Applying Co-word Analysis for Exploratory Studies: a Demonstration using Web-based Data on the Authenticity Construct

Richard G. Starr, Jr.* *The University of Auckland.* rg.starr@auckland.ac.nz
Karen V. Fernandez. *The University of Auckland.* k.fernandez@auckland.ac.nz

Keywords: authenticity, scale development, co-word analysis, multidimensional scaling

Abstract

Many marketing studies require exploratory analysis of construct-related terminology for the development of new measurement scales. The ability to conduct such analyses quickly, and with good external validity, is useful. This paper applies an underutilised technique, co-word analysis (Callon, Courtiale and Laville 1991), to web-based data on authenticity, a construct of increasing interest to consumer behaviour researchers. The technique allows accurate quantitative assessment, in unbiased naturalistic settings, of the interrelationships among the different terms used to describe a concept.

A list of key descriptors of authenticity was developed from a range of sources. Proximity data for these terms was developed using co-word analysis within a trademark database, and on the entire World Wide Web. The proximities were used to generate multidimensional scaling maps of the authenticity descriptors, providing useful insights for scale development. This demonstration suggests that the technique could be useful in a broad range of studies.

Introduction: Study Approach

Many marketing studies incorporate constructs with fuzzy boundaries. Authenticity is such a construct: idiosyncratic and socially-constructed, including personal and individual judgments, about an item or an experience, with many context-specific elements. This does not mean, however, that there are no commonalities across multiple individuals' experiences of authenticity. To the extent we can understand any common perceptions, we are likely to gain greater clarity about the concept itself, and better insights about how authenticity relates to marketing issues. Co-word analysis provides one way to begin this clarification by mapping individuals' impressions of the terminology which describes authenticity.

In this paper describing our initial phase of this work, we explicitly incorporated the commonality of others views' of authenticity. Our first step was to draw upon existing research and linguistic resources to derive an unquantified understanding of common synonyms and terms for describing authenticity.

Next, a co-word content analysis approach was used to develop proximity data for commonly-used terms for authenticity. This step created a contingency table describing the frequency of usage for individual terms, as well as the coincidence of different pairs of descriptors. Data was drawn from two sources, the United States Patent and Trademark Office database (covering purely commercial speech) and the World Wide Web as a whole (both commercial speech and common usage).

Finally, the contingencies were used as similarity measures to generate multidimensional scaling maps of the authenticity descriptors. These MDS maps were used to provide further insight into how descriptors of authenticity interrelate, as a precursor to formal scale development.

The Context for this Method

A substantial body of previous research in marketing suggests that authenticity is inherently in the eye of the beholder. The personal and idiosyncratic nature of authenticity lends itself to qualitative research, and most existing research is interpretive in nature. In one recent study, Beverland and Farrelly (2010), found that consumers use salient cues to make judgments about the authenticity of an object. This clearly suggests that some cues might lead to stronger perceptions of authenticity than others, and opens the door for a range of future research. However, in order to empirically examine the antecedents and consequences of authenticity, it is necessary for researchers to be able to measure the construct.

There is an emerging body of quantitative and experimental research into authenticity. Grayson and Martinec (2004), explicitly recognizing the personal nature of authenticity judgments, asked respondents to identify a site feature that they deemed authentic or inauthentic. This categorical treatment was suitable for their purposes, but it necessarily glosses over subtleties of the concept. We suggest it is possible to define and measure authenticity at a more refined level and develop a usable scale to assist in future research. Appropriate exploratory analysis is needed as one part of this process.

Prior marketing-related research on authenticity

Prior research generally agrees that authenticity is a difficult to define multidimensional construct with many context-specific aspects. Authenticity can be viewed as an assessment of a product or experience made by an evaluator in a particular context (Grayson and Martinec 2004). As such, it represents the interaction of an object, place, and person. Although it seems that authenticity should be an inherent property of an object, it is often difficult or impossible to demonstrate physical properties which indicate authenticity.

Despite the importance of this concept, the research around authenticity is contentious, leading Beverland and Farrelly (2010) to describe the nature of authenticity in consumption as "contested." A number of authors believe that authenticity is of increasing importance in a less-traditional postmodern world, while many postmodernists themselves might view the concept as spurious or irrelevant. Authenticity-related claims also appear to be of increasing importance in the marketing of both products and services.

Arnould and Thompson (2005) place research into authenticity within consumer culture theory (CCT) and specifically link it to personal authenticity. They summarise by saying "… the market produces certain types of consumer positions that consumers can choose to inhabit" (Arnould and Thompson 2005, p. 871). Authenticity is, then, a position that consumers can align themselves with or towards. Authenticity has been studied in a number of specific contexts, including tourist attractions (Grayson and Martinec 2004), luxury alcoholic beverages (Beverland and Luxton 2005; Beverland, Lindgreen and Vink 2008), and automobiles (Leigh, Peters and Shelton, 2006; Brown, Kozinets and Sherry 2003). The growing salience of authenticity is seen in greater visibility in the business press, including a fictionalised Harvard case study (Weinberger 2008) and entire book devoted to marketing via authenticity (Gilmore and Pine 2007). Overall, the research on authenticity is rich and detailed, but very context-specific, and difficult to generalise.

Method Description: Assessing how authenticity is lexicalized

We assume that concept of authenticity is related, but not identical, to the words we use to describe it. "Although words depend on the existence of concepts, the inverse is not true: concepts can, and do, lead a life independently of words (Fellbaum 1998, p. 8)." By reviewing

concept-related terms, including known synonyms and descriptors, it is possible to understand the range of ways that a given concept is lexicalized.

This study began with a search for "authentic" and "authenticity" in the WordNet 3.0 database (Miller et al 1990), a lexical database developed to be a comprehensive set of synonyms. The database contained four synonyms for authentic: *reliable*, *bona fide*, *unquestionable*, and *veritable*, and two synonyms for authenticity: *genuineness* and *legitimacy*. The Oxford English Dictionary (1989) confirmed that the WordNet synonyms appeared to be tangential to the underlying idea of authenticity. Only two of the six WordNet synonyms (*reliable* and *genuine*) appeared in the OED definitions. Because two of the terms (*veritable* and *bona fide*) are uncommon, a wider search was deemed necessary.

Prior research across many fields was examined, and it is apparent that the authenticity concept has been lexicalized using a fairly broad range of terms. Drawing on a range of sources, an initial list of 13 key descriptors was developed: *authentic, authenticity, original, real, certified, genuine, authorized, vintage, guaranteed, true, classic, licensed, and historic*. At this stage, terms to describe specific types of authenticity (iconic, indexical, expressive and so on) were not specifically sought. However, some of the terms chosen do relate, to a greater or lesser degree, to specific types of authenticity.

This list was judged to provide a workable initial range of authenticity-related vocabulary. However, it offers no indications of the strength of association between individual descriptors and the underlying concept, nor any idea of how these descriptors are conceptually related. To better understand which words are more- and less-associated with the authenticity concept, a content analysis was conducted to determine the frequency of these words, and their co-occurrences, in commercial discourse.

Authenticity in commercial speech

As a proxy for commercial discourse, the United States Trademark and Patent Office (USPTO) database was used as a sampling frame. It contains all U.S.-registered trademarks, along with a description of the language and basic claims to which the trademark holder is asserting copyright. The unit of analysis was individual trademark records. Initially, a targeted search for each of the 13 previously-derived individual authenticity descriptors was conducted using the dedicated search engine within the online USPTO trademark database.

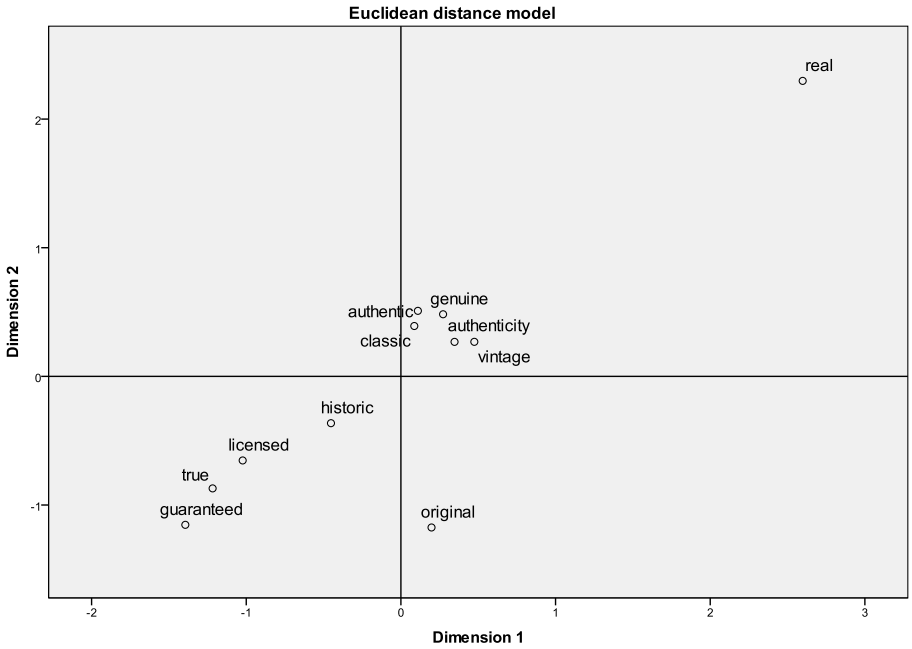
As a second step, the USPTO search engine was used to conduct a Boolean search for authenticity-related word pairs which appeared within the same trademark listing. The frequency of word pair co-occurrence was tabulated following standard co-word analysis protocols (Callon, Courtiale and Laville 1991). When a pair of words occurs within the same record, they are considered coincident and are counted. The total frequency of a word pair co-occurrence is a metric measure of their proximity or similarity. The words original and authentic occur in combination 56 times, for example. The entire correspondence table of word pairs was constructed, and a symmetrical data matrix (below) was developed.

Table 1: Proximity Data for Authenticity Word-Pairs in USPTO Data

	authentic	original	real	certified	genuine	authorized	vintage	guaranteed	true	classic	licensed	authenticity	historic
authentic		56	25	20	19	16	16	13	10	10	9	7	7
original	56		95	21	56	230	28	15	34	45	35	2	16
real	25	95		230	38	202	32	159	143	51	131	19	104
certified	20	21	230		7	1621	3	18	11	6	84	9	3
genuine	19	56	38	7		21	3	12	10	6	1	5	1
authorized	16	230	202	1621	21		2	11	22	13	57	4	3
vintage	16	28	32	3	3	2		13	5	60	1	1	25
guaranteed	13	15	159	18	12	11	13		2	2	4	1	0
true	10	34	143	11	10	22	5	2		18	5	5	0
classic	10	45	51	6	6	13	60	2	18		0	0	15
licensed	9	35	131	84	1	57	1	4	5	0		0	2
authenticity	7	2	19	9	5	4	1	1	5	0	0		0
historic	7	16	104	3	1	3	25	0	0	15	2	0	
total pairs	208	577	1204	2013	160	2186	173	237	255	216	320	46	169
total instances	2055	14571	85741	6850	1475	6862	2546	1998	9280	11100	2323	471	1777

This data was used with the metric ALSCAL multidimensional scaling procedure to map the relative position of the descriptors in a two-dimensional solution. When an initial MDS solution was developed, the terms *authorised* and *certified* mapped as outliers, which compressed the remainder of descriptors into a tight cluster. The outlying terms were removed and a new 11 item solution was calculated, yielding an r^2 =.583 and Young's S-stress =.448. This provides better understanding of how the descriptors relate to one another. The central ideas of *authentic*, *authenticity*, and *genuine* map are plotted at the centre of the map, lending credibility to the results. More legalistic concepts (*licensed*, *true*, and a*uthorised*) are relatively close together, but separated from the main descriptors. The term *real* is quite far from other descriptors, apparently because the word is used in two distinctly different contexts, one meaning "genuine," and the other meaning property ("real estate").

Figure 1: MDS Map of Authenticity based on USPTO Data



Authenticity on the World Wide Web

The same co-word analysis approach was employed using the entire World Wide Web as a sampling frame. Unlike the patent database, which is restricted to commercial speech, the WWW data contains a mixture of commercial and individual speech. In this case, individual webpages were the unit of analysis, and we derived the frequencies which describe the co-occurrence of a pair of terms within the same webpage. This does not strictly guarantee that the two descriptors are used in close proximity, nor that they are necessarily conceptually linked in the text of the webpage.

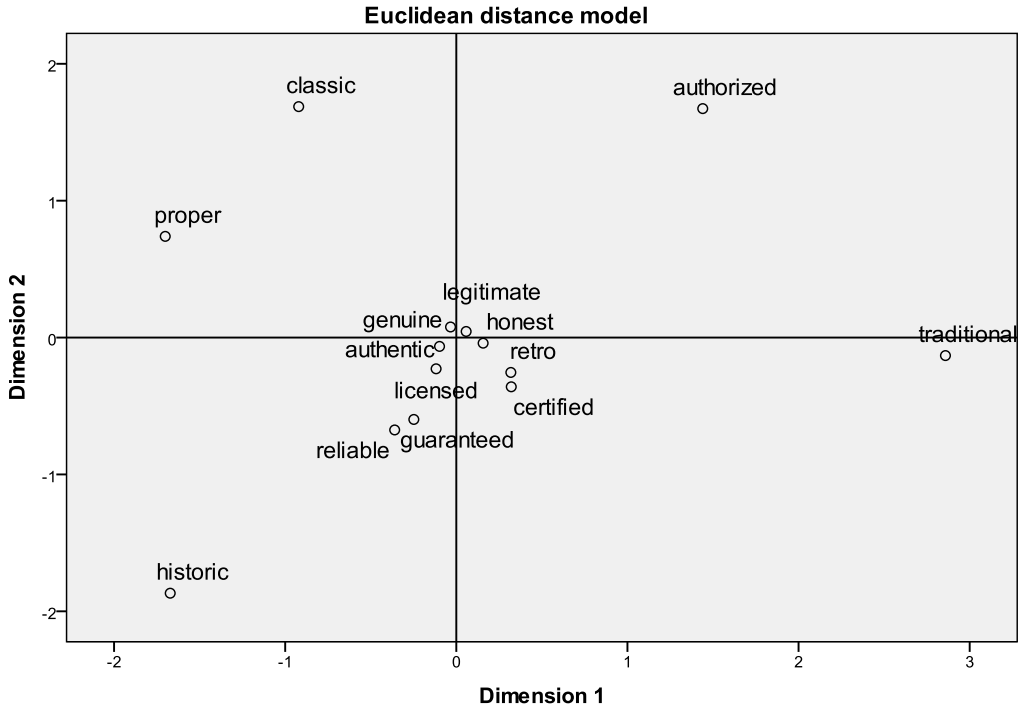
However, this is very similar to the type of association measured in co-citation analyses of academic research. A co-citation indicates only that one author has referred to another's work. It does not indicate whether the citation is favourable or unfavourable, or extensive or cursory. All of these analyses make the basic and reasonable assumption that stronger association implies stronger conceptual linkage. In the case of the WWW analysis, a spot check of a number of co-word occurrences confirmed that the word pairs are, on average, reasonably conceptually related. We believe these results are sufficient.

Data on word pair co-occurrences was collected using Google, the most commonly used web search engine. There are valid concerns about using commercial search technologies for

data collection. One key issue, the use of proprietary algorithms to order search results, was nullified by using total frequencies of co-occurrence, and ignoring relative rankings. However, a spot check of other search engines (Yahoo, Bing, and AltaVista) did reveal differences in both the number and relative proportion of individual word pairs. We proceeded on the assumption that this data has good internal validity: it is an accurate representation of the WWW as mapped by Google. The external validity of this data (how closely it actually corresponds to the full WWW) is not known. We believe the data is sufficiently accurate for exploratory research, with the potential for confirmation using more stringent techniques.

A 14 item solution was calculated from the Web data, yielding an r^2 =.478 and Young's S-stress =.574. As with the USPTO data, the general authenticity concept is in the centre of the map, reconfirming its centrality. More tangential descriptors map at the periphery.

Figure 2: MDS Map of Authenticity based on WWW Data



Conclusion

This paper demonstrates an underutilised technique, co-word analysis, in conjunction with multidimensional scaling. It enables a more rigourous approach to scale development by quantifying unbiased data which embodies the language consumers use to describe a concept. It is a useful precursor to other techniques. The demonstration explores one substantive construct, authenticity, but the technique is applicable to exploratory research in a broad range of contexts. When Callon, Courtiale and Laville first developed this technique, the internet was quite limited. Since that time, the growth of the internet has allowed broad access to databases of many kinds, including the web itself. Search engines allow the ability to quickly derive proximity data, either within a circumscribed context or of the entire web, which can be used in multidimensional scaling, or a range of other statistical analyses. Co-word analysis offers a fast and versatile approach for early-stage exploratory research into a range of marketing-related constructs.

References

Aaker, Jennifer L. (1997), “Dimensions of Brand Personality,” Journal of Marketing Research; 34 (August), 347-356.

Arnould, Eric J. and Craig J. Thompson (2005), “Consumer Culture Theory (CCT): Twenty Years of Research,” Journal of Consumer Research 31 (4), 868-882.

Beverland, Michael B. and Sandra Luxton (2005), “Managing Integrated Marketing Communication (IMC) Through Strategic Decoupling: How Luxury Wine Firms Retain Brand Leadership While Appearing to be Wedded to the Past,” Journal of Advertising, 34 (4), 103-116.

Beverland, Michael B., Adam Lindgreen, and Michael W. Vink (2008), “Projecting Authenticity Through Advertising: Consumer Judgments of Advertisers' Claims,” Journal of Advertising, 37 (1), 5-15.

Beverland, Michael B. and Francis J. Farrelly (2010), “The Quest for Authenticity in Consumption: Consumers' Purposive Choice of Authentic Cues to Shape Experienced Outcomes,” Journal of Consumer Research, 36 (1), 838-856.

Brown, Stephen, Robert V. Kozinets, and John F. Sherry, Jr. (2003), “Teaching Old Brands New Tricks: Retro Branding and the Revival of Brand Meaning,” Journal of Marketing, 67 (July): 19-33.

Callon, M., J. P. Courtiale, and F. Laville (1991), "Co-word analysis as a tool for describing the network of interactions between basic and technological research: the case of polymer chemistry," Scientometrics, 22 (1), 155-205.

Fellbaum, Christiane (1998), ed., WordNet: an electronic lexical database. Boston: Massachusetts Institute of Technology.

Gilmore, James H. and B. Joseph Pine II (2007), Authenticity: What Consumers Really Want, Boston: Harvard Business School Press.

Grayson, Kent and Radan Martinec (2004), “Consumer Perceptions of Iconicity and Indexicality and Their Influence on Assessments of Authentic Market Offerings,” Journal of Consumer Research, 31 (2), 296-312.

Leigh, Thomas W., Cara Peters, and Jeremy Shelton. (2006). “The Consumer Quest for Authenticity: The Multiplicity of Meanings within the MG Subculture of Consumption,” Journal of the Academy of Marketing Science, 34 (4), 481-493.

Miller, George A., Richard Beckwith, Christiane D. Fellbaum, Derek Gross and Katherine Miller (1990), “WordNet: An online lexical database,” International Journal of Lexicography, 3(4), 235-244.

Reisinger, Yvette and Carol J. Steiner (2006), "Reconceptualizing Object Authenticity," Annals of Tourism Research, 33 (1), 65-86.

United States Patent and Trademark Office Trademark Electronic Search System (TESS)(2010), http://tess2.uspto.gov/.

Weinberger, David (2008), "Authenticity: Is It Real or Is It Marketing?," Harvard Business Review, 2008 (March), 33-43.