

Correspondence analysis of genre preferences in UK film audiences

Nick Redfern,
Independent Researcher

Abstract

I analyze the genre preferences of UK film audiences, applying correspondence analysis to data produced for the BFI's report *Opening Our Eyes* on the cultural contribution of film. Although there is a weak correlation between the interactively coded variable 'gender-age' and genre, I identify clear patterns of genre preferences that extend the original interpretation of this data showing how viewer's preferences for different genres are associated with their gender and their age.

Keywords: Correspondence analysis, British cinema, film audiences, genre,

Introduction

Genre provides viewers with a first reference point for a film, and functions as a 'quasi-search' characteristic through which audiences assess product traits without having seen a particular film (Hennig-Thurau *et al.* 2001). In a market place comprising a large number of unique cultural products with no unambiguous reference brand, audiences form experience-based norms at the aggregate level of genre rather than the specific level of individual films (Desai & Basuroy 2005). Consequently, genre is the means by which the film industry alerts viewers that pleasures similar to those previously enjoyed are available without compromising the need for novel products; and empirical research has shown that genre is an important factor – if not the most important – in audiences' decision making about which film to see (Litman 1983, Da Silva 1998).

Understanding audience preferences for certain types of films is therefore a priority for film producers and distributors as this will be a factor in deciding which films to produce and how to market them effectively. In this short paper we analyze the genre preferences of UK film audiences, applying correspondence analysis to data produced by the British Film Institute's research into the cultural contribution of film in the UK. Specifically, we focus on

how genre preferences vary with gender and age when treated as a single composite variable.

The BFI dataset

In July 2011, the British Film Institute (BFI) published a report, *Opening Our Eyes* (BFI/Northern Alliance/Ipsos Media CT 2011), examining the cultural contribution of film in the UK.¹ This report analysed how audiences consume films and attitudes to the impact of film based on a series of qualitative 'paired depth' interviews and an online survey of 2036 UK adults aged between 15 and 74.

Question C.1 in the questionnaire invited respondents to express preferences for their favourite genres/type of films from a list comprising action/adventure, animation, art house/films with particular artistic value, comedy, comic book movie, classic films, documentary, drama, family film, fantasy, foreign language film, horror, musicals, romance, romantic comedy, science fiction, suspense/thriller, other, none, and don't know. Respondents were able to select as many genres as they wished, and the data represents the number of respondents expressing a preference for that genre. Figure 7 in the final report presents the breakdown of genre preferences by gender, concluding that male audience members exhibit stronger preferences for science fiction, action/adventure, and horror films while women preferred romantic comedies, family films, romances, and musicals.² In an additional detailed summary made available online, genre preferences were broken down by age group. These results showed younger respondents were more likely select comedy, horror, animation, and comic book as their favourite genres, whereas older audience members were more likely to select dramas, documentaries, and classic films.

The report did not present any findings regarding genre preferences based on the combination of the gender and the age of the subjects, and it is this interaction analysed here. In addition to publishing the final report the BFI has made the full set of result tables from the quantitative survey available to researchers freely online. Table 416 of this output contains the data on gender, age, and genre preferences, and is the basis for our correspondence analysis. We use nineteen of the categories listed above, with 'don't knows' excluded from the analysis. Table 416 lists the additional genre categories of westerns, historical, war, and gangster films, and these have been included in the category 'other.'

Correspondence analysis

Correspondence analysis (CA) is a multivariate technique for exploring and describing frequency data defined by two or more categorical variables in a contingency table. By calculating chi-square distances between the row and column profiles in a table, CA determines the (dis)similarity of the reported frequencies. CA aims to reveal the structure inherent in the data, and does not assume an underlying probability distribution. Consequently, CA requires that all of the relevant variables are included in the analysis and that the entries in the data matrix are nonnegative, but makes no other assumptions. CA

does not support hypothesis testing, and cannot be used to determine the statistical significance of relationships between variables. Here we describe the outputs of the correspondence analysis and their interpretation, and the reader can find introductions to the theory and mathematics of CA in Clausen (1998), Beh (2004), and Greenacre (2007).

The first output of the correspondence analysis is a table describing the variation in the contingency table, referred to as the *inertia*. The total inertia in the table is equal to the chi-square statistic divided by the total sample size: $\phi^2 = \chi^2/N$. This variation is decomposed into the principal inertias of a set of dimensions, each accounting for a percentage of the total inertia. For an $r \times c$ table, the maximum number of dimensions is $\min(r-1, c-1)$. The number of dimensions retained for analysis is based on the first k dimensions to cumulatively exceed a threshold (typically 80 or 90 per cent of the total inertia), all those individual dimensions accounting for more than $1/(\min[r, c] - 1)\%$ of the total inertia, or by reference to a scree plot of the inertias to determine where the drop in the percentage accounted for by a dimension drops away less rapidly. It is also dependent on our ability to give a meaningful interpretation to the dimensions selected. In selecting only a subset of the available we lose some of the information contained in the original table, but in discarding some dimensions we are able to see structure of the data more clearly for as little cost as possible.

As a form of geometric data analysis, correspondence analysis enables the information in a contingency table to be represented as clouds of points in low-dimensional graphical displays (see Le Roux & Rouanet 2005, Greenacre 2010: 79-88). The origin of the graph represents the average row (column) profile, and by assessing the distance of points from the centroid of the clouds we describe the variation within the table and their similarity. Row (column) points that lie close to the origin are similar to the average profile of the row (columns). Data points that lie far from the origin indicate categories for which the observed counts differ from the expected values under independence and account for a larger portion of the inertia. Points from the same data set lying close together represent rows (columns) that have similar profiles, and data points that are distant from one another indicate that the rows (columns) are remote. The distance between row points and column points cannot be interpreted as meaningful as they do not represent a defined quantity. The angle (θ) subtended at the origin defines the association between row and column points: when the angle is acute ($\theta < 90^\circ$) points are interpreted as positively correlated, points are negatively correlated if the angle between them is obtuse ($\theta > 90^\circ$), and points that subtend a right angle ($\theta = 90^\circ$) are not associated (Pusha *et al.* 2009).

In addition to the graphical displays, a detailed numerical summary of the correspondence analysis is produced. The *mass* of a row (column) indicates the proportion accounted for by that category with respect to all the rows (columns), and is simply the row (column) total of divided by the total sample size; while the *inertia* of a data point is its contribution to the overall inertia. The *squared correlation* describes that part of the variation of a data point explained by a particular dimension. The *quality* of a data point measures how well it is represented by the graph, and is equal to the sum of the squared

correlations of the dimensions retained for the analysis. The higher the quality of a data point the better the extracted dimensions represent it, and ranges from 0 (completely unrepresentative) and 1 (perfectly represented). The *absolute contribution* of a data point describes the proportion of the inertia of each dimension it explains, and is determined by both the mass of the data point and its distance from the centroid.

Gender, age, and genre preferences

Table 416 of the BFI’s results output presents counts of genre preferences sorted by gender, by age, and by gender and age. As our interest lies in the variation of genre preferences (19 categories) among UK audiences based on both gender and age we use only this last part of the table, treating ‘gender-age’ as an interactively coded variable with 10 categories combining all the levels of the variables gender (2 categories) and age (5 categories) (Greenacre 2007: 121-128). We apply correspondence analysis to this table using the ca package (version 0.33; see Nenadić & Greenacre 2007) in R (version 2.13.0).

Table 1 presents the 10 × 19 cross-tabulation of ‘gender-age’ with genre. The chi-square statistic for this table is 1312.28 (N = 13086, df = 162, $p < 0.01$), and we therefore conclude that there is a statistically significant association between gender-age and genre preferences for UK film audiences. However, there is only a weak correlation between ‘gender-age’ and genre preference, with just 10% of the variation in **Table 1** due to dependence: $\phi^2 = \chi^2/N = 1312.28/13086 = 0.1003$.

Table 1 Cross-tabulation of interactively-coded gender-age variable with genre. Cell counts represent the number of respondents in each group expressing a preference for a genre. Source: BFI/Northern Alliance/Ipsos Media CT:

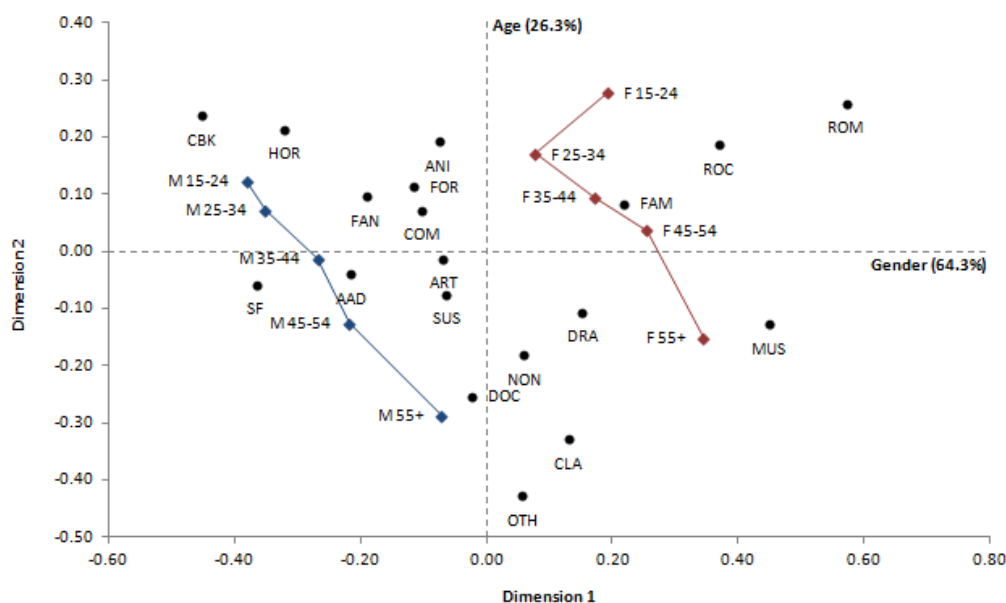
	Action/ Adventure	Animation	Classic	Comedy	Comic Book	Document ary	Drama	Family	Fantasy	Horror	Musicals	Romance	Romantic Comedy	Science Fiction	Suspense/ Thriller	Foreign	Art	Other	None
Male 15-24	133	69	36	142	55	49	67	49	72	71	18	14	32	92	105	23	16	2	3
Male 25-34	127	73	47	141	56	83	78	47	90	78	22	20	39	112	111	39	27	3	2
Male 35-44	129	66	78	138	50	73	95	66	76	63	23	21	44	116	110	36	36	4	8
Male 45-54	130	44	84	121	32	87	101	48	72	63	28	23	49	111	123	20	30	9	9
Male 55+	171	50	138	171	22	149	162	71	75	48	73	24	59	109	166	25	30	12	16
Female 15-24	85	58	29	134	22	40	88	68	56	63	53	97	114	38	84	22	20	2	9
Female 25-34	101	68	45	134	37	54	106	87	70	51	51	83	94	63	97	30	25	7	4
Female 35-44	88	82	86	139	33	78	130	115	78	48	69	94	113	63	105	35	33	1	12
Female 45-54	95	63	86	122	17	71	134	106	57	34	68	97	121	60	117	19	22	5	4
Female 55+	122	69	174	150	18	141	220	131	63	31	151	137	141	64	163	36	35	12	9

Table 2 shows the principal inertias, percentages, and cumulative percentage of each dimension, with a scree plot of the inertias. The first two dimensions account for 90.6 per cent of the inertia and the scree plot flattens out after the second dimension. Consequently, these dimensions were retained for analysis and the remainder were discarded. Figure 1 is the resulting symmetric map based on these two dimensions.

Table 2 Principal inertias of the correspondence analysis applied to **Table 1** explained by dimensions with scree plot:

Dimension	Principal Inertia	Per cent	Cumulative	Scree Plot
1	0.0645	64.3	64.3	*****
2	0.0264	26.3	90.6	*****
3	0.0036	3.6	94.2	*
4	0.0020	2.0	96.2	*
5	0.0015	1.5	97.6	*
6	0.0012	1.2	98.9	
7	0.0005	0.5	99.4	
8	0.0005	0.5	99.9	
9	0.0001	0.1	100.0	
Total	0.1003	100.0	100.0	

Figure 1 Symmetric correspondence analysis map of interactively coded 'gender-age' cross-tabulated with genre for UK film audiences:



Tables 3a and **3b** present the detailed numerical summary of the results for the rows (gender-age categories) and columns (genre categories), respectively.

Table 3a Detailed numerical summary of correspondence analysis by gender-age:

Name	Code	Mass	Quality	Inertia	Dimension 1			Dimension 2		
					Coordinates	Squared Correlation	Absolute Contribution	Coordinates	Squared Correlation	Absolute Contribution
Male 15-24	M 15-24	0.08	0.93	0.14	-0.38	0.84	0.18	0.12	0.09	0.05
Male 25-34	M 25-34	0.09	0.93	0.13	-0.35	0.90	0.18	0.07	0.04	0.02
Male 35-44	M 35-44	0.09	0.88	0.08	-0.27	0.88	0.10	-0.01	0.00	0.00
Male 45-54	M 45-54	0.09	0.85	0.07	-0.22	0.64	0.07	-0.13	0.22	0.06
Male 55+	M 55+	0.12	0.93	0.11	-0.07	0.05	0.01	-0.29	0.88	0.37
Female 15-24	F 15-24	0.08	0.86	0.11	0.19	0.28	0.05	0.28	0.58	0.24
Female 25-34	F 25-34	0.09	0.88	0.04	0.08	0.15	0.01	0.17	0.73	0.10
Female 35-44	F 35-44	0.11	0.76	0.05	0.17	0.58	0.05	0.09	0.18	0.04
Female 45-54	F 45-54	0.10	0.89	0.07	0.25	0.87	0.10	0.04	0.02	0.01
Female 55+	F 55+	0.14	0.97	0.21	0.35	0.81	0.26	-0.15	0.16	0.12

Table 3b Detailed numerical summary of correspondence analysis by genre:

Name	Code	Mass	Quality	Inertia	Dimension 1			Dimension 2		
					Coordinates	Squared Correlation	Absolute Contribution	Coordinates	Squared Correlation	Absolute Contribution
Action/Adventure	AAD	0.09	0.88	0.05	-0.22	0.85	0.07	-0.04	0.03	0.01
Animated	ANI	0.05	0.84	0.03	-0.08	0.11	0.00	0.19	0.73	0.07
Classic	CLA	0.06	0.95	0.08	0.13	0.13	0.02	-0.33	0.82	0.25
Comedy	COM	0.11	0.76	0.02	-0.10	0.51	0.02	0.07	0.25	0.02
Comic Book	CBK	0.03	0.92	0.07	-0.45	0.72	0.08	0.24	0.20	0.06
Documentary	DOC	0.06	0.91	0.04	-0.02	0.01	0.00	-0.25	0.91	0.15
Drama	DRA	0.09	0.99	0.03	0.15	0.67	0.03	-0.11	0.33	0.04
Family	FAM	0.06	0.80	0.04	0.22	0.70	0.04	0.08	0.10	0.02
Fantasy	FAN	0.05	0.93	0.03	-0.19	0.75	0.03	0.10	0.19	0.02
Horror	HOR	0.04	0.92	0.07	-0.32	0.64	0.07	0.21	0.28	0.07
Musical	MUS	0.04	0.94	0.10	0.45	0.86	0.13	-0.13	0.07	0.03
Romance	ROM	0.05	0.99	0.18	0.57	0.82	0.24	0.26	0.17	0.12
Romantic Comedy	ROC	0.06	0.96	0.11	0.37	0.77	0.13	0.19	0.19	0.08
Science Fiction	SF	0.06	0.96	0.09	-0.37	0.93	0.13	-0.06	0.02	0.01
Suspense/Thriller	SUS	0.09	0.77	0.01	-0.06	0.32	0.01	-0.08	0.45	0.02
Foreign	FOR	0.02	0.41	0.01	-0.12	0.21	0.01	0.11	0.20	0.01
Art House	ART	0.02	0.14	0.01	-0.07	0.14	0.00	-0.01	0.01	0.00
Other	OTH	0.00	0.60	0.01	0.06	0.01	0.00	-0.43	0.59	0.03
None	NON	0.01	0.16	0.01	0.06	0.02	0.00	-0.18	0.14	0.01

From **Table 3a** and **Figure 1** we see a clear horizontal separation between the male and female respondents, with points arranged vertically by age group from youngest to oldest within each gender category. Consequently, we interpret the principal axes in terms of the rows of **Table 1**, with the first dimension understood as gender and the second dimension as age. As gender accounts for 64.3 per cent of the total inertia compared to 26.3 per cent for age, this factor is dominant and explains the major part of the variation in **Table 1**. The

quality for the gender-age groups is high (see **Table 3a**), and these factors are well represented in two dimensions. The points for all gender-age groups are distant from the origin, indicating that no group is close to the average profile in either dimension and that all the groups contribute to the overall inertia.

From **Figure 1** we see the distance between the points representing male audience members greater as the age of the respondents increases. The points for males aged 15-24 and 25-34 are very close indicating they have similar row profiles and, therefore, similar genre preferences. The two middle-aged groups are distant from both the youngest and the oldest, while also being remote from one another. Males over the age of 55 are remote from the other age groups, indicating that their genre preferences are substantially different from those of younger male audience members. The points representing female respondents show a similar pattern with the middle aged groups distant from both youngest and oldest and with over-55s are remote from younger female audience members in their preferences. The greatest contrasts in genre preferences are observed when taking gender and age together: females over 55 are most different from males aged 15-24, and males aged 55+ are most different from young women.

A key difference between audience groups is how the importance of the factors of gender and age vary in explaining their genre preferences. Age becomes increasingly important in the representation of the points for male audience categories. The squared correlations for the three youngest male groups are greatest for dimension 1, indicating that their gender is more important in explaining their preferences than age; for males aged 45-54 gender is still the dominant component albeit to a lesser extent than younger cohorts and the influence of age becomes more apparent in the raised squared correlation for dimension 2; while for males aged 55+ age is the dominant factor. This pattern is not evident for female respondents, and looking at the squared correlations in **Table 3a** we see the opposite pattern to male audience members. The squared correlations for women aged 35-44, 45-54, and 55+ are dominated by the dimension of gender, whereas age is the main factor for the two youngest groups. However, it should be noted that for the females aged 15-24, gender does contribute substantially to the representation of this point.

Although the correlation between gender-age and genre preference is low, it is clear from these results that the variation within **Table 1** is highly structured in terms of the gender and age of the respondents. Describing the preferences of UK cinemagoers therefore requires taking *both* these factors into account and failure to do so leads to much useful information being obscured. The headline percentages reported by the BFI give only a partial picture of the genre preference of UK film audiences that fails to adequately capture that structure.

Turning to the genre categories themselves we see that the quality of these points is high (see **Table 3b**), indicating they are well represented in two dimensions and that gender and age are good predictors of the genre preferences of UK audiences. However, we note the quality of the representation for foreign (0.41) and art-house (0.14) films by these two dimensions is very low. This indicates gender and age do not explain variation in audience

preferences for these types of films, and that some other factor should be considered. Based on other data available in the BFI's results output, level of educational attainment is a better predictor of audience preference for these types of films: Table 20 of the results output cross-tabulates level of education and type of film most often watched, with 68 per cent of respondents selecting foreign language films educated to degree level. These two categories are typically applied to films to distinguish them from mainstream cinema (i.e. Hollywood films), and may not function as genre labels in the same context as terms such as 'comedy,' 'drama,' etc.

The quality of the categories 'other' and 'none' are also much lower than the mainstream genres, but as these points represent indistinct categories we do not discuss them further.

Gender is the most important factor in determining genre preference, with the cloud of points representing genres orientated along the first principal axis. Family films, romance, and romantic comedies are all associated with female audiences. In fact, 83 per cent of respondents to express a preference for romance films were female, and the corresponding figures are also high for family films (64%) and romantic comedies (72%). Musicals are also strongly associated with female audiences (71%), but this category is dominated by over 55s: over a quarter of respondents expressing a preference for this genre are in this age group. Drama also lies along the same direction as females over 55 indicating that this group is associated with this genre, but the distance from the origin is smaller reflecting a smaller effect. The proportion of males over 55 selecting drama films as a preferred genre is also greater than younger male viewers, but not to the same extent as their female counterparts. In fact, female viewers in each age group expressed a stronger preference for drama films than male viewers of the same age.

Genres associated with male audiences tend to be action-based and technology-driven. Of respondents expressing a preference for science fiction films, 65 per cent were male and there is little variation between age groups within this gender category. Consequently, this genre is very well represented by the first principal axis and age is not a significant factor. This is also the case for action/adventure films (58%), albeit it to a lesser degree as this point lies nearer the origin. Comic book, fantasy, and horror films are strongly correlated with male audiences, and lie along the same direction as males aged 15-24 and 25-34 indicating that age also a key factor here. The squared correlations for gender are the dominant factors for these genres, but age also contributes a substantial part of these points' representation.

It is interesting that genres we associate with male audiences appear to have broader appeal than genres we associate with female audiences. Dividing the cells by the column totals to give the proportion of respondents in each gender-age group expressing a preference for a genre, we see that no male age group accounts for more 4 per cent of the total for romance films compared to the very large proportion for female audiences noted above. Although female associated, family films do not show the extreme divide as romance films, romantic comedies, and musicals. For science fiction films, the female respondents

account for a total of 35 per cent of the expressed preferences for this genre, with each age group within this gender category contributing between 5 and 8 per cent of the total. This is also the case for comic book and action/adventure films. We conclude that so-called 'female genres' hold very little appeal to male audiences; and that while similar patterns are certainly evident for 'male genres' the effect is much smaller.

Three genres show high squared correlations with age. In all the cases the contribution of the first principal axis is small, and we conclude that gender is relatively unimportant in explaining audience preferences for these films. Animation is associated with under-35s, though female viewers aged 35-44 account 13 per cent of the column total in **Table 1** possibly due to selecting these films for family viewing. Documentaries and classic films are associated with over-55s. Of those expressing a preference for documentaries, 18 per cent were males over 55 and 17 per cent were females in the same age group. There is no specific trend among the other age groups, which show roughly equal levels of interest in these films. It is noticeable that proportion selecting classic films increases with age, though this may reflect the aging of the audience rather than a clear genre preference as the new films of one's youth become classics with time.

Two genres – comedy and suspense/thriller – lie near the origin. These points also have the lowest quality of the mainstream genres, though both are still well represented in Figure 1. Both dimensions contribute to the representation of these points, indicating that gender and age are relevant factors. Gender makes a larger contribution to comedy than age, with males under 35 slightly more likely to express a preference for this genre than males over 35 or female viewers; while for suspense/thrillers over 55s of both genders account for slightly greater proportion of the preferences expressed for this category. However, it is their closeness to the average profile that is most informative about these points, indicating that all gender-age groups enjoy these types of films. This does not mean that they are watching the same films *within* these genres – it is very unlikely males aged 15-24 are watching the same comedy films, for example, as women over 55; but the BFI's data cannot help us to explore this aspect.

Conclusion

This study analyzed the genre preferences of British film audiences. We have replicated the results originally presented by the BFI, and have extended them to reveal additional patterns in the data. Correspondence analysis enables us to obtain an overview of how different sections of the audience for films in the UK relate to one another, and to assess the relative importance of different factors in explaining the variation among audiences and their genre preferences. The study showed that gender is the dominant factor in determining audience preferences, with age an important but secondary factor. Most genres can be identified as either 'male' or 'female' with clear age profiles evident within gender categories, though preferences for animated films, classic movies, and documentaries are determined by age alone. These factors do not adequately explain

variation among audiences when applied to categories of films that lie outside mainstream cinema.

Biographical note:

Nick Redfern studied Film and History at the University of Kent, and completed his Ph.D. on regionalism in contemporary British cinema at Manchester Metropolitan University in 2006. He has published articles on regional film policy in the United Kingdom and the representation of the UK's regions in cinema in *The Journal of British Cinema and Television*, *Cyfrwng: Media Wales Journal*, *The Journal of European Popular Culture*, and *Transnational Cinemas*. He has taught film studies at Manchester Metropolitan University and the University of Central Lancashire, specializing in British cinema, Hollywood cinema, film analysis, and film theory. He is the author of the blog *Research into Film*. Contact: nickredfern@hotmail.com.

References

- Beh, Eric J., 'Simple correspondence analysis: a bibliographic review', *International Statistical Review*, 72 (2), 2004, pp. 257-284.
- BFI/Northern Alliance/Ipsos Media CT, *Opening Our Eyes: How Film Contributes to the Culture of the UK*, July 2011.
- Clausen, Sten-Erik, *Applied Correspondence Analysis: An Introduction*. Thousand Oaks, CA : Sage, 2008.
- Da Silva, Indra, 'Consumer selection of motion pictures', in Barry R. Litman (ed.), *The Motion Picture Mega-industry*. Boston : Allen and Bacon, 1998, pp. 144-171.
- Desai, Kalpesh Kausik, and Basuroy, Suman, 'Interactive influence of genre familiarity, star power, and critics' reviews in the cultural goods industry: the case of motion pictures', *Psychology and Marketing*, 22 (3), 2005, pp. 203-223.
- Greenacre, Michael, *Correspondence Analysis in Practice*, second edition. Boca Raton, FL: Chapman & Hall/CRC, 2007.
- Greenacre, Michael, *Biplots in Practice*. Bilbao: Fundación BBVA, 2010.
- Hennig-Thurau, Thorsten, Walsh, Gianfranco, and Wruck, Oliver, 'An investigation into the factors determining the success of service innovations: the case of motion pictures', *Academy of Marketing Science Review*, 6, 2001, [WWW document] URL <http://www.amsreview.org/articles/henning06-2001/pdf>, [visited 24 May 2011].
- Le Roux, Brigitte, and Rouanet, Henry, *Geometric Data Analysis: From Correspondence Analysis to Structural Data Analysis*. Dordrecht : Kluwer Academic Publishers, 2005.
- Litman, Barry R., 'Predicting success of theatrical movies: an empirical study', *Journal of Popular Culture*, 16 (4), 1983, pp. 159-175.
- Nenadić, Oleg, and Greenacre, Michael, 'Correspondence analysis in R, with two- and three-dimensional graphics: the ca package', *Journal of Statistical Software*, 20 (3), 2007, [WWW document] URL <http://www.jstatsoft.org/v20/i03/paper>, [visited 6 September 2011].
- Pusha Sonia, Gudi, Ravindra, and Noronha, Santosh, 'Polar classification with correspondence analysis for fault isolation', *Journal of Process Control*, 19 (4), 2009, pp. 656-663.

Notes:

¹ The report, the research questionnaire, the detailed summary, and the full set of result tables are available at: <http://www.bfi.org.uk/publications/openingoureyes/>, accessed 21 November, 2011.

² The report also presents results based on respondents' ethnic minority but these will not be discussed here.