/10.1509/jm.15.0420.

Niraj, R., & Singh, J. (2015). Impact of user-generated and professional critics reviews on Bollywood movie success. *Australasian Marketing Journal*, *23,* 179–187. http://doi.org/10.1016/j.ausmj.2015.02.001.

Rui, H., Liu, Y., & Whinston, A. (2013). Whose and what chatter matters? The effect of tweets on movie sales. *Decision Support Systems*, *55*(4), 863–870. https://doi.org/10.1016/j.dss.2012.12.022.

Samson, A.C., Zysset, S., & Huber, O. (2008). Cognitive humor processing: Different logical mechanisms in nonverbal cartoons—an fMRI study. *Social Neuroscience*, *3*(2), 125–140. https://doi.org/10.1080/17470910701745858.

Singh, V., Saxena, P., Singh, S., & Rajendran, S. (2017, May). Opinion mining and analysis of movie reviews. *Indian Journal of Science and Technology*, *10*(19), 1–6. https://doi.org/10.17485/ijst/2017/v10i19/112756.

Shaziya, H., Kavitha, G., & Zaheer, R. (2015). Text categorization of movie reviews for sentiment analysis. *International Journal of Innovative Research in Science, Engineering and Technology*, *4*(11), 11255–11262. https://doi.org/10.15680/IJIRSET.2015.0411065.

Świerczyńska-Kaczor, U. (2019). New challenges in the research of viewer experience – An overview of fMRI studies in neurocinematics. In A. Majer & T. Szczepański (Eds.), *Kultura produkcji filmowej: teorie, badania, praktyki* (pp. 114–134).

Świerczyńska-Kaczor, U., & Kotlińska, M. (2017). Could the movie be cute? Understanding the user-generated word-of-mouth by implementing text mining analysis on the movie market. In V. Novitzká, Š. Korečko, & A. Szakál (Eds.), *2017 IEEE 14th International Scientific Conference on Informatics, Informatics 2017, Proceedings* (pp. 377–382).

Świerczyńska-Kaczor, U., & Wachowicz, J. (2017). Understanding the film audience –Providing insight into the viewer's experience from text mining and manual text analysis of online film reviews. *Problemy Zarządzania*, *15*(4)*, 177–193.

Thorat, A.M., & Priya, R.V. (2018). Sentiment analysis of movie review using text mining. *International Journal of Pure and Applied Mathematics*, *119*(16), 3561–3566. Retrieved from http://www.acadpubl.eu/hub/.

Vrticka, P., Black, J.M., & Reiss, A.L. (2013). The neural basis of humour processing. *Nature Reviews Neuroscience*, *14*, 860–868. https://doi.org/10.1038/nrn3566.

Watson, K.K., Matthews, B.J., & Allman, J.M. (2007). Brain activation during sight gags and language-dependent humor. *Cerebral Cortex*, *17*(2), 314–324. https://doi.org/10.1093/cercor/bhj149.

Wild, B., Rodden, F.A., Grodd, W., & Ruch, W. (2003). Neural correlates of laughter and humour. *Brain*, *126*(10), 2121–2138. https://doi.org/10.1093/brain/awg226.