


The application of DICTION to analyse qualitative data: A luxury brand perspective

**Author:**Beate E. Stiehler-Mulder¹ **Affiliation:**

¹Department of Marketing Management, College of Business and Economics, University of Johannesburg, Johannesburg, South Africa

Corresponding author:Beate Stiehler-Mulder,
bstiehler@uj.ac.za**Dates:**

Received: 26 Apr. 2019

Accepted: 11 Oct. 2019

Published: 05 Dec. 2019

How to cite this article:

Stiehler-Mulder, B.E., 2019, 'The application of DICTION to analyse qualitative data: A luxury brand perspective', *Acta Commercii* 19(1), a783. <https://doi.org/10.4102/ac.v19i1.783>

Copyright:

© 2019. The Authors. Licensee: AOSIS. This work is licensed under the Creative Commons Attribution License.

Orientation: The evolution of technology has given rise to numerous computer-aided programs for analysing textual data, shaping the advancement of qualitative research.

Research purpose: This study aimed to demonstrate how some of the critique associated with text analysis programs, as well as the dilemmas experienced in conducting research in the social sciences (achieving either precision or depth in the results, but seldom both), may be overcome.

Motivation for the study: The South African luxury consumer market is described as complex and heterogeneous. Through the application of the text analysis software DICTION, this study demonstrated how this software can be applied to analyse in-depth interviews, and gain not only unique behavioural insights into a complex consumer group, but also produce quantitative output that allows for the development of distinct customer market segments.

Research design, method and approach: A qualitative approach was adopted to collect the data, while the analysis using the text analysis program DICTION produced both qualitative insights and quantitative scores. The latter were used to perform a cluster analysis and develop unique customer segments.

Practical/managerial implications: This study reveals how managers may swiftly obtain insights into lengthy transcripts of interview data, obtain quick answers about consumers' behaviour and develop consumer market segments using the output obtained from the text analysis software DICTION.

Contribution/value-add: The results demonstrate how both precision (statistical analysis) and richness (context from the transcripts) may be achieved in a research study, and provide a unique method to research, understand and segment a complex consumer market.

Keywords: text analysis software; qualitative research; DICTION; luxury brands; South Africa; consumer discourse; luxury consumer segmentation.

Introduction

In his landmark work on the dilemmatics of research, McGrath (1981) argues that research involves dealing not so much with a set of problems to be solved as with a set of dilemmas to be resolved. Fundamentally, there is no single 'true research method', particularly in the social sciences, which will guarantee the best outcomes in a given setting with a given set of resources. As McGrath (1981:179) noted, 'all research strategies and methods are seriously flawed; often with their very strengths in regard to one desideratum functioning as serious weaknesses in regard to other, equally important, goals'. For example, in order to be able to generalise, social scientists, and marketing scholars in particular, resort to survey research. This might enable them to state that, under a given set of circumstances and following a survey, marketing executives would act in a particular way, or that a large sample of consumers prefer brand A over brand B.

However, the method is flawed in that it is unable to offer any precision or control, and is also incapable of providing richness or context. If precision and control are the prime research objectives, then some kind of experiment is required; but, of course, this method also lacks any kind of richness or context, and now foregoes generalisability. Methodologies such as focus groups, case studies and in-depth interviews provide richness and context, but do so at the cost of losing generalisability, control and precision. Achieving all three research objectives is highly unlikely, if not impossible. As McGrath might put it, 'the researcher is always sitting on a three-horned dilemma'.

Read online:

Scan this QR code with your smart phone or mobile device to read online.

Not surprisingly, marketing scholars have tended to focus their research on a single objective – whether generalisability, precision or context – and have chosen their methodologies accordingly. More rarely do we see papers incorporating more than one approach to achieve more than one objective. However, when this does occur, these papers gain considerable attention and are widely cited. For example, Arnould and Price's (1993) research on the extraordinary experiences that consumers enjoyed while river rafting used a challenging mix of qualitative techniques, such as observation and in-depth interviews, in combination with robust qualitative tools in the form of survey questionnaires.

This article attempts to resolve at least two horns of McGrath's three-horned dilemma. In doing so, the article provides further quantitative insight into data obtained from a number of qualitative in-depth interviews. The role of qualitative research and content analysis is reviewed first, followed by the use of computer-aided analysis tools. The software DICTION is then discussed, and its ability to determine conversational tone to uncover underlying dimensions within consumer conversations, and its ability to quantify the output generated, are also noted. The qualitative study, consisting of a series of 16 in-depth interviews with consumers designed to elicit their views and feelings about luxury brands in the emerging South African market, is then described. How the text from each individual was analysed with DICTION is then described, showing how it enabled scores on five core dimensions of interest to be allocated to each respondent. In this way, an 80-cell contingency table could be constructed (16 respondents by five dimensions). These data served as input to a hierarchical clustering procedure that enabled the identification of four market segments (more distinct groups within the respondent sample) in relation to the dimensions. The limitations of the approach, the managerial implications and avenues for future research are then discussed.

Conceptual background

Qualitative research, content analysis and computer-aided software

Qualitative research methods have been an accepted approach in market research for decades. The methods' value has been proven to generate insight into consumer use of vocabulary, for example, to uncover consumer purchase motivations, and to guide the development of products and services (Bailey 2014). Qualitative data come from a number of sources, but arguably the most common sources remain focus groups or individual semi-structured in-depth interviews (Pistrang & Barker 2012). In-depth interviews, in particular, have been commended for their ability to obtain rich detail from one individual at a time, for removing socially desirable responses (often associated with the use of focus groups), and because they allow an interviewer to use projective techniques (Hair et al. 2013).

In-depth interviews are typically recorded and transcribed before analysis begins. It is argued that the analysis of

qualitative research resides within four families of approaches: *thematic analysis approaches* (aimed at identifying central ideas), *narrative approaches* (attention paid to the unfolding of events), *language-based approaches* (focusing on the underlying social rules that govern how language functions to achieve certain conclusions), and lastly, *ethnographic approaches* (characterised by broad data collection and the observation of participants) (Pistrang & Barker 2012). *Thematic analysis approaches* may be used to analyse a wide array of qualitative sources, but are arguably most often used to analyse interview data. As a structured approach, content analysis in particular is also regarded as a qualitative method that produces quantitative output, mainly because content analysis focuses on identifying themes or categories in the data that may be counted according to the frequency with which they occur (Pistrang & Barker 2012).

The evolution of technology and computer-aided programs for analysing textual data has further contributed to the proliferation of more systematic content analysis methods and, as a result, has shaped the advancement of qualitative research (Feltham-King & Macleod 2016; Krippendorff 2013). However, the use of computer software to analyse text has not escaped critique. According to Schmidt (2010), one of the main challenges of computer-aided software is these programs' questionable ability to 'understand' textual information. A pertinent problem identified includes the use of homographs, which are 'words with the same spelling but with different meaning' (Schmidt 2010:485). DICTION was developed as a mainframe text analysis program in the 1970s by Roderick P. Hart, and later evolved into a personal computer (PC)- and notebook-based utility (Diction 2019). This text analysis program claims to be valuable in overcoming some of the current critique associated not only with text analysis programs, but also with the challenges of achieving both precision (the quantification of data) and richness (context) in a research study.

The DICTION text analysis program

A major strength of DICTION lies in the fact that it was designed by a communications researcher. It has the ability to produce word counts, and it also depends on the subtle power of word choice and verbal tone, unlike most text analysis software, which largely relies on word counts only (Short & Palmer 2008). The dictionaries used in the program were constructed from the analysis of more than 20 000 texts, yielding a total corpus of 10 000 words (Sydserff & Weetman 2002). DICTION also provides the researcher with the opportunity to use predefined dictionaries, or as many as 30 custom dictionaries, built with words that the user has defined (e.g. topical or negative words) to suit their research needs (Diction 2019).

The program also provides a built-in normative database consisting of 50 000 previously analysed texts, across 50 different message types, which have been collected over many years of research. A user may choose to apply these general norms, or select from a further 35 sub-categories of

different text types (e.g. newspaper editorials, speeches and business reports) (Diction 2019) to obtain a comparative understanding of their findings. By default, the program also provides a 500-word average analysis for all long texts (shorter texts adhere to a 500-word norm), as the software developer argues that around 500 words are needed to convey a true understanding of a person's overall style of language habits (Diction 2019).

The results produced by DICTION are written to numeric files that may be used for later statistical analysis. The output options provided include raw totals, standardised scores, word and character counts and percentages, providing the user with a variety of ways to understand the text being processed (Diction 2019). The program also has an advanced lexical feature that makes a modest statistical accommodation for homographs (Short & Palmer 2008). Potentially confusing homographs, which represent around 10% of DICTION's search words, are weighed differentially based on the *Word Frequency Dictionary* by Easton (1940), while benign homographs (implying different ideas that belong to the same domain of meaning) are ignored by the program (Short & Palmer 2008). DICTION's ability to accommodate homographs therefore addresses Schmidt's (2010) general concern that text analysis programs are not able to identify and accommodate homographs in written texts.

DICTION also has the ability to read a variety of text formats, and can accept a large number of files within a single project using its power analysis tool for enhanced speed and reporting efficiency (Diction 2019). The user also does not have to spend an inordinate amount of time 'cleaning' text files by removing extraneous and commonly occurring words or common articles and pronouns, as DICTION simply ignores words that are not in the dictionaries being used. DICTION's lexical features (which allow for the study of vocabulary and word choice) are especially useful for providing richness to the text, and can also be applied to undertake an analysis of sentiment. Its core dictionaries identify five categories and 35 sub-categories (Diction 2019). The five main categories and their meanings include:

- *Activity*: Language featuring movement, change, the implementation of ideas and the avoidance of inertia.
- *Optimism*: Language endorsing some person, group, concept or event, or highlighting their positive achievements.
- *Certainty*: Language indicating resoluteness, inflexibility and completeness and a tendency to speak *ex cathedra*.
- *Realism*: Language describing tangible, immediate, recognisable matters that affect people's everyday lives.
- *Commonality*: Language highlighting the agreed-upon values of a group, and rejecting idiosyncratic modes of engagement.

The software can therefore be used effectively for the dual purpose of obtaining both precision in the form of statistical analysis and depth and richness in the form of textual analysis.

Analysing consumer conversations about luxury brands in South Africa using DICTION

The luxury goods market (which includes designer apparel, handbags, accessories, fine jewellery, watches, cosmetics and fragrances) is projected to grow to 25% of the global demand by 2025 in emerging economies. South Africa is the latest country to join the group of emerging BRICS countries (Brazil, Russia, India, China, South Africa), and represents the African continent in this group (BRICS 2015). It is also described as the prime shopping destination for luxury items on the African continent (Deloitte 2014; Euromonitor 2019), and is considered an important gateway for luxury brands entering the rest of Africa (Euromonitor 2016, 2019). South African luxury consumers are, however, poorly understood (Steinfeld 2015), and are described as heterogeneous. The country's middle- and high-income groups are argued to possess similar buying power to consumers in developed countries (Sonnenberg, Erasmus & Schreuder 2014), yet only a small portion of its population follows the developed markets' luxury trends (Atwal & Bryson 2014). Exploring luxury behaviours in this market is therefore important, but the complex nature of its consumers (Bills & Tabot 2017) suggests the need to use innovative research techniques to allow for an in-depth understanding of luxury behaviours.

Against this background, the complexity of South Africa as an emerging market specifically for luxury goods is acknowledged; and by incorporating DICTION as an analysis tool to explore this specific market, this study aims to address the following objectives:

- to uncover the underlying dimensions from consumers' discourse about luxury brands in the South African market
- to identify segments of luxury consumers based on the dimensions identified from their discussions about luxury brands
- to discover the most influential participants for each of the five dimensions.

Methodology

Importantly, the author of this article followed the guidelines of Fine and Kurdek (1994), and concedes that the data set used for the purpose of this article is an existing data set of transcribed interviews that was originally collected by Stiehler in 2014 and 2015, and published in Stiehler and Tinson (2015) and Stiehler (2016). However, based on the three objectives outlined above, this article has a distinctive focus in that it seeks to identify new insights from the collected data set using the software analysis program DICTION.

In collecting the initial data during 2014 and 2015, an extant segmentation approach of luxury consumers in South Africa recommended by Atwal and Bryson (2014) was used as a starting point to identify participants for the study. This segmentation approach argues that the South African luxury market consists of four segments: the *deluxe aspirers*,

the *self-made* or 'new money', the established *business magnate* and the *money aristocracy* (Atwal & Bryson 2014). A number of aspects were considered in identifying suitable participants to be interviewed. The geographical location identified to source participants was the Gauteng province in South Africa – the economic hub of the country and home to a number of luxury brands (Gauteng Online 2016). Respondents from among South Africa's economically active age groups, typically between 25 and 59 years of age, were considered (BRICS 2013), and the diverse racial profiles that exist in South Africa, as well as equal representation in terms of gender, were deemed important. The sample therefore included an equal number of white people and black people (the two dominant race groups in the Gauteng region) as participants, as well as men and women.

The detailed participant profiles that were required necessitated the use of a marketing research specialist to recruit the participants. The specialist considered the demographic and location specifications supplied, as well as the detailed descriptions of the four luxury consumer segments proposed by Atwal and Bryson (2014). The specialist compared these specifications and descriptions with an existing database, and candidates matching these were contacted and asked to complete a list of screening questions to ensure their suitability for the study. A total of 16 candidates were identified (four per luxury segment), and the sample size was deemed sufficient to ensure saturation (Guest, Bunce & Johnson 2006). Profiles of the participants are presented in Table 1. Pseudonyms were assigned to ensure the participants' anonymity.

The interviews were conducted by the author, using a semi-structured interview guide as a basis to direct the discussions, with each interview lasting from 40 min to 1 h. The questions posed started with general questions, such as: 'Where do you like to shop? How would you describe your lifestyle?' The discussions then shifted to uncovering consumers' perceptions of luxury brands, and included questions such as: 'What luxury brands have you bought recently? Why do you buy luxury brands? How does wearing luxury brands make you feel?' All interviews were recorded and transcribed,

producing a total body of text consisting of 36 433 words. For the purpose of this study, the text produced from these interviews was analysed using DICTION.

Ethical considerations

This article followed all ethical standards for research without direct contact with human or animal subjects.

Results

DICTION analysis

The text from each of the interviews was stored individually in a Word document. As indicated in Table 2, the words captured from the individuals ranged from a minimum of 994 to a maximum of 3426.

DICTION's five master variables (Activity, Optimism, Certainty, Realism and Commonality) were used in the analysis because these master categories provide the most general understanding of a given text. The scores obtained for each participant, by DICTION master category, are shown in Table 3. These were then used as input to a hierarchical clustering procedure that is described below.

The analysis of the interview transcripts in DICTION revealed that most of the participants used language relating to Optimism (mean = 49.51) when discussing luxury brands, suggesting an overall positive feeling associated with luxury brands in this market. Commonality (mean = 49.39) was the second highest dimension, suggesting that common values are shared and associated with luxury brands. The scores for the dimensions of Realism (mean = 48.62) and Activity (mean = 48.61) were almost equal. The fifth dimension, Certainty (mean = 43.54), had a noticeably lower score than the other four master categories, indicating that participants tended not to experience a sense of resoluteness about luxury brands.

Hierarchical cluster analysis

Cluster analysis is a multivariate technique that classifies cases into groups that are relatively homogeneous within

TABLE 1: Profiles of the participants interviewed.

Pseudonym	Luxury segment	Age (years)	Gender	Race
Candice	Deluxe aspirer	28	Female	White
Walt	Deluxe aspirer	26	Male	White
Amber	Deluxe aspirer	26	Female	Black
Steve	Deluxe aspirer	27	Male	Black
Abigail	Self-made	45	Female	White
Jonas	Self-made	44	Male	White
Joyce	Self-made	34	Female	Black
Max	Self-made	36	Male	Black
Annabel	Business magnate	40	Female	White
Jacob	Business magnate	48	Male	White
Amelia	Business magnate	45	Female	Black
Duncan	Business magnate	39	Male	Black
Wendy	Money aristocracy	56	Female	White
Pierce	Money aristocracy	45	Male	White
Mary	Money aristocracy	54	Female	Black
Izak	Money aristocracy	57	Male	Black

TABLE 2: Participants and number of words per interview.

Respondent	Number of words per interview
Izak	994
Mary	1046
Joyce	3418
Wendy	2157
Annabel	3426
Abigail	2050
Candice	2564
Pierce	1550
Jacob	3060
Jonas	2134
Walt	2672
Amelia	3077
Amber	2327
Duncan	2027
Max	2317
Steve	1614

themselves, and relatively heterogeneous between each other (Landau & Chis Ster 2010). It is a statistical technique of data analysis that is commonly employed in many fields, including biology, psychology and marketing. Hierarchical cluster analysis is a specific type of clustering that combines cases into homogeneous clusters by merging them together, one at a time, in a series of sequential steps (Blei & Lafferty 2009). The analysis employed is agglomerative, and is defined by the similarity of, or measurement of the distance between, cases used and the linkages between clusters (Bratchell 1989). The similarity measure is determined by the nature of the data being considered. The data originating from DICTION are continuous-type data, which are often best handled with squared Euclidean distance. At each step in the hierarchical clustering process, clusters with the smallest Euclidean distance are joined up with each other. In terms of linkage, a number of alternative approaches are possible, but average linkage is a natural compromise between single and complete linkage, and is therefore one of the most commonly employed techniques. As the name suggests, this technique averages all distance values between pairs of cases from different clusters.

Hierarchical clustering in Statistical Package for the Social Sciences (SPSS) was employed on the input data resulting from DICTION, which consisted of a table with five columns representing the master headings for each of the 16 respondents considered – see Table 3. The squared Euclidean distance was employed as a measure of similarity, while a between-group linkage was adopted as a clustering method. No transformations were necessary, as the data resulting from DICTION consisted of z-score values.

The resulting proximity matrix, which consists of the squared Euclidean distance calculated between all cases, indicates those cases with lower scores, which are the first to be joined together, while the agglomeration schedule shows how the hierarchical table progresses. An inspection of the agglomeration table shows a sharp increase in coefficients

TABLE 3: Scores obtained for each of the participants by DICTION master category.

Participant	Activity	Optimism	Certainty	Realism	Commonality
1 Amelia	48.07	49.40	41.07	50.03	49.34
2 Amber	46.03	52.99	44.52	47.83	48.96
3 Annabel	50.82	47.38	45.94	48.49	49.31
4 Abigail	46.83	49.03	44.42	48.57	49.67
5 Duncan	47.73	47.41	47.23	47.69	49.90
6 Izak	43.78	49.76	33.53	45.57	49.79
7 Joyce	47.53	51.75	44.07	47.37	46.40
8 Jonas	50.58	48.67	45.64	50.68	50.91
9 Jacob	50.51	49.96	42.77	50.02	49.25
10 Candice	45.93	50.58	43.10	47.76	49.69
11 Mary	50.52	47.50	38.58	52.78	47.64
12 Max	49.70	48.02	47.33	50.56	50.28
13 Pierce	50.12	49.42	45.41	50.74	48.28
14 Steve	49.83	49.71	41.24	46.40	48.84
15 Wendy	50.26	50.00	45.42	45.78	49.15
16 Walt	49.44	50.61	46.32	47.71	52.81
Total scores	48.61	49.51	43.54	48.62	49.39

after stages 12 and 13. A plot of values for stages with those of the coefficient column obtained from the agglomeration schedule provides a useful scree plot. Fitting a straight line on the scree side of the plot, and observing where this exists, suggests that the scree plot turns sharply upwards after stage 12 in the agglomeration process.

Figure 1 provides a visual display of the clustering process. Viewing the dendrogram from left to right shows cases that are similar to each other and joining up with each other earlier, against those that are dissimilar; vertical lines indicate stages where clusters join up. Long horizontal lines suggest clusters that are dissimilar from each other. The results portrayed in the dendrogram, together with those from the agglomeration schedule, suggest a cut-off point indicated by the dotted line, which results in a four-cluster solution. Details of the resulting four clusters are provided in Table 4.

Cluster 1 consisted of seven participants (Jonas, Max, Pierce, Annabel, Duncan, Wendy and Walt). This cluster’s highest scores for the five categories were obtained for Commonality (mean = 50.1) and Activity (mean = 49.81), followed by Realism (mean = 48.81), Optimism (mean = 48.79) and lastly, Certainty (mean = 46.2). The participants in this cluster were predominantly white people (five white people and two black people) and male (five men and two women). The second cluster also consisted of seven participants (Abigail, Candice, Amber, Joyce, Amelia, Jacob and Steve) and provided the highest scores on Optimism (mean = 50.3), followed by Commonality (mean = 48.9), and then Realism (mean = 48.3). Activity (mean = 47.8) had the second lowest score and Certainty (mean = 43.0) the lowest. This cluster was

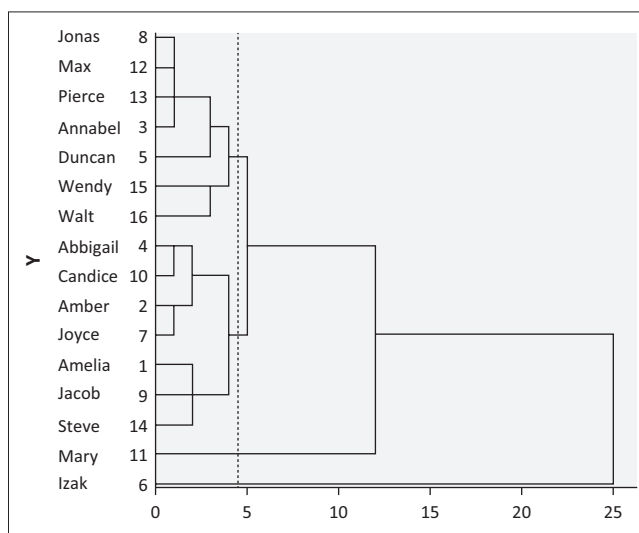


FIGURE 1: Dendrogram demonstrating a visual display of the clustering process.

TABLE 4: Means and standard deviations of DICTION meta-groups by cluster group.

Cluster	N	Activity	Optimism	Certainty	Realism	Commonality
1	7	49.8 (1.0)	48.8 (1.3)	46.2 (0.8)	48.8 (1.9)	50.1 (1.5)
2	7	47.8 (1.8)	50.3 (1.6)	43.0 (1.4)	48.3 (1.4)	48.9 (1.1)
3	1	50.5	47.5	38.6	52.8	47.6
4	1	43.8	49.8	33.5	45.6	49.8

dominated by female consumers (five women and two men), and was almost equal in terms of racial representation (four black people and three white people).

The third and fourth clusters both consisted of only one participant. Cluster 3 features participant Mary, a black female, who obtained the two highest scores for Realism (mean = 52.78) and Activity (50.52). The scores for Optimism (mean = 47.5) and Commonality (mean = 47.64) were close, while the score for Certainty was very low (mean = 38.58). Cluster 4 also consisted of a single participant, Izak, a black male, and like the individual in cluster 3, the scores obtained for Commonality (mean = 49.76) and Optimism (mean = 49.76) were close together, with Commonality being the highest. Realism (mean = 45.57) and Activity (mean = 43.78) were the third and fourth highest scores, with Certainty (mean = 33.53) having a very low score.

Discussion

The clusters identified in the results are discussed below. The researcher assigned proposed segment names for each cluster, based on the DICTION analysis results. These clusters are also described through the inclusion of quotes from the consumer interviews, to provide additional depth and richness to each cluster's description.

Cluster segment 1: Everyday luxury

The first segment's language is characterised by high Commonality and Activity in their discussions about luxury brands. The shared commonality seems to be rooted in the importance they assign to the ability of luxury brands to ensure acceptance and equality in different contexts: 'Luxury gives you a confidence boost ... and acceptance into certain circles' (Walt); and 'Wearing luxury brands means more to me in a business context' (Jonas). Activity is evident in the segment's use of different luxury brands as vehicles of flexibility that transform them to fit into any situation during their everyday lives: 'Luxury means ... I'm always ready for any occasions ... I'm flexible, I can go anywhere, I've got an outfit' (Annabel). Realism is evident in this segment from their references to using luxury to fit into the different roles presented by their everyday lives: 'If you don't have the nice gold bag, you don't want to go [to] the JHB country club ... it allows you access to a certain group of people' (Walt).

In some instances, luxury is used in their everyday lives, not only as a tool to allow them to look the part in a given situation, but also to excel in their appearance and to stand out in a specific context:

'Wherever I go ... I make it a point that I am noticed. So that to me means I want to be different to the people who are around me, in terms of what I wear.' (Max)

Optimism is seen in the enthusiasm with which luxury is discussed: 'Burberry is magnificent' (Pierce). It is also evident in the fact that luxury is endorsed for providing comfort in the actual fit of the clothing – 'I'm motivated by good quality

and that I'm comfortable' (Pierce) – and comfort, in the sense that luxury brands are of a higher quality and therefore last longer, allowing them to build stock that ensures preparedness for different situations:

'Everything that I have is good enough to handle this [any situation] ... for me that calmness is more than ... it's what money couldn't buy ... not waking up and going "Oh, what to wear, this doesn't look right ..."'. (Annabel)

'I'd rather pay a bit more ... than getting something and replace it two to three times'. (Pierce)

Even though Certainty was the lowest score overall, this segment (Everyday luxury) had the highest score for Certainty of all the segments. This is arguably because this segment wears luxury brands every day and purchases luxury brands more often than the other segments. However, low levels of Certainty are apparent in this segment's references to particular brands they have come to prefer through a process of experience with purchasing different brands: 'I know what I like and what I don't like' (Annabel); 'I'm yet to find a Polo shirt in which I felt 100% comfortable ... compared to Daniel Hechter which is virtually the same price ... every single one was comfortable' (Pierce).

Cluster segment 2: Luxury treasures

The second segment obtained the highest scores on the dimensions of Optimism and Commonality. Optimism was apparent from the discussions that revealed excitement about luxury, and about obtaining it because they have worked and saved for it: '... I'd like put it in the budget. When I'm saving up for it, I'd put there in my budget Gucci perfume, R500 this month' (Amber); 'When I was in Australia this year, I saw it and I bought it ... this specific watch I've looked at for nine years ... because it's expensive' (Jacob); '... you deserve it because you've made it ...' (Joyce). Commonality is seen in this segment's endorsing of luxury as rare, special, small treasures that are often bought as gifts for oneself: 'I call them [browsing for luxury] my mini holidays at my desk. My browsing is often online ... it's like a reward association that goes with it ...' (Joyce); '... for Christmas my ... from me to me with love ... is a pair of Tom Ford sunglasses' (Amelia).

Realism was also fairly high among this segment, but was more apparent in references made to the actual cost and price of obtaining luxury goods, and mentioning that luxury is expensive: 'I'd actually save up to get it ... I would' (Duncan); 'Quality drives me ... but I wait for the sales' (Amelia). Activity was not as high as in segment 1, and this is likely because of the segment seemingly preferring to buy luxury brands less regularly, being more careful, and opting for lasting purchases: 'I buy twice a year ... in July and Jan ...' (Amelia); 'My work shoes for example ... I have one pair ... and I've had them for three or four years ...' (Jacob); 'I think about it first ... before I buy anything ...' (Steve). Certainty is fairly low in this segment, possibly because this segment is more careful and does not buy luxury for everyday use (unlike segment 1).

Cluster segment 3: Functional luxury

The third segment consisted of only one participant who differed substantially from the other three segments, having high scores on Realism and Activity. A fairly high score on Commonality and Optimism showed similarity to the other segments, but differing again from segments 1 and 2 with a very low score on Certainty. The high score on Realism may be attributed to discussions about making functional purchases, and not necessarily being steered only by brand names:

‘The way it looks and the price motivates me, the functionality of the bag ... I’ll like compare Guess and Louise Vuitton, but I’ll take whichever is more functional for me. I’ll also compare price, but then again, if this one is more functional, I’ll buy regardless of price.’ (Mary)

The score on Activity is possibly because of the emphasis placed on how her preferences for different brands have changed over the years, and the change and growth that took place: ‘I used to be into Louise Vuitton, but I out-grew that. Right now I have a Polo bag, but I’ll also browse Guess’ (Mary).

Commonality and Optimism shared close scores. Optimism is seen in the endorsing of the positive qualities of how luxury brands make the individual feel: ‘I think [*how luxury makes me feel*] I’m one of the successful, confident people ... I do feel glamorous and sophisticated’. The very low levels of certainty could possibly be because of the fact that this participant is not as brand conscious – ‘I’m not a slave to brands ...’ (Mary) – and consults advertising as a source of education on brands and trends – ‘I’ll look at some adverts to see the trends’ (Mary) – thus not behaving like an authority on brands, but rather a realistic and functional shopper.

Cluster segment 4: Passionate luxury

The last segment also consisted of only one participant. In a similar way to the first and second segments, this segment had fairly high scores on Commonality and Optimism. Optimism was apparent in comments that suggested passion for luxury and what it symbolises – ‘I’m motivated to buy luxury because I love it ... luxury for me personally is exclusivity’ (Izak) – while Commonality, as in the first two segments, endorsed luxury for its durability and lasting properties: ‘I know when I wear it, it is quality and it is going to last’ (Izak). Realism and Activity were fairly low. The low Realism is potentially because of luxury brands being described more emotionally, personally, and with an emphasis on their symbolic value to the individual: ‘I feel glamorous, educated, important, people take me seriously ... it is part of how I market myself, my personal brand ... my kids would say, Daddy, you are always so stylish’ (Izak).

Lower scores on Activity could possibly be because of a preference for only purchasing classic luxury brands that come with a heritage, and not easily deviating from this:

‘You can’t buy a brand and not know where it comes from ... the originality ... the Gucci family, what happened to them ... the history of Giorgio Armani ... that also tells you a lot and it’s a seller as well.’ (Izak)

This segment had the lowest score on Certainty. This is interesting, because much passion and emotion concerning luxury brands is evident, but particular knowledge is supposedly lacking, which might be the cause of not always talking about luxury brands with authority:

‘Yes, you buy something and you realise you aren’t really happy. Like when I thought I bought exclusivity, thinking not all South Africans will have this and it wasn’t, I saw people wearing this ... these watches.’ (Izak)

Conclusion

This study aimed to demonstrate how some of the critique associated with text analysis programs (Schmidt 2010), as well as the challenges associated with the dilemmas in conducting research in the social sciences (McGrath 1981), may be overcome. This was achieved using the specialised text analysis software DICTION to demonstrate how both precision and richness (context) may be achieved in a research study. By analysing a corpus of text obtained from in-depth interviews with 16 luxury brand consumers in the South African market, the study identified four segments of luxury consumers. The segments were uncovered using DICTION’s textual analysis and quantified output, which was used to conduct a hierarchical cluster analysis. The analysis produced four clusters and, based on a review of each cluster’s descriptions and the interview discussions, the researchers were able to develop four unique luxury cluster segments: *Everyday luxury*, *Luxury treasures*, *Functional luxury* and *Passionate luxury*.

Of interest to managers is DICTION’s ability to provide them with swift insights into lengthy interview transcripts. The outputs produced by DICTION further allow managers to obtain quick answers to consumer behaviours through a review of their dimension scores, while the written transcripts may still be consulted to provide deeper insights into the scores produced. Unlike traditional quantitative analysis, detailed attention can be given to those participants indicated as outliers in the DICTION output by returning to their transcripts. Whether these participants can be truly regarded as exceptions in the market, or whether there may be more to these profiles, can then be determined, and a decision made on how to deal with these profiles. Furthermore, as DICTION’s output can be used to conduct a cluster analysis and to identify different market segments, managers can use DICTION to develop their own unique consumer segments in their markets of interest. Future research conducted and analysed with DICTION can then allow managers to group participants quickly according to their uniquely developed segments by simply reviewing their scores for each dimension in the output. DICTION’s ability to uncover latent aspects

of consumers' behaviour also provides managers with the opportunity to gain unique and deeper insights into the markets served.

Even though the sample size and the qualitative approach to collecting the data forgo the possibility of generalisation to the broader population (Zikmund & Babin 2007), the study demonstrates how both precision and richness (McGrath 1981) may be achieved.

A number of future research avenues are identified. A replication of this particular study on a much larger sample could be conducted to confirm the relevance of the segments identified, and to identify whether more potential segments reside in this specific market. DICTION could also be used to analyse consumer responses on more product and service categories, and researchers should also further explore its additional 35 sub-categories. If a large number of consumer responses can be collected, these responses can be analysed with DICTION, and the relationships between these scores and numerous other variables can be statistically determined.

The analysis approach undertaken in this article demonstrated the benefits of using the text analysis program DICTION and revealed how both precision and depth may be obtained while conducting research in the social sciences. By conducting this analysis in a country where luxury consumers are described as heterogeneous and complex (Atwal & Bryson 2014; Bills & Tabot 2017; Steinfield 2015), this article also contributes to the body of knowledge on luxury consumption on the African continent, and demonstrates how insights into a complex market may be obtained in a unique manner.

Acknowledgements

The author would like to thank Leyland Pitt and Albert Caruana for their help and inputs.

Competing interests

The author declares that no competing interests exist.

Authors' contributions

I declare that I am the sole author of this research article.

Funding information

Interviews collected were paid for by the researcher from the researcher's university research account.

Data availability statement

Data sharing is not applicable to this article as no new data were created or analysed in this study.

Disclaimer

The views expressed in this article are the author's own and not an official position of the institution.

References

- Arnould, E.J. & Price, L.L., 1993, 'River magic: Extraordinary experience and the extended service encounter', *Journal of Consumer Research* 20(1), 24–45. <https://doi.org/10.1086/209331>
- Atwal, G. & Bryson, D., 2014, *Luxury brands in emerging markets*, Palgrave Macmillan, London.
- Bailey, L.F., 2014, 'The origin and success of qualitative research', *International Journal of Market Research* 56(2), 167–184. <https://doi.org/10.2501/IJMR-2014-013>
- Bills, A. & Tabot, S., 2017, *Luxury brands in the South African emerging market: An empirical study of the relationship between consumer personality traits and purchase intentions towards luxury brands*, viewed 11 December 2018, from <https://iiespace.iie.ac.za/bitstream/handle/11622/277/Bills%20Ashleigh-%20Final%20Report%20%20%281%29.pdf?sequence=1&isAllowed=y>
- Blei, D. & Lafferty, J., 2009, 'Topic models', in A. Srivastava & M. Salami (eds.), *Text mining: Classification, clustering and applications*, pp. 71–94, Taylor & Francis Group, Boca Raton, FL.
- Bratchell, N., 1989, 'Cluster analysis', *Chemometrics and Intelligent Laboratory Systems* 6(2), 105–125. [https://doi.org/10.1016/0169-7439\(87\)80054-0](https://doi.org/10.1016/0169-7439(87)80054-0)
- BRICS, 2013, *BRICS joint statistical publication 2013*, viewed 02 May 2018, from <https://goo.gl/z34ypX>
- BRICS, 2015, *History of BRICS*, viewed 11 October 2018, from <https://infobrics.org/history-of-brics-2/>
- Deloitte, 2014, *Luxury brands turn to Africa as the next growth frontier – Deloitte*, viewed 22 July 2018, from <https://www2.deloitte.com/za/en/pages/consumer-business/articles/luxury-brands-powers-of-luxury.html>
- Diction, 2019, *DICTION, the in-text analysis programme*, viewed 08 January 2019, from <https://www.dictionsoftware.com/>
- Easton, H., 1940, *Word frequency dictionary*, Dover Publications, Mineola, NY.
- Euromonitor, 2016, *Luxury goods in South Africa*, viewed 27 September 2016, from <https://www.euromonitor.com/luxury-goods-in-south-africa/report>
- Euromonitor, 2019, *Luxury goods in South Africa: Analysis*, viewed 20 February 2019, from <https://www.euromonitor.com>
- Feltham-King, T. & Macleod, C., 2016, 'How content analysis may complement and extend the insights of discourse analysis: An example of research on constructions of abortion in South African newspapers 1978–2005', *International Journal of Qualitative Methods* 15(1), 1–9. <https://doi.org/10.1177/1609406915624575>
- Fine, M.A. & Kurdek, L.A., 1994, 'Publishing multiple journal articles from a single data set: Issues and recommendations', *Journal of Family Psychology* 8(4), 371.
- Gauteng Online, 2016, *The economy of Gauteng*, viewed 28 June 2016, from <https://www.gautengonline.gov.za/Business/Pages/TheEconomyofGauteng.aspx>
- Guest, G., Bunce, A. & Johnson, L., 2006, 'How many interviews are enough? An experiment with data saturation and variability', *Field Methods* 18(1), 59–82. <https://doi.org/10.1177/1525822X05279903>
- Hair, J.F., Celsi Wolfenbarger, M.W., Ortinau, D.J. & Bush, R.P., 2013, *Essentials of marketing research*, McGraw-Hill, New York.
- Krippendorff, K., 2013, *Content analysis: An introduction to its methodology*, 3rd edn., Sage Publications, Thousand Oaks, CA.
- Landau, S. & Chis Ster, I., 2010, 'Cluster analysis: Overview', in P. Peterson, E. Barker & B. McGaw (eds.), *International encyclopedia of education*, pp. 72–83, Elsevier Ltd, Oxford.
- McGrath, J.E., 1981, 'Dilemmatics: The study of research choices and dilemmas', *The American Behavioral Scientist* 25(2), 179–210. <https://doi.org/10.1177/000276428102500205>
- Pistrang, N. & Barker, C., 2012, 'Varieties of qualitative research: A pragmatic approach to selecting methods', in H. Cooper, P.M. Camic, D.L. Long, A. Panter, D. Rindskopf & K. Sher (eds.), *APA handbook of research methods in psychology: Vol. 2. Research designs*, pp. 5–18, American Psychological Association, Washington, DC.
- Schmidt, M., 2010, 'Quantification of transcripts from depth-interviews, open-ended responses and focus groups', *International Journal of Market Research* 52(4), 483–509. <https://doi.org/10.2501/S1470785309201417>
- Short, J.C. & Palmer, T.B., 2008, 'The application of DICTION to content analysis research in strategic management', *Organizational Research Methods* 11(4), 727–752. <https://doi.org/10.1177/1094428107304534>
- Sonnenberg, N.C., Erasmus, A.C. & Schreuder, A., 2014, 'Consumers' preference for eco-friendly appliances in an emerging market context', *International Journal of Consumer Studies* 38(5), 559–569. <https://doi.org/10.1111/ijcs.12120>
- Steinfield, L., 2015, *Consumer types versus stereotypes: Exploring social tensions in the luxury market of South Africa*, viewed 29 October 2018, from <https://eureka.sbs.ox.ac.uk/5364/1/2015-9.pdf>
- Stiehler, B.E. & Tinson, J.S., 2015, 'Opportunistic luxury branding: Understanding perceptions of brand authenticity in an emerging market context', *Journal of Global Business and Technology* 11(1), 39–55.
- Stiehler, B.E., 2016, 'Co-creating luxury brands in an emerging market: Exploring consumer meaning making and value creation', *Qualitative Market Research* 19(4), 395–415. <https://doi.org/10.1108/QMR-02-2016-0018>
- Sydserrf, R. & Weetman, P., 2002, 'Developments in content analysis: A transitivity index and DICTION scores', *Accounting, Auditing & Accountability Journal* 15(4), 523–545. <https://doi.org/10.1108/09513570210440586>
- Zikmund, W.G. & Babin, B.J., 2007, *Essentials of marketing research*, 3rd edn., Thomson South-Western, Natorp Boulevard Mason, Ohio.