

## **7<sup>th</sup> International Conference on Arts and Cultural Management**

A Possible Method of Quality Assessment of Film and Theatre Presentations in London

Richard Reavill  
Cass City of London Business School  
Bunhill Row  
London EC1

### **Synopsis**

The ultimate objective of this research is to determine the relationship between the quality of an arts product, (for example a film or a theatrical production), and its commercial success.

Firstly, it is necessary to develop a measure of product quality, and this is the subject of this paper. The Delphi Technique of evaluation using expert consultation provides the methodological basis for the research, and the arts critics of major London newspapers constitute the group of expert consultants to generate the quality assessment. The film and theatre offerings reviewed in the year 2000 are used to create a data-set of some 7600 items. Two published compilations of assessments of reviews provide the core of the analysis.

Initial consideration of the data indicates that the offerings cover the complete spectrum of quality. There is evidence of both high diversity and high consistency of critical views. The London film critics show a balanced overall view, the mean analysis being 5% positive. A similar analysis for the theatre critics suggests that on balance they take a much more generous view, 40% positive. Examination of the reasons for this significant difference could provide the basis of further research.

A detailed statistical analysis of the data-set is in progress to assess the consistency of individual critics, the possible subjectivity of the compiler who “reviews the reviews”, and any other interesting information that can be derived by data mining. Assessment of reviews of Lyric Theatre productions (opera and dance) has also been undertaken, though there are fewer productions and less data is available in this category. Also, the repertory nature and mainly subsidised origin of most of the output will limit its relevance for later assessments of “commercial” success.

Preliminary analysis suggests that the method adopted can give an acceptable measure of quality.

## Introduction

The author is an academic at Cass City of London Business School, but has a long established interest in the arts which has extended to research on the management of the lyric arts, (Gilhespy & Reavill, 2001; Reavill 1999, 2000, 2002) and dance criticism (Reavill, 1998). The research presented in this paper has as its ultimate objective the investigation of the relationship between the quality of an arts product, such as a film, a theatrical production (play, musical, etc.), an opera or a dance production, and its commercial success. For example, will a highly regarded film make a profit for its producer? Will the multi-million dollar Hollywood blockbuster movie repay its high costs, even if its quality attracts large audiences? Will a superior play attract a larger audience at a London theatre? Will a play that has been positively reviewed by the critics run longer than one that has been panned? Will a high quality musical at a London theatre achieve full houses over a period of many years and thereby repay the substantial investment required to produce it and ultimately generate a profit? Will a ballet or opera that has been praised by critics achieve a larger audience than one that has been abused? Though some products of the arts are subsidised, such as most opera and dance productions, the majority of products in cinema and the performing arts are commercially generated, and must conform to the rules of the market.

Consideration of activity outside the field of the Arts suggests that there is a strong relationship between the quality of a product or service, and its commercial success. This is well established, particularly if “quality” is defined as “fitness for the purpose”, as recommended by the quality guru, Philip Crosby (Crosby, 1979), and the customer assessment is dominant.

It is found generally in the commercial world of product manufacture and service supply that a higher quality product has a greater likelihood of commercial success. First impressions might suggest that this could apply in the Arts world. Further more detailed examination suggests that this assumption is either too crude, or even inaccurate, and that the real situation is more complex. Thus some focussed research has commenced to determine the real relationship between the quality of an arts product, and its commercial or quasi-commercial success.

To achieve this objective, two matters require resolution. Firstly, is a reliable assessment of the quality of an arts product, a film, a play, a musical, an opera, or a ballet, possible? The quality of artefacts such as washing machines and electronic calculators can be assessed against quantifiable (“hard”) criteria. Scientifically based test procedures can be devised to assess the quality of the artefact in quantitative terms. Tolerances can be measured for mechanical components, reliability tests can be established for electronic equipment. The quality of a service can be assessed by how it is perceived by its recipient, the customer. Techniques such as questionnaire surveys are helpful to establish the view of the customer, and convert this to a quasi-quantitative rating. The customer decides whether the service is “fit for his (or her) purpose”, and though possibly prejudiced and irrational, this subjective opinion is generally all that matters.

What methods can be used to assess the more subjective matter of whether a film or play is “good”, “bad” or moderate? This is a matter that is much more difficult, perhaps even impossible, to quantify. It involves complex multiple criteria, high potential diversity of opinion, and personal taste. It is therefore a “soft” problem. One of the ancillary purposes of the research presented in this paper was to examine whether this consideration of “softer” (less easily quantifiable) aspects of quality could be taken a stage further into the area of some products of the “arts”. The particular products selected for consideration were cinema products, (films, movies); theatre productions (generally plays and musicals presented at London theatres); operas; and ballet and dance performances. The reason for this particular eclectic selection will become apparent later in this paper.

Secondly, can a reliable indicator be found for the commercial performance of an arts product? For example, are the receipts at the cinema box-office a true measure of the commercial success of a film? Certainly this represents a measure of customer support, but the film may have cost £2 million or £20 million. Can the length of run of a theatre production serve as a good measure, if costs of production and the capacity and rents of the theatres vary widely? A financial account of the film or theatre project, a profit and loss account, would provide the ultimate test, but would this data be readily available? A subsidised opera or ballet can play to moderate or full houses, but a capacity attendance for a popular work may occur, even with a poorly regarded production. The management of the opera or ballet company will have built their professional assessment of the popularity of the piece into the number of performances scheduled into a repertory season, which will further distort a possible assessment.

Putting aside for the moment the problems of obtaining a rigorous assessment of commercial performance, this paper looks primarily at the quality assessment issue. One way to make a realistic quality assessment of a film or theatre production would be to survey the views of customers by means for an appropriately worded questionnaire survey. Another method would be to canvass the view of experts. It would be interesting to do both and compare the results, but a customer survey would involve a massive expenditure of resources to obtain a sufficiently comprehensive and meaningful set of results. Therefore the use of the opinions of experts is the method adopted in this work.

### Methodology

One established method for the assessment of complex issues by the use of expert opinion is the Delphi Technique, and it is this provides the methodology for this investigation. There is some justification, other than the expediency of having available expert opinion, for the use of this method. The Delphi Technique is a method used to assist the solution of complex problems, frequently of a technical or specialist nature (Van Grundy, 1988). The technique was developed in the 1950s by the RAND Corporation to harness the opinion of a range of experts on the potential damage from atomic bomb attacks. That was a very specialised investigation, but the technique has been used more generally for technological and economic forecasting, and in other areas where the views of knowledgeable individuals were essential to gaining understanding of a particular specialist and complex matter. The mechanics of the technique require the

identification of a group of up to 25 people expert on the subject. A questionnaire is prepared concerning the substantive issues. This is sent to the experts, and their responses analysed. On the basis of this information a second questionnaire is prepared and the process repeated. Further repetitions may ensue. Ultimately the information generated is assessed against the requirements of the exercise.

The technique draws its name from the practice in Ancient Greece of consulting the oracle at Delphi when an important enterprise was being considered. The modern Delphi Technique is a more complex version of the original method. The ancient method involved the ritual slaughter of a sacrificial animal, and the prognosis on the wisdom of the proposed action would be determined following an inspection of its entrails. The modern version of this procedure is hopefully a little more rational.

### Use of Delphi Technique for Data Gathering

For the purposes of applying the Delphi technique to the problem of assessing the quality of films, plays, musicals, operas, ballets and dance productions, there is an established body of experts readily available: the critics who write for the major London newspapers. UK national newspapers are printed for circulation in London and elsewhere, and one evening newspaper is published in London. The newspapers considered in this research are: Evening Standard; Daily Mail; The Times; Daily Telegraph; The Guardian; The Independent; Financial Times; Daily Express; Sunday Times; Sunday Telegraph; the Observer; and The Independent on Sunday. All these newspapers publish reviews of most new films, plays, musicals, ballets, operas and dance performances that open in the capital during the year. The film, theatre, opera and dance critics of the major London newspapers are independent, self-selecting, and only informally recruited for this exercise. They are unpaid for this study since their opinions are in the public domain. They provide a permanent Delphi consultants group, which is already established to generate information about the perceived quality of arts products.

Perhaps this is more precisely a double Delphi system, since the oracular output has interpretations in duplicate. Also recruited unwittingly to assist in the collation of the data are the two arts journalists of the Evening Standard and the Sunday Telegraph who prepare and publish independent tabular compilations of the views of the various newspaper critics. The Evening Standard produces a weekly compilation entitled "What the Critics Said", normally collating the views of most of the critics of the twelve newspapers mentioned earlier. The Sunday Telegraph publishes a similar compilation, entitled "Culture Vulture", which is limited to the views expressed in the eight dailies.

Both compilations consider the categories of Cinema (so termed in the "Culture Vulture" compilation, but termed "Film" in the "What the Critics Said" compilation); Theatre (which includes Musicals as well as plays); Opera; and Ballet (so termed in "Culture Vulture", but called "Dance" in "What the Critics Said"). The Ballet/Dance productions cover the spectrum from classical ballet to mainstream dance, but rarely consider modern experimental works. The paper considers the four categories, but as the Ballet and Opera categories have fewer productions, also considers the joint category of "Lyric Theatre".

This database of published material was assessed for the full year 2000. The year 2000 was chosen because the investigation started in mid 2001, the scale of the available data was significant, and the availability of published economic data about the film and theatre productions was anticipated in late 2002 at the earliest. From the reviews published during the full year 2000, a data set of some 7600 items has been prepared. The two independent compilations use a simple three-element Likert scale, as this allows a presentation of greatest visual immediacy.

### London Newspapers

Most newspapers published in London are national newspapers, printed for circulation in England and elsewhere. The national newspapers generally publish a London edition, usually the last edition printed as it is required for local distribution. Many such newspapers have Arts Editors, Arts Sections or Pages, and retain the services of specialist critics who review new films, theatre productions, art exhibitions, books, music, etc.

London newspapers could be classified into broad-sheets and tabloids, the former being A2 size publications, the latter being A3 size. The broad sheets tend to be the more “serious” newspapers, the tabloids the more populist mass circulation newspapers. A further sub-classification of the tabloids would be the “Red-Tops”. The most populist of the tabloids (and those perhaps most “down-market”, most pictorial, and the simplest to read), have red banner titles. These newspapers do not feel their readers are greatly interested in culture and the arts, do not publish reviews of films, plays, etc, and therefore do not contribute to this research. Another classification is in terms of the time and frequency of publication. Newspapers are published daily on weekdays Monday to Saturday, weekly on Sundays, or in the evenings on weekdays.

Major Newspapers Published in London – Table 1

<u>Publication Time</u>	<u>Broad-sheets</u>	<u>Tabloids</u>
Daily:	The Times The Guardian	The Daily Mail The Daily Express
Morning:	The Independent The Financial Times The Daily Telegraph	“Red Tops”: The Sun The Daily Mirror The Daily Star
Evening:		The Evening Standard
Sunday:	The Sunday Times The Sunday Telegraph The Observer The Independent on Sunday	

There are other newspapers published in London, and elsewhere in towns and cities in the UK. However, those listed above are the most significant in terms of stature and circulation, and 12 of them with consistent policies for reviewing arts offerings will be considered further in this paper.

### Newspaper Reviews

The extent to which newspapers review new productions is highly variable. The broad-sheets have more space, and therefore tend to review more productions and in greater depth. The dailies tend to print immediate reviews, often the day after the opening of the production, or the following day. The immediacy may have the disadvantage of a rushed judgement. The Sunday broad-sheets have generally both more space and more time for the critic to consider his or her comments. The tabloids tend to review fewer shows and the reviews are shorter. The advantage of a more succinct text may be balanced by an inability to consider details.

Most of the longer, in-depth reviews (or “notices”) tend not to have any summary quantitative assessment, but some reviews include such an assessment as part of their headline. Popular methods include a star rating of 1; 2; 3; 4; or 5 stars in ascending order of quality. Another version of this rating is used by the Guardian newspaper, and scores the offering 1 - 5, also in ascending order of quality. The categories are designated: 5 Unmissable; 4 Uncommon; 3 Uncontraversial; 2 Undesirable; and 1 Unprintable. This is essentially a five element Lickert Scale, though it might be argued that the designation “Unprintable” is inaccurate for the assessment of an already published review, however damning. Perhaps “Unspeakable” would be a better term that maintains the alliteration. Another system indicates 0 adequate; \* good; \*\* very good; \*\*\* outstanding; x poor, again a five element Lickert Scale. The Evening Standard has used the latter method for its main reviews alongside its compilation column which uses a simpler method (see later). This newspaper also publishes a weekly entertainment colour-supplement on Thursdays which contains summary reviews including films, coded to a system of: \*\*\* unmissable; \*\* excellent; \* good. Those not classified are presumably moderate or poor.

### Compilation of Reviews

Two London newspapers, the Sunday Telegraph and the Evening Standard, publish compilations of the views of various critics of the new shows (or new “openings”) for the previous week. Both cover new films opening in London or on general release and new theatrical productions, (plays, musicals, opera, ballet, dance). The Sunday Telegraph compilation is published of necessity on Sunday and is headed “Culture Vulture – Last Week’s Openings”. It reviews up to 6 films, 6 theatre productions and sometimes a ballet and/or opera. The compilation draws on the views of the critics of the 5 leading broad-sheet daily papers, the two non-Red-Top tabloids, and the Evening Standard. The assessment is on the basis of “Great Carrion” (i.e. good); “On the Turn” (i.e. adequate) or “Rotten” (i.e. poor). The visual presentation allows an “at a glance” assessment of the general critical view. It can indicate a consensus (or lack of one) among the critics and

the reader can quickly learn which presentations might be worth seeing, and which should be avoided.

The Evening Standard column is entitled “Last Week’s Openings - What the Critics Said”, and it generally publishes its compilation in the Tuesday edition. The format is very similar to that of “Culture Vulture”, with the shows judged to be “Good” (i.e. good), “OK” (i.e. adequate) or “Awful” (i.e. poor). The breadth of coverage is less than that of the Sunday Telegraph, as normally only about 5 films and 4 plays are considered, plus the occasional opera or “dance” production. However, the number of newspapers cited is greater, as the Tuesday publication day allows inclusion of the reviews of the four broadsheet Sunday papers. The same daily newspapers are cited as for the Sunday Telegraph compilation.

Essentially, both compilations use a simple three-element Likert scale, as this allows a visual presentation of great immediacy. The fact that both compilations use material from the same daily papers is helpful for analysis, but the greater number of shows considered in the Sunday Telegraph compilation means that only those considered in both compilations can be directly compared.

#### Creation of a Data-Set

A data base has been prepared for the year 2000 from the “Culture Vulture” compilation of the Sunday Telegraph, and the “What the Critics Said” compilation of the Evening Standard.

The Society of London Theatre (SLT) provides a definitive list of theatre productions each year in its report Box Office Data Report-2000. The British Film Institute (BFI) publishes a list of films presented for the first time, or formally reissued, each year. Not all films and theatre productions are reviewed, and there are many films that achieve only modest distribution, as is indicated by BFI data on gross earnings.

For cinema, 239 films were listed for the year 2000 by the BFI, 264 were included in the “Culture Vulture” compilation and 219 were included in the “Last Week’s Openings” compilation. There are occasional omissions from each compilation present in the other compilation, and 269 films are present in one or both compilations.

For theatre, the SLT list includes the following categories:

Plays; Musicals; Opera; Ballet; Dance.

These are generally included in the compilations. The SLT also lists three other categories not included in the compilations.

Children’s Shows; Entertainments; Sunday Performances.

217 productions were included in the “Culture Vulture” compilation, and 113 were included in the “Last Week’s Openings” compilation. For the Lyric Theatre productions, 100 were included in the “Culture Vulture” compilation, but only 26 were included in the “Last Week’s Openings” compilation. The complete data set has over 7600 items.

## Use of the Data-Set for Quality Assessment

It has been stated earlier in the Introduction that quality assessment of artefacts can be performed objectively against definitive and quantitative criteria. Assessment of service performance involves a more subjective assessment in that it involves the degree of satisfaction of the recipient of the service. Even so, some criteria can be formulated against which the recipients will judge the service. The same can be said for the views of the critics of films and theatre productions, but it will be difficult to get consensus as to the criteria to be included and their relative importance. However, it can be argued that the critics are a body of experts whose tastes may be more educated than those of the general film and theatre patrons, but are unlikely to be dissimilar. Furthermore, each individual critic might feel that he or she is writing for the readership of the particular newspaper. The views of individual critics may vary according to their personal tastes and values, but they are writing for a wide range of readers, whose tastes and values will also range widely. Arrays of reviews can be seen which are consistently very favourable, favourable, uncommitted, unfavourable or very unfavourable, or other arrays of view which are more mixed. This might suggest that some shows will be very acceptable or unacceptable to the great majority of potential customers, and others may be more controversial, with strong views both for and against.

There is a second human judgement being made in this exercise, that of the individual who prepares the compilation. This individual is an experienced arts journalist, often an Arts Editor of the newspaper. He makes an assessment of the text of the review, and puts it into one of 3 categories: good; adequate; bad. This is a 3 point Likert Scale, and could be considered rather crude. A 5 point Likert Scale might give more precision, and further work on the data base in the future will examine the translation of assessments on the “5 stars” system to the 3 division scale, where the former are available.

The assessment of quality in the two compilations is converted to a pseudo-numeric scale by assigning the score -1 to an unfavourable (“poor”) assessment, 0 to an intermediate (“adequate”) assessment and +1 to a favourable (“good”) assessment. Thus, if the views of the critics on the offerings of the year 2000, are evenly spread on a continuum from very good to very bad, the overall aggregate would come to 0. Hence the degree of positivity (0 to 1.0) or negativity (0 to -1.0) gives a pseudo-quantitative assessment of the favour or lack of favour of the critic towards the offering.

## Initial Assessment of the Data-set

Both the compilation of the Sunday Telegraph, and that by the Evening Standard consider the same eight daily papers; so cross correlation is possible. The consistency of the two compilations is very high for the cinema and theatre offerings. Coverage of the available data by the two compilations is also very high for cinema and theatre. The Lyric Theatre (Opera and Ballet/Dance) offerings are fewer in number, and the coverage by the two compilations varies. This is discussed in more detail later.

The coverage of the available films and stage productions is generally found to be slightly greater for the broad-sheet newspapers than for the tabloids. Some variation was detected in the benchmarks adopted by the critics of each newspaper. For example, for film criticism, The Times gave overall the most favourable reviews, and the Daily Mail the least favourable. However, the range of views from one extreme to the other only represented 17% of the spectrum, and the overall assessment was very close to the mean at 5% positive. However, for theatre productions, the critics were much more generous at 40% positive. This result suggests some more questions. For example, are the theatre productions presented in London generally of a higher quality compared to the films shown in the capital, or do London's theatre critics have lower standards than its film critics? The lyric theatre offerings are even more favourably regarded, with the "Culture Vulture" compilation recording a 44% positive view.

The preliminary conclusion is that the assessment method chosen is capable of providing a reasonable measure of the quality of these arts products. This will ultimately allow the research to proceed, provided suitable measures of economic performance are found. Initial examination of the relationship between the gross quality data on films and the available data on box-office receipts found no correlation, probably due to limited box-office data. On films where box-office data was available, some correlation was apparent, but it is clear that film quality may not be the primary driver of commercial success. The cinema, theatre, and lyric theatre data sets will now be considered in greater detail.

### The Data-Set for Film Reviews

The "Culture Vulture" compilation for the year 2000 lists 264 films, and the "What the Critics Said" lists 219. The "Culture Vulture" compilation has 50 films not considered by "What the Critics Said" and omits 5 films included in the other compilation. Therefore the films common to both compilations were 214, and the total films considered were 269. This shows a level of commonality between the two lists of about 80%. The criteria for inclusion is that the majority of the newspapers surveyed should have reviewed the film. The "Culture Vulture" compilation surveys 870 newspapers, rarely publishing an assessment below 5 as the following analysis shows:

Newspaper Reviews	8	7	6	5	4	<4	Total
Films in Compilation	193	52	14	2	3	0	264

Newspapers reviews per compilation entry = 7.63 (maximum possible = 8)

Total reviews considered 2014.

The compilation draws on a wider range of newspapers, including 4 Sunday broad-sheets, a total of 12 newspapers. However its criteria for inclusion appear more stringent than for "Culture Vulture", so fewer films (219) are included in the compilation.

Newspaper Reviews	12	11	10	9	<9	Total
Films in Compilation	160	41	13	5	0	219

Newspapers reviewed per compilation entry = 11.63 (maximum possible = 12)

Total Reviews considered 2546

Thus, as far as films are concerned, the “Culture Vulture” compilation of the Sunday Telegraph surveys a wider range of film offerings, but limits its analysis to the views of 8 weekday newspapers. The “What the Critics Said” compilation of the Evening Standard surveys fewer film offerings, but employs a wider range of critical opinion by including the views of the critics of the 4 Sunday broad-sheets as well as the same weekday newspapers considered by the Sunday Telegraph. Thus the number of film offerings considered by the “Culture Vulture” compilation is greater than that of the “What the Critics Said”, but the latter considers more critical reviews in total.

There are a number of causes of divergence between the number of films considered by the two compilations. 8 films are present exclusively in “Culture Vulture” at the beginning and end of the year and this may be due to differing views of year-end demarkation, and the tendency for the Evening Standard to have more limited publication over the Christmas/ New Year period.

Generally, the films omitted from one compilation and included in the other have fewer newspaper reviews, and so may be omitted for that reason. Some 25 films could be so assessed. Five films are re-issues and therefore may or may not be re-reviewed. However, some omissions appear arbitrary, and may be based on space considerations or other less rational criteria.

#### Preliminary Analysis of the Data Set for Films

The “Culture Vulture” compilation was very consistent, appearing every Sunday throughout the year 2000, a total of 53 compilations. For films, it cited 2104 reviews out of a possible maximum of 2112, though this maximum may be illusory as a particular newspaper may have omitted to review a particular film. The coverage of the 264 films, by the various newspapers, and the positivity of the views expressed in the articles by their critics, were as follows:

##### Daily Broad-sheets:

	Guardian	Times	Telegraph	Independent	Financial Times
Coverage %	100.0	99.2	98.1	96.6	93.2
Positivity	0.076	0.183	0.112	-0.059	0.061

##### Tabloids:

	Standard	Mail	Express
Coverage %	95.4	91.3	89.0
Positivity	0.123	-0.153	0.068

##### Grouped Data:

	Daily Broad-sheets	Tabloids	All Newspapers
Coverage %	97.4	91.9	95.4
Positivity	0.075	0.014	0.053

The cumulative data for the full year shows some variation in the coverage of film releases by the 8 selected newspapers. The Guardian achieves the maximum possible coverage (100%), but it must be understood that this is 100% of those films included in the “Culture Vulture” compilation. There are other films released, and indeed the Guardian may have reviewed films which the Sunday Telegraph compiler may not have elected to include. However, it gives a basis for comparison. The Express achieves the lowest coverage (89.0%). The broad-sheet newspapers generally had higher coverage than the tabloids, which might be expected as they usually have more space. The exceptions are the Financial Times, which as a newspaper specialising in business news devotes limited space to the Arts, and the Evening Standard, which has a policy of high coverage of a wide range of London entertainment.

The cumulative data also shows a small variation between the views of the various newspaper film critics. The Times gives overall the most favourable reviews (score 0.183) and the Daily Mail the least favourable (-0.153), with the other newspapers distributed between these extremes. However, this range (-0.153 to 0.183) is less than 17% of the total range of possible scores, so the divergence of view is modest. Though there may be differences of opinion between critics as to the worth of a particular film, over the period of a year, and some 260+ films, their cumulative view is very close, and about 5% positive of the zero position on the spectrum. This suggests that the critics’ benchmark is very close to the middle of the range, and their assessments between good; adequate; and bad; are evenly balanced.

A “critics cumulative quality assessment” of each of the 264 films in the “Culture Vulture” compilation was obtained by taking the mean of the score assigned by each critic. Overall mean scores from -1.00 to 1.00 were obtained, showing that the critics could be unanimous in their praise (or condemnation) of a particular film, and show every shade of opinion between these two extremes. With a maximum of 8 data sources and a 3 point Likert scale, the assessment is likely to be rather crude as there are a limited number of possible mean scores.

The “What the Critics Said” compilation of the Evening Standard appeared 49 times during 2000, and listed 219 films. It cited 2544 reviews out of a theoretical maximum of 2628, and the individual newspaper coverage, and the positivity of the critic’s views reported in the compilation, were as follows:

Daily Broad-sheets:

	Guardian	Times	Telegraph	Independent	Financial Times
Coverage %	99.5	98.6	100.0	98.6	95.4
Positivity	0.115	0.231	0.096	-0.024	0.024

Tabloids:

	Standard	Mail	Express
Coverage %	97.7	94.5	90.9
Positivity	0.224	-0.121	0.065

#### Sunday Broad-sheets:

	S. Times	S. Telegraph	Observer	Independent on S
Coverage %	98.2	91.8	99.5	96.8
Positivity	-0.139	0.075	0.087	0.000

#### Grouped Data:

	Daily Broad-sheets	Tabloids	Sunday Broad-sheets	All
Coverage %	98.4	94.4	96.6	96.8
Positivity	0.081	0.058	0.005	0.050

The reduced number of films reviewed, and the possibility of a slight difference of assessment standard by the two compilers could account for the small shift of all the positivity assessment. For example, for the “all newspapers” assessment:

	Daily Broad-sheets	Tabloids
Culture Vulture	0.075	0.014
What the Critics Said	0.081	0.058

The very small change represents a consistent “shift” for all the newspapers. The small changes in the coverage figures are directly due to the reduction in the number of films cited, and have no component of compiler subjectivity. An overall higher figure for coverage by the broad-sheets and the tabloids is noted for “What the Critics Said” compared with “Culture Vulture”. The implication of this small change is that the larger number of films surveyed by the “Culture Vulture” compilation tended to include films of marginal interest, which may or may not have been reviewed by the contributing newspapers.

#### The Data Set for Theatre

Both compilations included assessment of performances in the London theatre. The “Culture Vulture” compilation for theatre is significantly more comprehensive than that of “What the Critics Said” and is generally equivalent to that for films. The “What the Critics Said” analysis of theatre covers a significantly narrower range of theatre offerings compared with “Culture Vulture.”

Initial analysis of the theatre listings suggests that the aggregated critical opinion is much more positive. The “Culture Vulture” compilation indicates a positivity score of 0.402, and the “What the Critics Said” reports 0.420. This shows that the theatre productions of the year 2000 were much more highly regarded than the cinema productions. This begs more questions than it answers. Are the productions in the theatre generally higher quality, or do the theatre critics have lower standards?

The “Culture Vulture” compilation for the year 2000 lists 217 theatre productions, and the “What the Critics Said” lists 113. The commonality of the two theatre compilations is substantially lower than that for cinema, so a direct comparison of the overall figures would not be valid.

The “Culture Vulture” compilation surveys 870 newspapers, rarely publishing an assessment below 4 as the following analysis shows:

Newspaper Reviews	8	7	6	5	4	<4	Total
Productions in Compilation	56	33	36	52	34	5	217
Newspaper reviews per compilation entry = 6.02 (maximum possible = 8)							
Total reviews considered 1306							

The “What the Critics Said” compilation draws on a wider range of newspapers, including four Sunday broad-sheets, a total of 12 newspapers. However its criteria for inclusion appear more stringent than for “Culture Vulture”, so fewer theatre productions (113) are included in the compilation.

Newspaper Reviews	12	11	10	9	8	7	6	<6	Total
Productions	32	17	16	15	8	6	11	8	113
Newspapers reviews per compilation entry 9.16 (maximum possible = 12)									
Total reviews considered 1074									

### Preliminary Analysis of the Data Set for Theatre Productions

The “Culture Vulture” compilation was very consistent, appearing every Sunday throughout the year 2000, a total of 53 compilations. For theatre productions, it cited 1306 reviews out of a possible maximum of 1736.

The coverage of the 217 theatre productions, by the various newspapers, and the positivity of the views expressed in the articles by their critics, were as follows:

#### Daily Broad-sheets:

	Guardian	Times	Telegraph	Independent	Financial Times
Coverage %	83.9	91.7	86.6	77.0	50.2
Positivity	0.396	0.558	0.415	0.305	0.404

#### Tabloids:

	Standard	Mail	Express
Coverage %	88.5	65.4	58.1
Positivity	0.332	0.367	0.402

#### Grouped Data:

	Daily Broad-sheets	Tabloids	All Newspapers
Coverage %	77.9	70.8	75.2
Positivity	0.421	0.367	0.402

The cumulative data for the full year shows some variation in the coverage of theatre productions by the 8 selected newspapers. The Evening Standard achieves the highest coverage, and the Financial Times the lowest coverage. The broad-sheet newspapers generally had higher coverage than the tabloids.

The cumulative data also shows a small variation between the views of the various newspaper theatre critics. The Times gives overall the most favourable reviews, and the Express the least favourable, with the other newspapers distributed between these extremes. However, this range (0.558 to 0.302) is less than 13% of the total range of possible scores, so the divergence of view is modest, and lower than that for cinema (<17%). Though there may be differences of opinion between critics as to the worth of a particular theatre production, over the period of a year, and 113 productions, their cumulative view is quite close, and about 40% positive of the zero position on the spectrum. This suggests that the critic's benchmark is well into the positive area of the range, and their assessments are quite favourable.

A "critics cumulative quality assessment" of each of the 217 theatre productions in the "Culture Vulture" compilation was obtained by taking the mean of the score assigned by each critic. Overall mean scores from -1.00 to 1.00 were obtained, again showing that the critics could be unanimous in their praise (or condemnation) of a particular production, and show every shade of opinion between these two extremes.

The "What the Critics Said" compilation of the Evening Standard appeared 49 times during 2000, and listed 113 theatre productions. It cited 1074 reviews out of a theoretical maximum of 1356, and the individual newspaper coverage, and the positivity of the critic's views, as reported in the compilation, were as follows:

Daily Broad-sheets:

	Guardian	Times	Telegraph	Independent	Financial Times
Coverage %	89.4	96.5	92.0	84.1	62.8
Positivity	0.356	0.615	0.442	0.253	0.451

Tabloids:

	Standard	Mail	Express
Coverage %	100.0	78.8	59.3
Positivity	0.398	0.528	0.224

Sunday Broad-sheets:

	S. Times	S. Telegraph	Observer	Independent on S
Coverage %	54.0	78.8	73.5	81.4
Positivity	0.442	0.461	0.398	0.424

Grouped Data:

	Daily Broad-sheets	Tabloids	Sunday Broad-sheets	All
Coverage %	85.0	79.4	71.9	79.2
Positivity	0.427	0.398	0.431	0.421

With major differences in the number of theatre productions reviewed, (217 and 113), direct comparison between the two compilations would be invalid. Productions common to both compilations are being identified for review. However, assessment according to newspaper type is very close, but this may be a coincidence.

	Daily Broad-sheets	Tabloids
Culture Vulture	0.421	0.367
What the Critics Said	0.425	0.400

### The Data Set for the Lyric Theatre

The offerings of opera, ballet and dance productions are more limited, with only 105 productions noted, but “Culture Vulture” has a good coverage of what is available. The coverage of these Lyric Theatre productions by “What the Critics Said” is very limited. The “Culture Vulture” compilation cited 236 reviews of ballet and 278 reviews of opera. The “Culture Vulture” compilation for the year 2000 lists 47 ballets and 53 operas, a total of 100 lyric theatre productions. “What the Critics Said” lists 15 ballets and 11 operas, a total of 26 productions, so the coverage by “What the Critics Said” is about a quarter of the productions. The “Culture Vulture” compilation has 35 ballets and 42 operas not considered by “What the Critics Said” and omits 3 ballets included in the other compilation. Therefore the ballets common to both compilations were 12, and the operas 11. The total ballets considered were 50 and the operas 53. This shows a level of commonality between the two lists of about 24% for ballet and 21% for opera. The criteria for inclusion is that the majority of the newspapers surveyed should have reviewed the film. The “Culture Vulture” compilation for Lyric Theatre considered 524 reviews, rarely publishing an assessment on less than 4 reviews as the following analysis shows:

Reviews Assessed	8	7	6	5	4	<4	Total
Ballets in Compilation	1	8	31	29	26	5	100

Reviews per compilation entry = 5.14 (maximum possible = 8)

Total reviews considered 514.

The “What the Critics Said” compilation draws on a small part of the lyric theatre productions of 2000, (25%), so it is not felt that there is much value on further analysis of this compilation.

## Preliminary Analysis of the Data Set for Lyric Theatre

The coverage of the 105 ballet and opera productions, by the various newspapers, and the positivity of the views expressed in the articles by their critics, were as follows:

Ballet productions:

Daily Broad-sheets:

	Guardian	Times	Telegraph	Independent	Financial Times
Coverage %	72.3	93.6	74.5	78.7	70.2
Positivity	0.559	0.341	0.229	0.351	0.273

Tabloids:

	Standard	Mail	Express
Coverage %	83.0	19.1	10.6
Positivity	0.641	0.444	1.000

Grouped Data:

	Daily Broad-sheets	Tabloids	All Newspapers
Coverage %	77.9	37.6	62.8
Positivity	0.350	0.642	0.415

Opera productions:

Daily Broad-sheets:

	Guardian	Times	Telegraph	Independent	Financial Times
Coverage %	83.0	86.8	90.6	66.0	69.8
Positivity	0.341	0.587	0.333	0.571	0.459

Tabloids:

	Standard	Mail	Express
Coverage %	69.8	22.6	5.7
Positivity	0.486	0.667	0.333

Grouped Data:

	Daily Broad-sheets	Tabloids	All Newspapers
Coverage %	79.2	32.7	61.8
Positivity	0.452	0.519	0.466

The variation is wider in the coverage of the ballet and opera productions than for cinema and theatre, and this will increase the variation in the overall views of the critics.

The variety of notices for the ballet productions has the Telegraph giving the least friendly reviews (0.229) and the Standard as the most positive (0.641). The range is 0.412, which represents 20.6% of the total spectrum. The Express figure is discounted, as the coverage is only 10% of the total productions.

The range for the opera productions again has the Telegraph as the least friendly reviewer (0.333) and the Mail as the most positive (0.667). The range is 0.334, which represents 16.7% of the total spectrum. The Express figure is again discounted, though it equates to that of the Telegraph.

### Future Work

More detailed analysis of the data is being performed with the aid of the SSPS software package. Compiler bias and the consistency of the individual critics are being examined. Some detailed checks of the data set will be undertaken where minor anomalies have been found, and the full statistical analysis will be completed. The question of the subjectivity of the compiler will be addressed by comparison of those parts of the data-set which are common to each compilation. Further analysis will be undertaken to determine the consistency of quality assessment by individual critics.

### Conclusions

The use of the compilations of film, theatre, and lyric theatre reviews has been justified as an expert information source based on the principles of the Delphi technique. A data set has been compiled using the assessments of the quality of the films, plays, etc. which is currently being subjected to further detailed analysis. Initial analysis of the film, theatre and lyric theatre productions in London in 2000 and reviewed by the major London newspapers cover a very wide range of quality. There is evidence of both high consistency, and high diversity of critical view of the various productions. More detailed analysis of the data set will indicate the extent of the consistency of the various critic's views of the quality of the films presented in 2000. The initial analysis suggests that the film critics have a very balanced overall view: 5% positive. Theatre critics have a more generous view: 40% positive. Ballet critics assess at 42% positive, and opera critics at 47% positive.

It is concluded that the quality assessment method being developed is capable of providing a reasonable measure of the quality of arts products. This will allow the overall research plan to proceed towards the assessment of the relationship between product quality and commercial performance of arts products.

## References

Crosby, Philip B., (1979) *Quality is Free* McGraw-Hill, New York

Gilhespy, Ian, and Reavill, Richard (2001) "Sponsorship, Social Responsibility, and the Arts in Britain" *6<sup>th</sup> International Conference on Arts & Cultural Management*, Brisbane, July 2-4, 2001.

Reavill, Richard, (1998) "A Hard Nut to Crack", *Dance Now*, Vol. 7, **1**, pp. 21-26

Reavill, Richard, (1999) "The Problems of the Lyric Theatres in London - A Stakeholder View" *5<sup>th</sup> International Conference on Arts & Cultural Management*, Helsinki, June 13-17, 1999.

Reavill, Richard, (2000) "Towards an Integrated System", *Dance Theatre Journal*, Vol. 15, **4**, pp. 39-43.

Reavill, Richard, (2002) "Stretton under Attack", *Dance Europe*, **56** October 2002, pp. 26-27

Van Gundy, Arthur B. (1988) *Techniques of Structured Problem Solving*, 2<sup>nd</sup> ed., Van Nostrand Reinhold pp. 323-8