

*The integration of netnography and text mining for the representation of
brand image in fashion blogs*

Belinda Crawford Camiciottoli
Assistant Professor
University of Florence
Department of Business Science
Via delle Pandette, 9
50127 Florence (Italy)
Tel: +39 055 4374722
Email: belinda.crawford@unifi.it

Simone Guercini
Full Professor
University of Florence
Department of Business Science
Via delle Pandette, 9
50127 Florence (Italy)
Tel: +39 055 4374704
Email: simone.guercini@unifi.it

Silvia Ranfagni
Assistant Professor of Marketing
University of Florence
Department of Business Science
Via delle Pandette, 9
50127 Florence (Italy)
Tel: +39 055 4374688
Email: silvia.ranfagni@unifi.it

Abstract

The primary objective of this study is to propose an interdisciplinary and innovative methodological approach to investigate the relations among brand, market actors and company decision-makers. The methodology is seen as an instrument that can contribute to the generation and representation of market knowledge from the viewpoint of company policy makers. Moreover, the paper illustrates an initial application to specific brand realities in the fashion industry. The innovation of the study lies in its integration of netnography with techniques of text mining to analyze perceptions of brand image in fashion blogs as a form of social media. The results are interpreted from the perspective of fashion industry decision-makers. In other words, the authors experiment a new methodology to determine if and how it may constitute a valid instrument for gaining insights into the market and for making decisions about brand management. We underline that because the aim of our paper was to develop and experiment a new methodology, we did not formulate specific research hypothesis.

Keywords: netnography, text mining, brand, fashion blogs

1. Brand associations: a synthesis of meanings in the mind of consumers

Consumer brand knowledge is a strategic resource to be analyzed, controlled and managed over time (Romaniuk and Gaillard, 2007). Companies need a clear understanding of how and why consumers evaluate brands on the basis of brand knowledge stored in their memories (Supphellen, 2000). Consumer brand knowledge can be conceptualized as a set of associations related to a brand or, more specifically, a node in consumer memory to which a variety of associative links are connected (Anderson, 1983; Keller, 1993). According to Biel (1991), the personal meanings that consumers attribute to a brand are synthesized into brand associations which are formed by the components perceived to underlie the brand's image. Keller (1993) classifies these components as attributes of the product itself (product-related), or attributes linked to the purchase and consumption of the product, such as price information, product appearance, usage and user imagery (non-product related). However, brand associations may also be less tangible, in terms of perceived benefits (Broniarczyk and Alba, 1994), product experiences, as well as feelings, thoughts and overall evaluations of brand-related information (Keller, 2003).

Brands can be linked to many brand associations as they may be perceived in different ways by consumers. Companies may in fact decide to increase brand attributes by implementing their "conceptual extension" (Park et al., 1991) independently from "merchandise extension" of products identified by the same brand (Hatch and Schultz, 2001): the same brand may represent many attributes but be applied to few products, or it may represent few attributes and be applied to many products (brand extension). At the same time, consumers recognize the brand for few or even for only one specific brand association. This may happen when companies carry out a branding strategy focalized on few or on a single attribute, assuming that "the power of a brand is inversely proportional to its scope" and that "a brand becomes stronger when you narrow its focus" (Ries and Ries, 1998, pp. 9 and 17). However, regardless of the choice of expansion versus focus of brand attributes, companies look for strong, positive and unique associations. Such associations strengthen a brand and therefore increase its equity (Broniarczyk and Gershoff, 2003; Bridges et al., 2000; Chen, 2001). In other words, brand associations determine the extent of a brand's leverage and suggest how a brand's equity can be leveraged in

the marketplace (Tauber, 1988; Aaker, 1996). For companies, the development of a direct association between a brand and specific attributes (whether exclusive to that brand or not) is a primary objective of a transfer brand image strategy (Gareth, 2004). The perception of brand uniqueness forms the basis of brand differentiation and has an influence on consumer attitude; it may produce a positive impact on consumer choices (Carpenter et al., 1994), thus contributing to maintaining and improving brand performance (Romaniuk and Gaillard, 2007). This perception has a positive influence on brand performance if it allows the brand to overcome the competition. According to Ries and Ries (1998, p. 14) “marketers often confuse the power of a brand with the sales generated by the brand. But sales are not just a function of a brand’s power. Sales are also a function of the strength or weakness of a brand’s competition”. Thus, for companies, what is important is that consumers a) perceive brand associations as distinctive characteristics compared to competitors' brands, and b) develop a brand image that corresponds to brand identity as much as possible (Aaker, 2003; Keller, 1993).

In the branding process, the brand identity defined by corporate marketers may not match consumer brand perception (brand image). More specifically, company and consumer brand associations may not correspond (Venkatraman, 1989). Companies, for example, may make positioning decisions by implementing a broad differentiation in terms of brand association, which is not necessarily perceived by consumers. In these cases, the reduction of the brand association extension through a clear and effective communication is necessary to foster the development of a more focalized brand image (Madhavaram et al., 2005). As a consequence, marketers are also forced to take drastic decisions on brand associations in order to strengthen their brand differentiation from competitors. If, however, the brand differentiation incorporated in the brand identity is not transferred in the brand image, the risks of failure may be high. The concept of matching between brand identity and brand image can be found in early research on brands (Brown, 1950; Tyler, 1957). These studies in fact emphasize that the brand itself cannot have a real consistency if it is not compared with the perception that consumers have of it (Gardner and Levy, 1955). The analysis of brand perception in terms of brand association and its comparison with the brand identity defined by company is the basis for evaluating brand performance (Keller, 2003).

In the next section, to contextualize our study, we briefly discuss the phenomenon of online consumer communities, with particular reference to the fashion industry. In the following sections, we present the innovative methodological approach on which the study is based,

incorporating elements of both netnography and text mining to identify brand associations that emerge from an online fashion community. We continue with a detailed description of the methodology, after which we present and discuss of the results of an application of the methodology, and conclude with discussion of the managerial implications of the findings.

2. The online community of fashion blogs

Our study analyses consumers in interactional social contexts based on digital platforms, among which are blogs, forums, wikis and social networks, which are now considered new marketplaces where consumers and users interact to produce and mutually exchange information (Muniz and O'Guinn, 2001; Szmigin et al., 2005). These virtual environments therefore constitute a context that generates and codifies a rich source of data (Cova 1997), reflecting a convergence of actors who may assume a variety of roles (consumers, current or potential customers, enthusiasts, experts, etc). It is thus possible to study the complex interactions of consumers with the market and, in particular, with brands and companies (De Valck, 2005).

Our analysis focuses on fashion blogs. Consumers of fashion now extensively interact by means of digital platforms (cf. Rickman and Cosenza, 2007; Boyd Thomas et al., 2007). They produce consumer knowledge in written form, facilitating in this way well-articulated expressions of brand-related perceptions among fashion consumers. As a consequence, fashion marketers have begun to recognize that trend watching and word-of-mouth monitoring of the online community are important tools for keeping up with today's fashion-conscious and fickle consumers (Kim and Jin, 2006). In fact, the analysis of an online fashion community can promote the development of the relational and distinctive power of a brand that represents a strong point of convergence between consumer and market (Fournier, 1998). In internationalization processes, fashion companies consider consumer knowledge as a strategic resource to discover market diversities and valorize brand identity. Given the important role of consumer brand knowledge and perceptions in the fashion industry, we consider the online fashion community to be an important context for brand association research. We exploit this environment to study brand associations in the consumer's view by means of an innovative methodology that integrates qualitative and quantitative techniques. This methodology can also provide managers with an accessible instrument useful for analyzing brand associations,

identifying the network of strong, favorable, and unique brand associations in consumer memory, and redefining branding strategies.

3. Analyzing brand associations: an integrated approach

Brand associations have been investigated using various techniques, ranging from focus groups to in-depth personal interviews and ethnography (cf. Henderson et al., 2002; Roedder John et al., 2006). The methodology that we propose differs from existing techniques (Henderson et al., 2002; Roedder et al., 2006); it is characterized by the absence of face-to-face contact with consumers who are observed as they interact in an online social context (in line with netnography), and by an in-depth analysis of the language they produce during these interactions (through techniques of text mining). In our view, this two-pronged approach can offer important insights into the role of brand associations in marketing decisions.

3.1 Netnography

Netnography is a qualitative research methodology proposed by Robert V. Kozinets (Kozinets, 2002). It can be characterized as ethnography on the Internet in that it adapts ethnographic techniques to analyze and understand consumer behaviors that emerge from texts produced by online communities (Kozinets, 2002). Netnography uses information available from online forums to gain insights into “tastes, desires, relevant symbol systems, and decision-making influences of particular consumers and consumer groups” (Kozinets 2002, p. 61). Like ethnography, netnography is also based on the observation of the interactive processes in a community of participants, but utilizes computer-mediated discourses rather than data collected from live encounters (Arnould and Wallendorf, 1994). However, netnography distinguishes itself from ethnography as a more unobtrusive and naturalistic research method. In consumer environments, virtual interactions are considered by users to be natural. In fact, they are not fabricated by external observers for different objectives, but emerge through spontaneous online communications (Rheingold, 1993). Moreover, with its focus on online communities, the netnographic approach actually facilitates the researcher-observer’s access to participants, while reducing time and cost factors (Kozinets, 2006). In terms of research process, netnography

normalizes criteria for specific investigative phases. It begins with an important preparatory phase (called *entrée*) to identify the online community on the basis of the research objectives, and to acquire information about its content, ranking and participants (members). This is followed by the data collection (copying textual messages exchanged among participants of online communities into analyzable files), data analysis (identifying categories or topics within the content), and finally data interpretation (Kozinets, 2006). The last step also includes conventional procedures to ensure that the research is reliable and "trustworthy" (Lincoln and Guba, 1985).

The netnographer is a discoverer of meanings linked to consumption. Netnographers do not observe people, but rather analyze the computer-mediated texts that they produce, without knowing exactly the identity of the people belonging to social groups. In this way, the researcher's role goes beyond observing and analyzing computer-mediated interaction to also recontextualize the conversational acts that emerge from the texts produced.

Netnography can be applied to different interactional contexts in which groups of consumers converge and engage on issues of common interest. These may be related to specific products or certain brands. Regarding products, some recent studies include netnographic analyses of Dutch and Flemish online food cultures (De Valcke, 2005), users of coffee (Kozinets, 2002) and open source community products (Hemetsberger and Reinhardt, 2006). Concerning brands, netnographic researchers have investigated online communities linked to the well-known television series *Star Trek* (Kozinets, 2006) and *Napster* (Giesler, 2006). Netnography is a particularly effective way to analyze brand associations in communities where conventional forms of access may be difficult (Langer, 2003; Pires et al., 2003). As noted by Kozinets (2010), netnography can be fruitfully combined with other research approaches to gain further insights into perceptions of online communities. In the following subsection, we introduce text mining as a complementary tool to netnography.

3.2 Text mining

While netnography is based on the qualitative observation of online discourse, text mining is a quantitative method used to extract information from relatively large amounts of textual data (Witten, 2005). From a disciplinary perspective, text mining falls under the over-arching area of

natural language processing (NLP) as a relatively new area of linguistic theory and practice. It utilizes instruments associated with the field of *corpus linguistics*, i.e., software applications that can extract language-related information from a collection of texts compiled in electronic form, i.e., a *corpus*, or *corpora* in the plural form. More specifically, computer applications are used to generate, analyze and manipulate texts stored in electronic form. Today text mining has numerous applications, including not only the analysis of language and linguistic trends, but also the production and updating of dictionaries, machine translation and speech recognition.

The distinguishing feature of text mining is its capacity to derive new types of information from textual sources (e.g. frequencies, semantic categories), thus going a step beyond simple information retrieval (Hearst, 1999). This is accomplished by the automatic insertion of metadata (i.e., data about data) into text files, often in the form of descriptive tags that label items according to specific criteria, such as semantic field or part-of-speech category. In this way, it is possible to discover important patterns and trends across textual data that could not otherwise be exposed. Clearly, text mining can provide a wealth of information about communicative situations of interest to researchers. For example, the perceptions of online communities towards brands can be systematically investigated with text mining procedures that analyze the language in the texts they produce.

The potential of text mining to offer new insights into consumer perceptions has begun to be tapped in some recent marketing research. In a practice-oriented article, Rickman and Cosenza (2007) briefly illustrate how text mining tools can be applied to trend forecasting by tracking ‘buzz’ (i.e. key words and phrases) in fashion weblogs. Chen’s (2009) case study of an online complaint forum targeting the credit card industry highlighted the effectiveness of text mining software in distinguishing important issues for customers and underlying reasons for their dissatisfaction.

Consumer reviews have been the focus of some important studies highlighting the quantitative capacities of text mining in marketing research. Using customer reviews of electronic products on Amazon.com, Archak, Ghose and Ipeirotis (2007, p. 56) combined text mining and econometrics to “extract actionable business intelligence from the data and better understand the consumer preferences and actions”. This approach allowed the authors to determine the importance that consumers attribute to various product features, as well as the polarity and strength of their opinions, and the implications for pricing strategies. More recently, Lee and Bradlow (forth.) implemented text mining to automatically generate and examine

product attributes for purposes of market structure analysis. Specifically, on customer review websites they exploited existing Pro-Con (i.e., favorable vs. unfavorable opinions) summary functions to extract and process phrases used by customers to assess product attributes. Text mining was proposed as a way to complement existing methods of market structure analysis, such as expert buying guides, user surveys and proprietary market research reports. However, as the authors point out, text mining is able to reveal new product attributes that cannot be detected with traditional approaches that can indicate the presence of unique submarkets.

3.3 A targeted selection of elements of netnography and text mining

Our methodological approach borrows key elements from both netnography and text mining, adapting them to the aims of this study. From text mining, we utilize software applications to extract information that helps identify potential types of brand associations that may emerge from texts produced by online communities. From netnography, we adopt the technique of observing an online community through the texts they produce, some specific research phases, and inductive coding procedures to classify brand associations. In this way, the two approaches can be seen as complementary, even if each provides an autonomous and unique contribution to the analysis. However, following Langer and Beckman (2005), we depart from Kozinets' (2010) recommendations for netnographic studies concerning the issues of informed consent and researcher participation. The online sources that we utilize can be characterized as public communication as they are freely accessible without registration. Therefore, we did not contact participants for permission to conduct research based on texts which we assume they knowingly render public. At the same time, we do not report information that could reveal their identities in order to protect their privacy. In addition, in an effort to conduct our research in the most unobtrusive way possible, we did not become members of the online community itself. Our research does not aim to gain an in-depth understanding of the culture of the fashion community, but rather to investigate consumer brand associations. In our view, this objective would neither be particularly facilitated nor enlightened by the active participation of the researchers, and could even be hindered if perceived by the online community as intrusive (cf. Hudson and Bruckman, 2004).

4. Methodology: a structured process of brand association analysis

In this section, we provide a step-by-step description of the implementation of the integrated approach presented above. More specifically, we articulated the analysis into the following phases: identification of data source, data collection and compilation, data analysis and data interpretation.

4.1 Identification of data source

The initial step was to identify a source of online data that would be appropriate for the fashion brand focus of our research. With reference to the entrée phase of netnographic research (Kozinets, 2002), we opted to access a fashion blog rather than a fashion forum. Fashion blogs contain texts written by both experts (opinion leaders) and enthusiasts (consumers), and are therefore often considerably richer than those typically found in fashion consumer forums. At the same time, fashion blogs can be seen as a community or “ecosystem” in which everyone is a “real” consumer. Among the myriad of fashion blogs found on the Internet today, we chose Style.com as the most suitable for our research goals for several reasons. First, it consistently ranks highly according to well-established criteria used to identify successful blogs, including membership, Alex traffic data, number of indexed pages and incoming links. Posts/comments on Style.com are also available for approximately three years which enabled the collection of a sufficient amount of data. In addition, an initial perusal of Style.com posts showed that they dedicate considerable space to well-known fashion brands, unlike other top-ranking fashion blogs that focus more on street fashion, celebrities or gossip. Finally, Style.com is a multi-authored blog written by both staff and guest contributors, thus better reflecting the idea of an online community with respect to some popular single-author fashion blogs.

4.2 Data collection and compilation

After identifying Style.com as the data source, we decided to focus our analysis on the brand associations of three leading Italian fashion companies: Valentino, Dolce & Gabbana and Giorgio Armani. All three are globally-recognized brands and have well-consolidated processes of internationalization. They also represent brands that are closely identified with iconic personalities and the world of luxury, thus offering the type of rich and articulated comments that have been linked to consumers of fashion (Xun and Reynolds, 2010). We then collected data from all the blog posts that contained comments about the three brands, i.e., where the brand names appeared at least once. This selective process was facilitated by the tagging and search tools provided on the Style.com website. The data was then compiled into a small fashion blog corpus made up of three subcorpora representing each individual brand. Table 1 provides an overview of the data sources used in this study.

Table 1. Data sources for the Fashion Blog Corpus

Brand	N. blog posts	N. user comments	Blog timeframe	Word count
Valentino	36	38	Aug 2008 – July 2011	8836
Dolce & Gabbana	48	103	Sept 2008 – July 2011	13,809
Giorgio Armani	41	60	Aug 2008 – July 2011	11,884

4.3 Data analysis

In an exploratory phase of analysis, we used the text-mining software *WMatrix* (Rayson, 2008) to obtain some initial clues as to what types of brand associations might be found in the Fashion Blog Corpus (hereinafter FBC). The software automatically assigns each word in the corpus (34,529 words) to a pre-established semantic fields, and then generates a tag cloud to illustrate fields that are significantly more frequent when compared to a larger normative corpus, i.e., the British National Corpus sampler of spoken English (982,712 words). Statistical significance is determined by the log-likelihood (LL) measure incorporated in *Wmatrix*. The log-likelihood value takes into account the word frequencies of the two corpora (observed values) and calculates expected values (see the Appendix for formula details).

The results of the global semantic analysis served as a launching pad for a netnographically-inspired study of the FBC in order to identify specific brand associations. Although Style.com blog posts are pre-classified under broad topics (e.g., trends, shopping

information, events and personalities in the fashion world), these were not sufficient to clearly distinguish brand associations. Therefore, we carefully read the FBC to first pinpoint each mention of the brand, and then establish categories of brand associations on the basis of contextual cues. We then manually coded the categories into corpus text files, which were later reviewed separately by the individual researchers to control for reliability. Inter-rater reliability was 88.6%, i.e., the percentage of brand associations that were coded in the same way by the researchers out of the total number identified in the corpus. Consensus was reached for the remaining percentage. In this way, the identification of brand association categories combined the ‘deductive’ knowledge from the semantic field analysis with the ‘inductive’ knowledge from netnographic coding procedure.

For additional comparative insights, we then returned to text mining tools of *WMatrix* to extract information for each brand subcorpus. In particular, we compared frequencies of the key semantic fields identified, as well as specific words that had been allocated to them.

4.4 Data interpretation

The multi-faceted analysis described above led to progressively fine-tuned results, which combined specific competencies associated with netnography and text mining. In order to properly contextualize the results and understand their implications from the perspective of company management, we interpreted them with reference company performance indicators, also as a way to suggest possible changes in branding strategies. (Levin et al., 1996; Madhavaram et al., 2005).

5. Results and Discussion

5.1. Semantic field analysis: signaling brand associations

The semantic analysis performed to suggest possible types of brand associations is illustrated in *WMatrix* screenshot reproduced in Figure 1. The cloud shows semantic fields in the FBC that have statistically higher frequencies when compared to a normative corpus of

general spoken English (see methodology section).

Figure 1. Semantic field cloud for the FBC (Source: Rayson, 2008)

To be statistically significant, the fields must have a log-likelihood LL value above 6.63 which is cut-off for 99% level of confidence. Although all the fields shown in the cloud have LL values higher than 6.63, those in larger fonts are more significant than those in smaller fonts. For example *Clothes_and_personal_belongings*, *Personal_names* and *The_Media:_TV,_Radio_and_Cinema* have respective LL values of 447.64, 441.34 and 247.73, while *Colour_and_colour_patterns*, *Long,_tall,_and_wide* and *Time:_New_and_young* have respective LL values of 19.35, 18.05 and 30.72. The higher the LL value, the more significant is the difference. As can be seen from the figure, many highly significant semantic fields can be linked to products, people and events linked to the fashion world, thus indicating a network of potential important brand associations for consumers of fashion.

5.2 Netnographic analysis: identifying brand association categories

During the manual coding of the FBC, we distinguished three categories of brand associations which emerged consistently: *product-related attributes*, *non-product related attributes*, and *designer identities*. While the first two categories clearly reflect Keller's (1993) framework, the numerous personal comments about designers suggested the need for a specific category of *designer identity*, corroborating Boyd Thomas et al.'s (2007) work which also found references to individual designers in their analysis of an online fashion forum. Table 2 provides a description of the categories, along with illustrative examples from the FBC text files.

Table 2. Brand association categories in the FBC

Categories	Description	Examples
Product-related attributes	Comments relating to the distinctive characteristics of the brand, products campaigns and collections	<ul style="list-style-type: none"> - Dolce & Gabbana is a classic brand with beautiful timeless pieces. - In a perfect world, we'd be dancing the night away in these bow-embellished Valentino pumps. - I love the Giorgio Armani Fall 2010 collection. The black beret and glasses give it that uniform look.
Non product-related attributes	Comments relating to celebrities who wear products and social events that involve/ promote the fashion house	<ul style="list-style-type: none"> - Paltrow went hard-edged in a Giorgio Armani tailored blazer and shorts suit with jet black accessories. - Expect plenty of Valentino: label heavy and jewelry designer Carlos de Souza is doing the list this year, and the house is co-sponsoring the event. - Dolce & Gabbana's massive new book, "Diamonds and Pearls," seems like a straightforward celebration of the pair's gran amor for embellishment.
Designer identity	Comments relating to the fashion designers associated with the brand	<ul style="list-style-type: none"> - Valentino is looking a tad too tan lol. - Alessandra Facchinetti, formerly of Gucci and Valentino, has found new life working on Tom Ford's womenswear. - The Gabbana half of Dolce & Gabbana hates strawberries in winter and buying fur coats in July. - The famously controlling Mr. Armani is letting go just a little.

Of particular interest is the close conceptual correspondence between the categories identified and the highly significant semantic fields that appeared in the cloud generated for the FBC. (see Figure 1). For example, *Clothes_and_personal_belongings* can be mapped onto as product-related attributes, while *The_Media:_TV,_Radio_and_Cinema* and *Personal_names* can be aligned with non product-related attributes and designer identity. This marked overlapping serves not only to reinforce the validity of the categories identified through the netnographic approach, but also to demonstrate its complementarity with other research methods (Kozinets, 2010).

5.3. Comparative analysis: exploring differences in fashion brand associations

A comparative analysis of brand association categories identified through netnographic

methods across the three brand subcorpora is shown in Table 3. The data are presented in both raw frequency counts of brand associations coded into the corresponding text files and the normalized value of instances per 1000 words (PTW). This parameter provides a more accurate profile of variation since the three subcorpora have different lengths (i.e., total word counts), as shown in Table 1.

Table 3. Distribution of brand association categories across the FBC

	Product-related attributes		Non product-related Attributes		Designer identity		Totals	
	N	PTW	N	PTW	N	PTW	N	PTW
Valentino	32	3.62	11	1.24	34 ^a 13 ^b	3.84 1.47	90	10.17
Dolce & Gabbana	80	5.79	27	1.95	20	1.44	127	9.18
Armani	50	4.20	27	2.27	37	3.11	114	9.58

a = Valentino Garavani

b = Other house designers

On general level, the table shows a substantial homogeneity in terms of the frequency of brand associations that emerge from the three fashion blogs, ranging from 9.18 to 10.17 instances PTW. This can be an indicator of a relatively high level of competition among the three brands (Punj, 2002). However, at the level of individual brands, there are some interesting differences. Valentino has a higher level of brand associations linked to designer identity (5.31 PTW), especially Valentino himself, contributing 3.84 PTW to this total. This is somewhat surprising considering the fact that the company has made a strong effort to distinguish the brand identity from the person of Valentino after he relinquished ownership (Burrese and Ranfagni 2011). Apparently, fashion bloggers still tend to associate the brand with its founding designer. Dolce & Gabbana has relatively high levels of product-related attributes (5.79 PTW) which refer to brand associations of a more tangible nature. Product-related attributes also characterize Armani brand associations (4.20 PTW). However, we also find a relatively strong designer identity (3.11 PTW). Considering that Armani also has the highest level of non-product related attributes (2.27 PTW), it would seem that consumers perceive its network of brand associations in a more balanced way in comparison with the other two brands (Broniarczyk and Gershoff, 2003).

Tables 4 and 5 provide a more fine-tuned picture of brand association differences among

the three fashion brands which emerged from additional text mining processes performed with *WMatrix*. Table 4 lists the top three semantic fields identified for each fashion brand along with frequency counts of the total number of words assigned to the field in parentheses. Table 5 lists the top five words contained in each of the top three fields.

Table 4. Distribution of top three semantic fields across the FBC

Valentino	Dolce & Gabbana	Armani
Personal names (386)	Clothes and personal belongings (614)	Clothes and personal belongings (436)
Clothes and personal belongings (158)	Substances and materials: solid (242)	Personal names (366)
Arts and crafts (65)	General appearance and physical properties (82)	Arts and crafts (90)

Table 5. Distribution of top five words within top three semantic fields across the FBC

Fashion brand/top semantic fields	Top five words within semantic fields (listed by frequency)
<u>Valentino</u>	
Personal names	Valentino, Paolo Piccioli, Maria Grazie Chiuri, Alessandra Facchinetti, Kate Moss
Clothes and personal belongings	fashion, wear, dresses, denim, couture
Arts and crafts	designer, design, draw, artist, art
<u>Dolce & Gabbana</u>	
Clothes and personal belongings	jackets, fabrics, dresses, shirts, denim
Substances and materials: solid	leather, silk, velvet, cotton, fur
General appearance and physical properties	black, colors, white, red, blue
<u>Armani</u>	
Clothes and personal belongings	fashion, dress, pants, couture, wear
Personal names	Giorgio Armani, Karl Lagerfeld, Thom Browne, Lady Gaga, Lanvin
Arts and crafts	designers, design, pictures, art, pictures

As can be seen from Table 4, the top position of *Personal names* for Valentino tends to confirm that its brand associations are closely linked to the identity of the designer himself. Similarly, the three top semantic fields of Dolce & Gabbana indicate strong tangible brand associations, while Armani's semantic fields present a more balanced profile. Moreover, from examining the top words contained in the semantic fields listed in Table 5, with respect to

Valentino and Armani the distinctly concrete dimension of Dolce & Gabbana's brand association emerges even more clearly. Also of interest in the Armani brand associations are the presence of the names of other fashion designers/brands, i.e., Lagerfeld, Browne and Lanvin. This indicates that these brands are mentioned by fashion bloggers in concomitance with Armani, which could have implications this fashion brand in terms of competition.

5.4. Performance analysis: interpreting brand association categories

As the final step, we collected data on sales revenue for the three companies to explain how trends in brand association categories can be interpreted in relation to sales revenues. We obtained data from the AIDA database which provides financial data for Italian companies and published annual reports, as illustrated in Table 4. The data available at the time of writing for all three brands covered the two-year period from 2008 to 2009 comprised in the blog timeframe (see Table 1). The figures for Valentino refer only to the Valentino brand, and not to other brands that are part of the Valentino Fashion Group (e.g. Missoni, Marlboro Classics).

Table 4. Sales revenue of the three fashion brands (in millions of euros)

Company	Sales revenue 2008	Sales revenue 2009	Delta
Valentino	260	231	-11,16%
Dolce & Gabbana	953	909	-4.62%
Giorgio Armani	1,620	1,518	-6.30%

The table shows that all three brands suffered losses in sales revenue when comparing 2009 with 2008. One obvious explanation for this result could be the global financial crisis that occurred during the blog timeframe. In fact, a study based on a sample of Italian high-end fashion brands reported an average decline in sales revenues of 5.3% from 2008 to 2009. However, the differences in the delta values of the three brands could also point to other reasons. As can be seen, Valentino experienced the highest percentage of loss (-11.6) compared to Giorgio Armani (-6.30) and especially to Dolce & Gabbana (-4.62). It is interesting to note that Valentino brand associations were mostly concentrated in identity of the designer himself (see Table 3), while Dolce & Gabbana and Armani had higher frequencies of brand associations linked to more tangible product-related attributes. This could be related to bloggers' continuing

intense interest in the designer himself, despite of the company's current efforts to transition towards an identity that is less dependent on its founder (Burresi and Ranfagni, 2011). Thus, by comparing delta values and brand association categories a company in fact may discover a low correspondence with intended brand knowledge and intervene accordingly.

To summarize, the three types of brand associations linked to *product attributes*, *non product-related attributes* and *designer identity* were consistently present across the three fashion brands investigated, suggesting that they form a 'core' of consumer brand knowledge of luxury fashion brands. However, when these brand associations are primarily tangible in nature, brand performance may be enhanced. Conversely, a prevalence of less tangible brand associations may be related to lower performance levels.

6. Conclusions and managerial implications

This paper offers a new methodological approach to understanding consumer brand knowledge by identifying brand associations. It combines techniques of qualitative netnography and quantitative text mining to analyze texts produced by a community of online consumers. In this sense, we may even characterize our approach as a sort of 'online textography', borrowing the concept of "textography" (Swales, 1998) as a methodology that merges in-depth textual analysis and ethnography, which we then extend to online text sources.

This approach has important managerial implications. First of all, it investigates brand associations by exploiting information on an online community without directly involving the consumer or using complex mathematical and statistical techniques (Roedder John et al., 2006). In other words, it provides managers with a more accessible method for analyzing brand associations to better understand how they are stored in the consumer's memory. It is important to point out that the application of this method requires a strong integration between language skills and managerial skills. At the same time, the quantitative and qualitative dimensions of our approach can be articulated into different levels of complexity, thus allowing managers to choose the desired degree of depth for analyzing brand associations. The approach may orient companies in the definition of the differentiating power of their brand; more specifically it offers analytical tools to: a) determine whether the categories of brand associations in the mind of the

consumer correspond to the categories of attributes that define the company's competitive positioning, b) reveal possible new extensions of brand associations represented by novel categories identified in the online communities, and c) analyze the impact that new categories of brand associations (as expressed in specific marketing actions, such as new forms of communication, distribution and sponsorship) have on existing brand representations (Sjödin and Törn, 2006). With regard to point b), we underline the usefulness of our method to better understand existing brand associations or identify new ones, also to decide how to transfer specific attributes from an existing brand to new types of products (Park et al., 1991; Monga and John, 2010). Our approach also has the advantage of allowing managers to perform a combined analysis of categories of brand associations between competitors. This process is facilitated by using easily accessible information at relatively low costs, which can then be analyzed with targeted research methods. An analysis of competition can be achieved by evaluating the distribution of brand associations across categories. In this study, such a distribution highlighted some interesting differences among the three brands in terms of category frequency. The methodology therefore enabled not only the identification of types of brand associations in the online fashion community, but also how different types of brand associations may vary in relation to different brands. The methodology can acquire a more prescriptive and normative value for a company if the analysis of brand associations is combined with indicators of business performance (e.g., sales, profitability, market share). This would delineate a system of interactions between these indicators and categories of brand associations that could direct branding strategies.

While our study has contributed key insights into the brand associations of an online community, it is not without limitations. It is based on relatively small corpora that correspond to three fashion brands, and thus the findings are not broadly generalizable. However, large amounts of data that target specific brands from both consumer and corporate sources are currently not readily accessible. It is also important to recognize that detailed qualitative analysis (a key component of this study) is only possible with relatively small corpora. Thus, we see the value of this approach in its capacity to signal important brand association trends of interest to marketers. It could be further strengthened if we could engage companies, perhaps in the context of personal interviews, in forms of collaboration to interpret the results and to understand how to use them.

In the next step of our research, we aim to develop our methodological approach by exploiting additional text mining techniques to gain further insights into the brand associations of online communities. The idea is to expand the analysis to include not only data from fashion blogs, but also from company communications (Ross and Harradine, 2011). In this way, we will be able to compile parallel corpora in order to identify brand descriptors (Plummer, 1985) within brand associations and measure to what extent they correspond. The final objective is to define a synthetic indicator of the degree of correspondence (i.e. match or mismatch) between consumer and company-defined brand associations. This can provide managers with an accessible method to enhance their knowledge of consumer brand associations, and especially to understand how they emerge from their online communications.

References

Aaker, David A. (1996), *Building Strong Brands*. New York, The Free Press.

Aaker, David (2003), The Power of Branded Differentiator, *MIT Sloan Management Review*, Fall:83-87.

Amould, Eric J. and Melanie Wallendorf (1994), Market-Oriented Ethnography: Interpretation Building and Marketing Strategy Formulation, *Journal of Marketing Research*, 31 (November), 484-504.

Anderson, John R. (1983), *The Architecture of Cognition*, Cambridge, MA: Harvard University Press.

Archak, Nikolay, Anindya Ghose, and Panagiotis G. Ipeirotis (2007), Show Me the Money!: Deriving the Pricing Power of Product Features by Mining Consumer Reviews, in Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (San Jose, California, USA: ACM, 2007), 56-65.

Biel, A.L. (1991), *The Brandscape. Converting Image into Equity*, Admap, pp.41-46.

Boyd Thomas, Jane, Cara Okleshen Peters and Holly Tolson (2007), An Exploratory Investigation of the Virtual Community MySpace.Com. What Are Consumers Saying about Fashion?, *Journal of Fashion Marketing and Management*, 11, 4, 587-603.

Bridges, Sheri, Kevin Lane Keller, and Sanjay Sood (2000), Communication Strategies for Brand Extensions: Enhancing Perceived Fit by Establishing Explanatory Links, *Journal of Advertising*, XXIX, 4 (Winter), 1-11.

Broniarczyk, Susan M. and Andrew D. Gershoff (2003), The Reciprocal Effect on Brand Equity and Trivial Attributes, *Journal of Marketing Research*, XL (May), 161-175.

Broniarczyk, Susan M. and Joseph W. Alba (1994), The Importance of Brand in Brand Extension, *Journal of Marketing Research*, 23, (May), 214-228.

Brown, William F. (1950), The Determination of Factors Influencing Brand Choice, *Journal of Marketing*, 14, (5), 699-706.

Burresi, Aldo and Silvia Ranfagni (2011), Congruity Brand Image and Multichannel distribution in Fashion Companies, Proceedings of the 10th International Marketing Trends Conference (Paris, January 20-22).

Carpenter, Gregory S., Rashi Glazer, and Kent Nakamoto (1994), Meaningful Brands From Meaningless Differentiation: The Dependence on Irrelevant Attributes, *Journal of Marketing Research*, 31(August): 339-350.

Chen, Cheng-Hsui (2001), Using Free Association to Examine the Relationship Between the Characteristics of Brand Associations and Brand Equity, *Journal of Product & Brand Management*, 10 (7), 439 – 451.

Chen, Kuan C. (2009), Text Mining E-Complaints Data From E-Auction Store with Implications for Internet Marketing Research, *Journal of Business & Economics Research*, 7,5 (May), 15-24.

Cova, Bernard (1997), Community and Consumption: Towards a Definition of the “Linking Value” of Product or Services, *European Journal of Marketing*, 31,3/4, 297 – 316.

De Valck, Kristine (2005), *Virtual Communities of Consumption: Networks of Consumer Knowledge and Companionship*, ERIM PhD Series: Research in Management.

Fournier, Susan. (1998), Consumers and their Brands: Developing Relationship Theory in Consumer Research, *Journal of Consumer Research*, 24, (March), 343-373.

Gardner Burleigh B. and Sidner J. Levy (1955), The product and the Brand, *Harvard Business Review*, March-April, pp. 33-39

Gareth, Smith (2004), Brand Image Transfer Through Sponsorship: A Consumer Learning Perspective, *Journal of Marketing Management*, 20, 457-474.

Giesler, Markus (2006), Consumer Gift Systems: Insights from Napster. *Journal of Consumer Research*, 33, 2 , 283-90.

Hatch, Mary J. and Majken Schultz (2001), Are the strategic stars aligned for your corporate

brand?, *Harvard Business Review*, (February), 128-134.

Hearst, Marti A. (1999), Untangling Text Data Mining, in the Proceedings of ACL'99: the 37th Annual Meeting of the Association for Computational Linguistics, University of Maryland, June 20-26, 1999. <http://people.ischool.berkeley.edu/~hearst/papers/acl99.pdf>. (Accessed July 9, 2011).

Hemetsberger, Andrea, and Christian Reinhardt (2006), Learning and Knowledge-Building in Open-Source Communities: A Social-Experiential Approach, *Management Learning*, 37, 2, 187-214.

Henderson, Geraldine R., Dawn Iacobucci and Bobby J. Calder (2002), Brand Diagnostics: Mapping Branding Effect Using Consumer Associative Networks, *European Journal of Operational Research*, 111 (December), 306-327.

Hudson, James M. and Amy Bruckman (2004), 'Go Away' Participants Objections to Being Studied and the Ethics of Chatroom Research, *The Information Society*, 20, 2, 127-139.

Keller, Kevin Lane (1993), Conceptualizing, Measuring and Managing Customer-Based Brand Equity, *Journal of Marketing*, 57 (January), 1-22.

Keller, Kevin Lane (2003), Brand Synthesis: The Multidimensionality of Brand Knowledge. *The Journal of Consumer Research*, 29, 4 (March), 595-600.

Kim, Hye-Shin and ByoungHo Jin (2006), Exploratory Study of Virtual Communities of Apparel Retailers, *Journal of Fashion Marketing and Management*, 10,1:41-55.

Kozinets, Robert V. (2002), The Field Behind the Screen: Using Netnography for Marketing Research in Online Communities, *Journal of Marketing Research*, XXXIX (February): 61-72.

Kozinets, Robert V. (2006), Click to Connect: Netnography and Tribal Advertising. *Journal of Advertising Research*, September: 279-288.

Kozinets, Robert V. (2010), *Netnography. Doing Ethnographic Research Online*. Thousand Oaks Sage Publications.

Langer, Robert. (2003), Skin two: (un)-covering the skin in fetish carnival, Proceedings of the European Association for Consumer Research Conference (EACR), Dublin, 5-7 June.

Langer, Roy and Suzanne C. Beckman (2005), Sensitive Research Topics: Netnography Revisited, *Qualitative Market Research*, 8, 2, 189-203.

Lee, Thomas Y. and Eric T. Bradlow (forthcoming), Automated Marketing Research Using Online Customer Reviews, *Journal of Marketing Research*.

Levin, Aron, M. J. Charlene Davis, Irwin Levin (1996), Theoretical and Empirical Linkages Between Consumers' Responses to Different Branding Strategies, *Advances in Consumer Research*, 23, 296-300.

Lincoln, Yvonna S. and Egon Guba G. (1985), *Naturalistic Inquiry*. Beverly Hills, CA: Sage Publications, Inc.

Madhavaram, Sreedhar, Vishag Badrinarayanan, and Robert E. McDonald (2005), Integrated Marketing Communication (IMC) and Brand Identity as Critical Components of Brand Equity Strategy, *Journal of Advertising*, 3,4 (Winter), 69-80.

Monga, Alokparna Basu and Deborah Roedder John (2010), What Makes Brands Elastic? The Influence of Brand Concept and Styles of Thinking on Brand Extension Evaluation, *Journal of Marketing*, 74, 80-92

Muniz, Albert M., Jr., and Thomas C. O'Guinn (2001), Brand Community. *Journal of Consumer Research*, 27 (March), 412-32.

Park, C.Whan, Milberg, Sandra and Robert Lawson (1991), Evaluation of Brand Extensions: The Role of Product Feature Similarity and Brand Concept Consistency, *Journal of Consumer Research*, 18, (2), 185- 193.

Pires, Guilherme D., John Stanton, and Bruce Cheek (2003), Identifying and reaching an ethnic market: methodological issues, *Qualitative Market Research*, 6,4, 224-35.

Plummer, Joseph T. (1985), How Personality Makes a Difference, *Journal of Advertising Research*, 24, 6, 27-31.

Punj, Girish and Junyeon Moon (2002), Positioning options for achieving brand association: a psychological categorization framework, *Journal of Business Research*, 55, 4, 275-283.

Rayson, Paul (2008), From key words to key semantic domains, *International Journal of*

Corpus Linguistics, 13, 4, 519-549.

Rheingold, Howard (1993), *The Virtual Community: Homesteading on the Electronic Frontier*. Reading, MA: Addison Wesley.

Rickman, Tracy Anna and Robert M. Cosenza. (2007), The Changing Digital Dynamics of Multichannel Marketing. The Feasibility for the Weblog: Text Mining Approach for Fast Fashion Trending, *Journal of Fashion Marketing and Management*, 11, 4: 606-621.

Ries, Al and Laura Ries (1998), "The 22 Immutable Laws of Branding", Harper Collins Publishers, London

Roedder, Deborah, J., Barbara Loken, Kim Kyeongheui and Monga Basu Alokparna. (2006), Brand Concept Maps: A Methodology for Identifying Brand Association Network, *Journal of Marketing Research*, XLIII (November), 549-563.

Romaniuk, Jenni and Elise Gaillard (2007), The Relationship Between Unique Brand Associations, Brand Usage and Brand Performance: Analysis Across Eight Categories, *Journal of Marketing Management*, 23,4, 267-284.

Ross, Jill and Rod Harradine (2011), Fashion Value Brands: The Relationship Between Identity and Image, *Journal of Fashion Marketing and Management*, 15, 3,306 – 325.

Sjödin, Henrik and Fredrik Törn (2006), When Communication Challenges Brand Associations: a Framework for Understanding Consumer Responses to Brand Image Incongruity, *Journal of Consumer Behaviour*, (5), 32-42.

Supphellen, Steinar (2000), Understanding Core Brand Equity: Guidelines for in Depth Elicitation of Brand Associations, *International Journal of Market Research*, 3,3: 319-338.

Szmigin, Isabelle, Louise Canning and Alexander E. Reppel (2005), Online community: enhancing the relationship marketing concept through customer bonding, *International Journal of Service Industry Management*, 16, 5, 480 – 496.

Swales, John M. (1998), *Other Floors, Other Voices. A Textography of a Small University Building*. New Jersey. Lawrence Erlbaum Associates, Inc., Publishers.

Tauber, Edward M. (1988), *Brand Leverage: Strategy for Growth in a Cost Controlled World*.

Journal of Advertising Research, 28, (August-September), 26-30.

Tyler, William D. (1957), The Image, the Brand, and the Consumer, *Journal of Marketing*, 22, 2, 162-165

Venkatraman, Venkat N. (1989), The Concept of Fit in Strategy Research: Toward Verbal and Statistical Correspondence, *The Academy of Management Review*, 14,3 (July), 423-444.

Witten, Ian H. (2005), Text mining, in *Practical Handbook of Internet Computing* Munindar Paul Singh, ed. Boca Raton: Chapman & Hall/CRC Press, 14-1-14-22.

Xun, Jiyao and Jonathan Reynolds (2010), Applying Netnography to Market Research: The Case of the Online Forum, *Journal of Targeting, Measurement and Analysis for Marketing*, 18, 1,17-31.

Appendix

Log-likelihood is calculated according the following formula:

Legend. N: number of words in the datasets, O: the observed values, E: the expected values.
More explanation about the calculation of log-likelihood can be found at: <http://ucrel.lancs.ac.uk/llwizard.html>