Measuring Asymmetries in Brand Associations Using Correspondence Analysis

Anna Torres & Michael Greenacre

Universitat Pompeu Fabra

Anna Torres' current address is Department of Business, Universidad Carlos III, calle Madrid 126, 28903 Getafe, Madrid.

E-mails: anna.torres@econ.upf.es, michael@upf.es.

1

Abstract

Correspondence analysis is introduced in the brand association literature as an alternative tool to measure dominance, for the particular case of free choice data. The method is also used to analyse differences, or asymmetries, between brand-attribute associations where attributes are associated with evoked brands, and brand-attribute associations where brands are associated with the attributes. An application to a sample of deodorants is used to illustrate the proposed methodology.

1 Introduction

This paper is motivated by a marketing research study of brand images for a set of deodorants and the hypothesis that the way of formulating the study may generate changes in the apparent brand images that result.

Our interest is related to the brand imaging problem but focussed specifically on capturing and measuring a possible asymmetry in the results of two analyses. These alternatives do not differ in the techniques used, but in the methods of data collection. These are called "attribute-by-attribute" and "brand-by-brand" ways of collecting data. While in the "brand-by-brand" method people are asked to make associations between evoked brands and listed attributes, in the "attribute-by-attribute" method the attributes are presented one at a time and people associate them with evoked brands. These alternatives have already been compared by Barnard & Ehrenberg (1990), but focusing on different ways of measuring associations: free choice, equivalent to evoked brands analysis, scaling and ranking (the last two considered forced-choice). The concepts required to understand the duality existing in the brand association, in other words, the orientation of brand association (brand to attribute or attribute to brand) appear in Farquhar & Herr (1992).

We propose correspondence analysis (CA) (Benzécri et al., 1973, Greenacre, 1984) as an alternative, in *naming methods* (also called *sorting, free choice* in Joyce, 1963), to measure *instance dominance* and *attribute dominance* (Farquhar & Herr, 1992) between a couple composed of a brand and an attribute.

CA can also be used to determine if asymmetry exists between the two mental processes for this particular study. By asymmetry we mean differences between the brand-by-brand and the attribute-by-attribute way of thinking and establishing connections. We will propose a particular way of coding data in CA (Greenacre, 2000) to describe asymmetries. Joyce (1963) showed that the free choice data, the attribute-by-attribute method was the best able to discriminate between similar brands. We will try to see if these results are also corroborated in our case.

Although Farquhar & Herr (1992) indicated that "*a brand may be associated with specific product categories as well as with specific product attributes, related customer benefits, and various summary evaluations*", most of the literature refers only to product categories and so to *brand leveraging* activities. The reason can be that the study of asymmetires in this case is richer since it lets to establish predictions on buying process as well as possibilities and limitations in the stretchability of and established brand (Farquhar & Herr, 1992 and Farquhar, 1989). Nevertheless, the application we are going to do for the brand-attribute tables

3

can be extended to the brand-category product case.

2   Brand-attribute association

There exist different types of brand association such as association with product attributes, customer benefits, relative price, the uses or applications, association with particular clients, different celebrities, life styles, categories of product or countries (Aaker, 1991). A problem can arise when researchers neglect the directionality of brand associations. Brand management involves two activities, which are complementary in deciding about the limits of a brand's stretch: brand building and brand leveraging. On the one hand, *brand building* activities focus on establishing favourable attitudes and strengthening the relationship from the brand to a particular category, product attribute, customer benefit, or usage situation (Farquhar and Herr, 1993). On the other hand, *brand leveraging* activities must consider the strength of existing associations directed towards the brand (Farquhar, Herr & Fazio, 1990 and Farquhar, Han, Herr, & Ijiri, 1992). Different reasons such as the relatively high cost of launching and building new brands, the unavailability of satisfactory trademarks in some markets and the strong competition for distribution within the trade provoke an interest in the creation of new associations instead of generating a new brand (Farquhar & Herr, 1992). The point is that the strengths of the existing directional associations can limit the stretchability of the brand. For example, Farquhar, Han, Herr & Ijiri (1992) point out the fact that strong associations between a brand and a product category can make it risky to extend a *master brand* directly to other product categories (e.g. the risk of diluting the core associations and eroding the customer base). We are interested in measuring the directional associations, brand-to-attribute and attribute-to-brand, where the strengths of the directional relations can be asymmetric (Farquhar & Herr, 1992).

When measuring brand-attribute association, there exist different definitions of *dominance* (Ashcraft, 1978; Barsalou, 1983; Farquhar, Herr & Fazio, 1990; Loftus, 1973; Mervis & Rosch, 1981). Farquhar & Herr (1992) cites Loftus (1973), using her definitions of instance dominance and category dominance, but inverting the definitions for no apparent reason. For example, Farquhar & Herr (1992) use the term "instance dominance" to define what Loftus (1973) called "category dominance". Both agree that dominance refers to the strength of the directional association between a brand and an associate, in our case an attribute. We prefer the original usage of the terms by Loftus (1973), that is we shall use "attribute dominance" to refer to the strength of the directional association from a brand to an attribute, and "brand dominance" as the

4

strength of the directional association from an attribute to a brand. We shall measure attribute dominance by the number of people who gave the attribute in response to the brand and, in a parallel way, brand dominance by the number of people who gave the brand in response to the attribute, with appropriate adjustments for size in order to normalize the measures. Further, high and low instance (i.e., brand) or category (i.e., attribute) dominance fixing a particular proportion. For example, Loftus (1973) established that *"stimuli were classified as high instance dominance if at least 70% of the sample gave the instance in response to the category and low instance dominance if fewer than 26% gave the instance in response to the category"*. On the other hand, Ashcraft (1978) defined that the *"high dominant property had a frequency value of at least 50%, and the low dominance property had a frequency of less than 50%, where frequency rate was defined as the number of subjects generating the property divided by the number of subjects generating the most frequent property for that category member"*. As was noted by Farquhar & Herr (1992) directional brand associations would not be of interest if the strengths of the two relationship were equal, in other words, if asymmetry did not exist.

3 Data

The original data sets were two tables of frequencies generated from two different samples. The sample size in each case was approximately equal to 200.

The format of the tables is the same in both cases. The rows are 10 brands of deodorants and the columns 11 attributes or benefits of using them. Given the confidentiality, we are not able to show the brand names of the deodorants. Instead, we will denote them by $B1$ until $B10$. All the attributes are *positive* versions (Barnard & Ehrenbert, 1990), included in one of the categories from the list of "key operational attributes", given by Joyce (1963). The attributes are the following:

$A1$: prevents body odour all day,
$A2$: keeps me dry all day,
$A3$: does not irritate my skin,
$A4$: has a long lasting fragance,
$A5$: has a pleasant fragance,
$A6$: leaves me feeling confident,
$A7$: leaves no marks on my clothes,
$A8$: can use all over the body,
$A9$: is portable/can carry around,
$A10$: is quick to apply
$A11$: costs a little less than most others.

Respondents were asked a series of questions about the brands, such as "most often used brand", "a brand one would switch to", "a brand one would not use again", and "brands used nowadays". Any mention of, or recognition of, a brand would include that brand amongst a respondent's personal list of "evoked brands".

The original tables were composed of frequencies representing the number of people who made a positive association between a brand and an attribute. To obtain the first table, called "brand-by-brand" (B), for each one of the evoked brands, the subjects had to associate or not the list of the attributes. To get the second one, named "attribute-by-attribute" (A), for each attribute, they stated whether it applied to each of the brands they had evoked. Barnard & Ehrenberg (1990) already spoke about "free-choice" questioning (equivalent to different evoked brands for each subject) as a *"better option with respect to "forced-choice" since respondents are not forced to make a response for each brand-attribute pairing. For respondents unfamiliar with a brand, the "forced-choice" procedure can give rise to errors due to, say, guessing or yea-saying"* (Hughes, 1969; Morrison, 1979). The tables are the following:

6

Table 1: Brand-by-brand (B)

	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	Total evoked
B1	53	49	61	40	49	38	47	27	36	41	11	(101)
B2	36	20	34	61	72	36	41	62	49	47	13	(104)
B3	41	32	43	20	46	24	30	12	44	56	30	(113)
B4	34	29	33	17	31	18	18	15	13	36	7	(76)
B5	43	40	35	23	28	37	29	14	18	45	18	(93)
B6	52	40	56	39	63	46	40	28	31	54	21	(116)
B7	96	96	81	58	81	75	58	36	51	80	26	(155)
B8	21	25	26	10	24	15	16	12	13	24	3	(55)
B9	26	38	31	22	32	21	21	8	24	32	9	(60)
B10	11	5	14	7	11	5	7	5	10	6	1	(20)

For example, the cell ($B4$, $A1$) tells us that 34 subjects out of the 76 who evoked brand $B4$ made a positive association with respect to attribute $A1$: *Prevents body odour all day.*

Table 2: Attribute-by-attribute (A)

	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	Total evoked
B1	51	44	59	36	54	41	36	37	58	62	14	(98)
B2	26	23	33	50	52	31	33	63	54	57	17	(91)
B3	33	28	40	24	41	30	33	23	60	49	24	(106)
B4	28	23	28	19	24	22	19	15	24	36	5	(63)
B5	36	32	32	19	22	29	21	18	32	50	7	(85)
B6	28	26	43	29	35	26	26	26	41	54	22	(94)
B7	83	76	79	51	68	74	48	40	70	96	29	(152)
B8	25	26	31	19	20	23	17	19	35	35	13	(61)
B9	21	20	29	15	24	20	22	8	34	28	4	(56)
B10	5	5	10	4	8	4	7	5	10	10	4	(22)

In this case, for the couple composed of $B4$ and $A1$, 28 subjects made the association, out of a total evoked of 63 (notice that in this case, the total evoked for $B4$ is lower than in the brand-by-brand analysis).

4 Preliminary description of dominance and asymmetries.

To describe dominance relations, Loftus (1973), Ashcraft (1978) and Farquhar & Herr (1992), measure the strength of association simply by the frequency with which the item was mentioned, with normalizations that do not include information related with evoked brands. For example,

Ashcraft (1978) defined frequency as *"the number of subjects generating the property divided by the number of subjects generating the most frequent property for that category member"*. For example, if we take $A3$ for the attribute-by-attribute table (A), the original values are

$$[59, 33, 40, 28, 32, 43, 79, 31, 29, 10].$$

Since the highest value corresponds to $B7$ we should take this value as reference point and transform the data as percentages referred to it:

$$[74.7, 41.8, 50.7, 35.4, 40.5, 54.4, 100, 39.2, 36.7, 12.7],$$

High dominant property, according to this particular author, are given by those with values higher than 50%. Thus for $A3$ we observe high dominant property with respect to $B1, B3, B6$ and $B7$.

The previous definition of dominance, as happens with other definitions, do not take into account the "halo effect". This is when some brands, just because they are more well known, are more evoked compared to the others, receiving higher frequencies of association. This effect is also described in Barnard & Ehrenberg (1990) as "usage effect". They note that it appears in the "free choice" way of collecting data. They say that *"many people tend to associate a positive attribute with larger brands than associate it with smaller brands. The explanation is that a larger brand has more claimed users than a smaller brand, and that users of a brand are more likely than nonusers to give a positive attribute response"*. The consequence is that these larger or more well known brands will always receive the higher values and so, they will be establish the dominant relations. Since we have information on evoked brands, we can improve the definition of dominance proposing to analyze percentages calculated from frequencies with respect to the evoked brands and then determine dominance from these percentages. The new tables allow us to treat each brand as equal and so to measure brand dominance independently if they are more or less well known.

The following table shows percentages referred to the brand-by-brand (B) original data with respect to the evoked brands:

Table 3: Percentages relative to evoked brands for brand-by-brand table(B)

	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11
B1	52.5	48.5	60.4	39.6	48.5	37.6	46.5	26.7	35.6	40.6	10.9
B2	34.6	19.2	32.7	58.7	69.2	34.6	39.4	59.6	47.1	45.2	12.5
B3	36.3	28.3	38.1	17.7	40.1	21.2	26.5	10.6	38.9	49.6	26.5
B4	44.7	38.2	43.4	22.4	40.8	23.7	23.7	19.7	17.1	47.4	9.2
B5	46.2	43.0	37.6	24.7	30.1	39.8	31.2	15.1	19.4	48.4	19.4
B6	44.8	34.5	48.3	33.6	54.3	39.7	34.5	24.1	26.7	46.6	18.1
B7	61.9	61.9	52.3	37.4	52.3	48.4	37.4	23.2	32.9	51.6	16.8
B8	38.2	45.5	47.3	18.2	43.6	27.3	29.1	21.8	23.6	43.6	5.5
B9	43.3	63.3	51.6	36.7	53.3	35.0	35.0	13.3	40.0	53.3	15.0
B10	55.0	25.0	70.0	35.0	55.0	25.0	35.0	25.0	50.0	30.0	5.0

For example, the cell ($B4, A1$) tells us that 44.7% of the subjects who evoked brand $B4$ associated it with attribute $A1$.

The attribute-by-attribute table, similarly expressed relative to evoked brands, is as follows:

Table 4: Percentages relative to evoked brands for attribute-by-attribute table (A)

	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11
B1	52.0	44.9	60.2	36.7	55.1	41.8	36.7	37.8	59.2	62.3	14.3
B2	28.6	25.3	36.2	54.9	57.1	34.1	36.3	69.2	59.3	62.6	18.7
B3	31.1	26.4	37.7	22.6	38.7	28.3	31.1	21.7	56.6	46.2	22.6
B4	44.4	36.5	44.4	30.2	38.1	34.9	30.2	23.8	38.1	57.1	7.9
B5	42.3	37.6	37.6	22.4	25.9	34.1	24.7	21.2	37.6	58.8	8.2
B6	29.8	27.7	45.7	30.9	37.2	27.7	27.7	27.7	43.6	57.4	23.4
B7	54.6	50.0	52.0	33.6	44.7	48.7	31.6	26.3	46.1	63.2	19.1
B8	41.0	42.6	50.8	31.1	32.8	37.7	27.9	31.1	57.4	57.4	21.3
B9	37.5	35.7	51.8	26.8	42.9	35.7	39.3	14.3	60.7	50.0	7.1
B10	22.7	22.7	45.5	18.2	36.4	18.2	31.8	22.7	45.5	45.5	18.2

For example, the cell ($B4, A1$) tells us that 44.4 of the subjects who evoked brand $B4$ named it in response to $A1$.

For attribute $A3$ there exists brand dominance in the cases of $B1$, $B7$, $B8$ and $B9$ (using 50% as the cut-off point). Notice how the brand dominance changes when the halo effect is eliminated. Brands $B3$ and $B6$ were all highly evoked brands. $B8$ and $B9$ are not highly evoked brands but have high dominance when associations are expressed to the evoked brands.

We think that percentages relative to evoked brands express information which is more correct in terms of determining brand dominance instead of total frequencies, which suffer from the halo effect.

We now give some other examples to motivate the idea of asymmetry. If we establish that brand dominance exists for percentages bigger than 50%, taking the cell $(B1, A3)$ in the attribute-by-attribute analysis (table 4), we find a value of 60.2%, indicating that a dominance relation between the pair exists. For the same pair $(B1, A3)$ in the brand-by-brand analysis (figure 3) the percentage is very similar (60.4%), so in this case asymmetry does not exist. But if we take another pair such as $(A1, B10)$, in the brand-by-brand case, the proportion of people who made the association is equal to 55.0%, while for the attribute-by-attribute study it takes a value of 22.7%. These percentages are quite different, one being more than double the other. So in this case asymmetry in the directional associations exists.

Remember that all percentages are referred to positive versions of the attributes. In other words we capture, for each brand-atttribute pair, the number of people who evoked the brand and established a positive association between the brand and the attribute. We can also show the counterpart, which will collect the people who evoked the brand but did not establish the positive association. In terms of the table, we will have a positive and a negative version for each attribute. For example, in the brand-by-brand study, 52.5% of the people who evoked $B1$ associated it with $A1_+$: *prevents body odour all day*. This means that 47.5% evoked it but did not make the association, in other words, they consider that brand $B1$ is related with $A1_-$: *does not prevent body odour all day*. Later in correspondence analysis we shall take both the positive and negative versions of the attributes into account.

The question arises what our definition of dominance might be if we did not have the information on evoked brands. If the halo effect is, as Barnard & Ehrenbert (1990) claim, such that *"users of a brand are more likely to give a positive attribute response"*, we might consider expressing the observed frequencies of positive attribute responses relative to their totals across the attributes. This total can be considered as a surrogate for the "total evoked". Figures 1 and 2 show the relationship between total responses and total evoked brands for matrices B and A respectively, and it is clear that there is a strong linear relationship between these two totals.

Figure 1 : Total number of associations versus total evoked brands for the brand-by-brand data (table B).
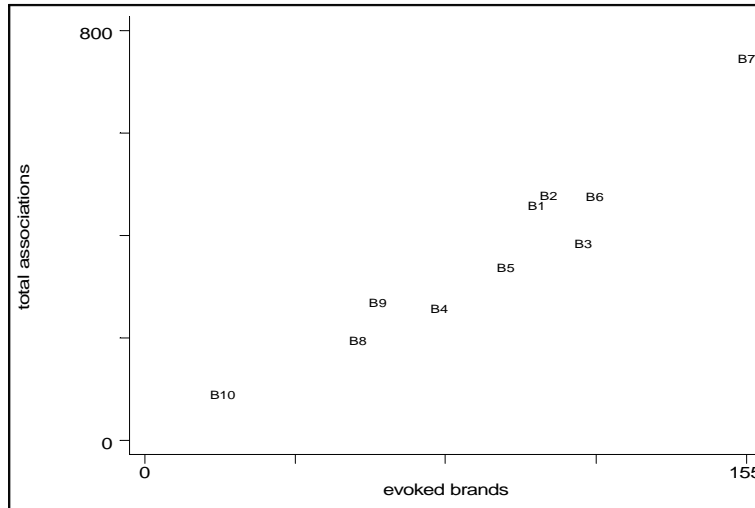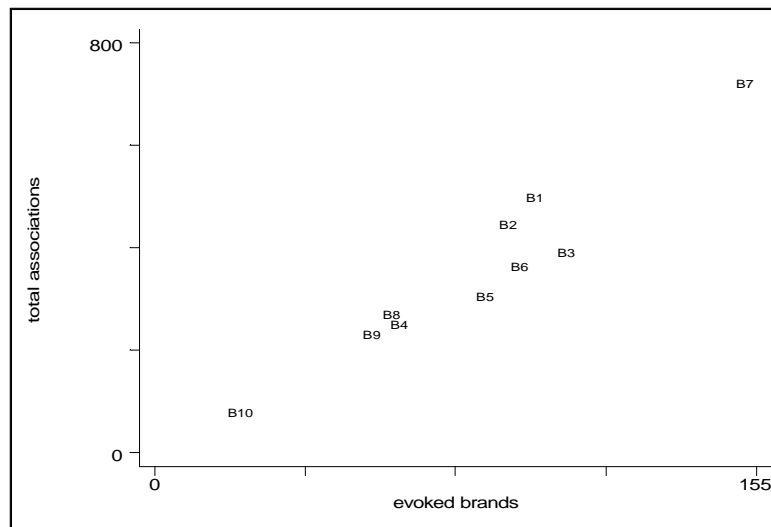

Figure 2 : Total number of associations versus total evoked brands for attribute-by-attribute data (table A).

If we divide the observed frequencies by their totals, we obtain a set of relative frequencies summing up to 1, called a *profile*. In this case, it would be the profile of a brand across its attributes. It is such brand profiles that are visualized in correspondence analysis (Shahim & Greenacre, 1988). Correspondence analysis is thus a tool which can help us display the dominance relations in terms of profiles. We shall have brand profiles for the brand-by-brand (B) matrix, as well as the attribute-by-attribute (A) matrix and we would like to compare two

11

sets of profiles to see to what extent they agree or disagree with one another. As we shall show, methodology exists for techniques such as correspondence analysis to compare two matrices and this will allow us to summarize the possible asymmetries in the dominance relations resulting from the two ways of collecting the data.

5 Correspondence analysis: a tool to describe dominance.

Correspondence analysis (CA) is a statistical technique used to describe and interpret categorical data with an elegant but simple graphical display which permits more rapid interpretation and understanding of the data (see, for example, Greenacre, 1993). When frequency tables are the matrices to be analyzed, then the algorithm of CA can be understood as finding independent factors, each one constructed from the attributes according to the total number of answers or times that each attribute has been associated with one or other brand. The factors define orthogonal dimensions of a perceptual map, where the brands and the attributes are represented by points projected onto the map. The centre of the map, or centroid, can be interpreted as an "average brand" characterized for having an average percentage of association with each attribute. Then, dominance relations between a brand and an attribute will be established with respect to that mean, which we shall call "relative dominance". The implication of this new definition is that we can describe dominance relations with respect to attributes with low averages, while with the previous definitions they might not be recognized as having dominance relations. Further, this sets us free in the sense of not having to fix a percentage threshold (e.g. 50%) for deciding whether dominance relations exist or not.

In matrix notation CA of brand profiles can be written as the following matrix decomposition of the transformed data matrix

$$(\mathbf{R} - 1\mathbf{c}^T)\mathbf{D}_c^{-1/2} = \mathbf{F}\mathbf{V}^T \qquad \text{where } \mathbf{F}^T\mathbf{D}_r\mathbf{F} = \mathbf{I}, \ \mathbf{V}^T\mathbf{V} = \mathbf{I},$$

where $\mathbf{R}$ is row profile matrix of the brands, $\mathbf{c}$ the centroid or "mean brand", $\mathbf{D}_c$ the diagonal matrix with relative attribute frequencies, called column "masses" on the diagonal, $\mathbf{F}$ the principal coordinates for the brands, and $\mathbf{D}_r$ the diagonal matrix with relative brand frequencies on the diagonal. Notice that the column masses are identical to the elements of the average row profile. As an example, for the brand-by-brand case, the data matrix of brand profiles and the average brand profile would be

Table 5: Row profile matrix for brand-by-brand table (*B*)

	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	sum
B1	11.8	10.7	13.3	8.9	10.9	8.4	10.4	6	8	9.1	2.4	100
B2	7.7	4.2	7.3	13	15.2	7.7	8.6	13.2	10.4	9.9	2.9	100
B3	10.7	8.3	11.3	5.4	12.2	6.3	8	3.3	11.6	14.9	8	100
B4	13.6	11.5	13	6.7	12.4	7.3	7.3	6.1	5.2	14.2	2.7	100
B5	13	12.1	10.7	7.1	8.5	11.3	8.8	4.2	5.4	13.6	5.4	100
B6	11.1	8.6	11.8	8.4	13.3	9.8	8.6	5.9	6.6	11.5	4.4	100
B7	13.1	13.1	10.9	7.8	10.9	10.1	7.8	4.8	6.9	10.9	3.6	100
B8	11.1	13.1	13.7	5.2	12.8	7.9	8.5	6.4	7	12.8	1.5	100
B9	9.8	14.4	11.8	8.4	12.1	8	8	3	9.1	12.1	3.4	100
B10	13.4	6.1	17.1	8.5	13.4	6.1	8.5	6.1	12.2	7.3	1.2	100
average	11.5	10.2	12	8.1	12.2	8.3	8.5	6	8.3	11.4	3.5	100

For example, from this table we can see high values of association (relative to the mean) between *B*2 and *A*8. While the percentage for the cell (*B*2, *A*8) is equal to 13.2%, the mean takes a value of 6%. Then, the map will show dominance relations by the bigger differences between a percentage of association and the mean. The operation is repeated for the attribute-by-attribute (*A*) case:

Table 6 : Row profile matrix for attribute-by-attribute table (*A*)

	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	sum
B1	10.4	9	12	7.4	11	8.4	7.4	7.6	11.8	12.5	2.8	100
B2	6	5.2	7.5	11.4	11.8	7.1	7.5	14.3	12.2	13.1	3.9	100
B3	8.5	7.1	10.4	6.3	10.7	7.7	8.5	6	15.7	12.6	6.3	100
B4	11.4	9.6	11.4	7.8	9.9	9.1	7.8	6.2	9.9	14.8	2.1	100
B5	12	10.8	10.8	6.3	7.4	9.7	7.1	6	10.8	16.8	2.3	100
B6	7.9	7.3	12.1	8.1	9.7	7.3	7.3	7.3	11.5	15.2	6	100
B7	11.7	10.6	11	7.2	9.6	10.4	6.8	5.5	9.8	13.4	4	100
B8	9.7	10.2	12.1	7.4	7.8	9	6.7	7.4	13.5	11.2	5	100
B9	9.5	9	12.9	6.7	10.4	9	9.7	3.5	15.2	12.4	1.7	100
B10	7	7	14	5.5	11	5.5	9.8	7	13.7	13	5.5	100
average	9.4	8.6	11.3	7.5	10	8.4	7.8	7.2	12.3	13.5	3.9	100

An example of relative brand dominance for this table could be the couple composed of *A*10, *B*5 which shows a higher association (16.8%) with respect to a mean value of 13.5%.

These might be considered the correct tables to describe relative dominance relations for cases where evoked brands are not available. But in this case we do have information on evoked brands and we should make

use of it. The data matrices to be analyzed would thus be the percentages with respect to the evoked brands (tables 3 & 4). Furthermore, we will consider these percentages corresponding to positive association as well as those corresponding to negative association in the CA analysis.

The tables are thus composed of 10 rows (brands) and 22 columns (11 positive versions and 11 negative versions of the attributes). The row totals or brand totals are all equal to 1100, being the sum of 100 for each pair (positive and negative pole) of the 11 attributes. At the same time the doubling has an implication for the attributes, which have a constant sum of 1000 if we add up negative and positive poles across all 10 brands. In other words, both brands and (doubled) attributes have constant masses in the CA.

We shall show brands and attributes in different maps to be able to appreciate better all the points. The relative positions will indicate the association between the two sets of points.
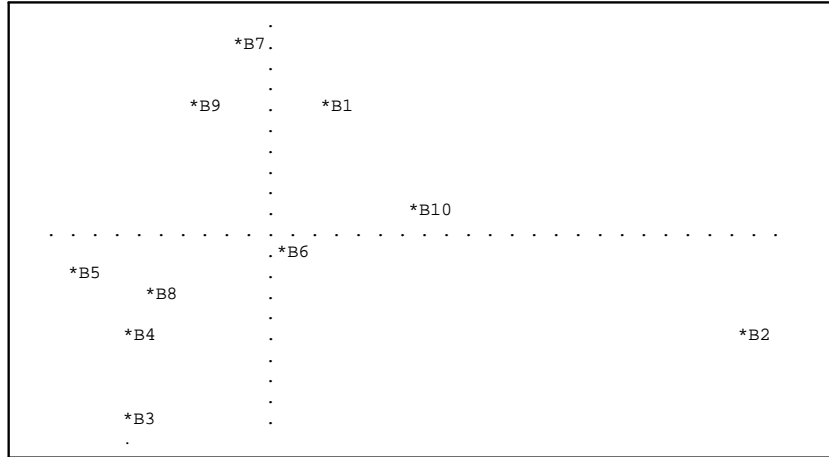
6 Results.

6.1 Attribute dominance for the brand-by-brand table (B).

The tables of results of the analysis are given in the Appendix. We separate attributes and brands in two maps, showing each set in principal coordinates (figures 3 and 4). The first principal axis accounts for 44.0% of the total inertia while the second principal axis accounts for 23.4%. Together they represent 67.4% of the total inertia.

In attribute dominance relations, people first consider the evoked brands and then name the listed attributes they associate with each brand. We first show the symmetric map of brands.
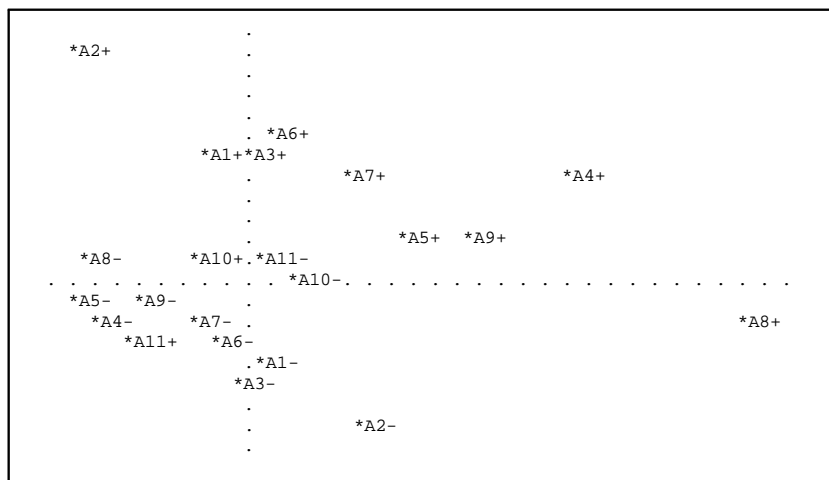
14

Figure 3: Correspondence analysis of doubled brand-attribute associa-
tion expressed as percentages: principal coordinates of brands.

```
                              .
                  *B7.
                              .
                              .
              *B9     .      *B1
                              .
                              .
                              .
                              .
                              .       *B10
                              .
. . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
                         .*B6
      *B5                   .
           *B8              .
                              .
        *B4                   .                            *B2
                              .
                              .
                              .
        *B3                   .
                    .
```

In general terms we can describe the following. On the right hand
side of the map, we find $B2$, while on the left hand side, opposite to $B2$,
are situated $B5$, $B8$ and $B4$. On the top, we see $B7$, $B9$ and $B1$, and
on the bottom, $B3$. $B6$ appears almost in the centre and $B10$, next to
the first principal axis, towards the right hand side.

The second map we offer is the map of doubled attributes in principal
coordinates:

Figure 4: Correspondence analysis of doubled brand-attribute associ-
ation expressed as percentages of evoked brands: principal coordinates of
doubled attributes.

```
                       .
      *A2+             .
                       .
                       .
                       .
                       . *A6+
               *A1+*A3+
                       .       *A7+               *A4+
                       .
                       .
                       .       *A5+   *A9+
      *A8-      *A10+.*A11-
. . . . . . . . . . . *A10-. . . . . . . . . . . . . . . .
    *A5-   *A9-        .
     *A4-     *A7- .                               *A8+
        *A11+   *A6-
              .*A1-
           *A3-
                .
                .       *A2-
                .
```

15

We have two points for each attribute, the positive pole opposite the negative pole, on a line through the origin. On the one hand, the positive pole represents the group of people who evoked the brand and established a positive association between the brand and the attribute. On the other hand, the negative pole represents people who evoked the brand and did not establish the association between the particular attribute and the brand. Furthermore, if distance between the poles were to be calibrated in a 100 equal intervals from 0 at the negative pole to 100 at the positive one, then the average rating could be read off at the point where the line crosses the centre of the display. For example, the average rating on $A8$: *can use all over the body*, has an average association equal to 24% (see *sum* in the data matrix of the Appendix and divide the row by 10, the number of brands, to get the mean). Hence the origin is situated nearer to the negative pole of this attribute. For this same attribute, we can see that it is the one for which the poles are the most separated. The reason is the higher variability in the percentage responses for this attribute. Other attributes with high variability are $A2$: *keeps you dry all day*, with an average of 40.7% and $A4$: *has a long lasting fragance*, with an average of 32.5%. An example of small variability is $A10$: *quick to apply* with an average equal to 45.7%. The relative directions of the lines connecting opposite poles indicate correlations between the attributes. For example, $A11$: *costs a little less than most others*, is mostly negatively correlated with all other attributes. The reason can be that since all the attributes are profits or positive versions, you are disposed to pay more for all of them, which could be the negative pole of attribute $A11$.

In general, figure 4 shows attributes $A8$, $A4$, $A9$ and $A5$ along the horizontal principal axis and $A2$, $A1$, $A3$, and $A6$ along the vertical axis. Positive poles are in the upper right part of the map and negative poles in the lower left, except for attribute $A11$ mentioned above.

From both maps, we can interpret the existing associations. On the one hand, given the positions previously defined, $B2$ shows relative attribute dominance with respect to $A8_+$: *can be used all over the body*, since once we show $B2$ to the people who evoked this brand, the proportion of them who gave the attribute in response to the brand is higher than the mean association established for this attribute and all brands. At the other end of the first principal axis, $B5$ shows an association with the negative pole of attribute $A5_-$: *has not a pleasant fragrance*, since between all brands, it is the one with a lower percentages of association. $B8$, as well as $B3$ and $B4$ show an association with respect to $A4_-$: *has not a long lasting fragance*. $B3$ shows relative attribute dominance with respect to $A11_+$: *costs a little less than most others*. Going to

16

the second principal axis, we find relative attribute dominance relations between $B7$ and $B9$ with respect to $A2_+$: *keeps you dry all day*. They could be considered brands in direct competence in terms of this profit or core attribute. $B1$ shows relative attribute dominance with respect to $A3_+$: *does not irritate my skin* and $A7_+$: *leaves no marks on clothes*. Finally, $B10$ shows relative attribute dominance with respect to $A9_+$.

6.2 Brand dominance for the attribute-by-attribute study (A).

Figures 5 and 6 show the corresponding maps of attributes and brands respectively, both in principal coordinates. In this case, where people is asked attribute-by-attribute, we measure relative brand dominance, in other words, higher proportions of people with respect to the mean, who gave the evoked brand in response to a particular listed attribute. Given the symmetry in the treatment of brands and attributes for the correspondence analysis in the symmetric map, we are able to already establish a comparison between this analysis and the previous one, to infer some possible asymmetries.

While the total inertia has decreased, the proportion explained by the first two principal axes has increased. The first one explains 46.4% of the total inertia while the second one explains 31.4%. The distribution of attributes in the space seems quite similar compared to the brand-by-brand analysis, but there are some differences.

We will start by the description of attributes, which are the first that appear in a brand dominance study. The more evident changes that appear with respect to the previous figure 4 are the change in $A10$: *is quick to apply* and $A11$: *costs a little less than most others* positions. While $A10_+$ appeared before in the left hand side and quite near to the centroid, now it appears in the right hand side further from the centroid. Attribute $A11_+$ appears in the third quadrant, quite far from the centroid, which indicates an increase in its correlation with the second principal axis. The same attribute, in the previous analysis, appeared in the fourth quadrant and among all the negative attribute poles. We can also see an increase of variability for $A1$: *prevents body odour all day*, as well as an increase in its correlation with the second principal axis. Its position is now very similar to $A2$: *keeps me dry all day*. Apart from these differences the points in figure 5 appear rotated compared to those in figure 4.

Figure 5: Correspondence analysis of doubled brand-attribute associations expressed as percentages of evoked brands: principal coordinates of doubled attributes.
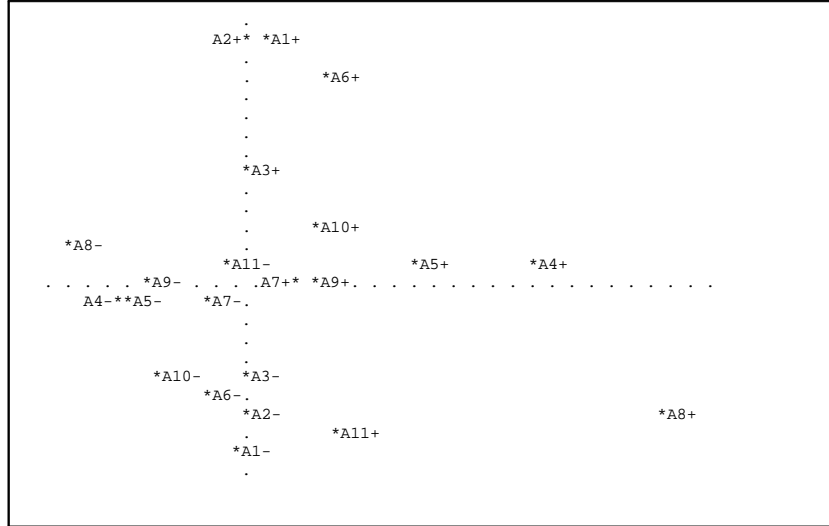


```
                              .
                         A2+* *A1+
                              .
                              .         *A6+
                              .
                              .
                              .
                              .
                         *A3+
                              .
                              .
                              .         *A10+
   *A8-                       .
                       *A11-              *A5+        *A4+
. . . . . . *A9-  . . . .A7+* *A9+. . . . . . . . . . . . . . . . . . . .
    A4-**A5-      *A7-.
                         .
                         .
                         .
     *A10-    *A3-
          *A6-.
              *A2-                                      *A8+
              .        *A11+
           *A1-
              .
```

Figure 6 shows the position of brands, for which we can deduce relative brand dominance relations. Before going to the description of dominance relations, we will list the changes with respect to figure 3. $B10$ changes its position since now it appears in the fourth quadrant while in the previous analysis, it was situated in the second one. While $B5$, $B4$ and $B9$ where situated in the fourth quadrant, they go up and now are located in the first one. Since the inertia of the second principal axes increases, all brands appear more dispersed.

Figure 6: Correspondence analysis of doubled brand-attribute associations expressed as percentages of evoked brands: principal coordinates of brands.
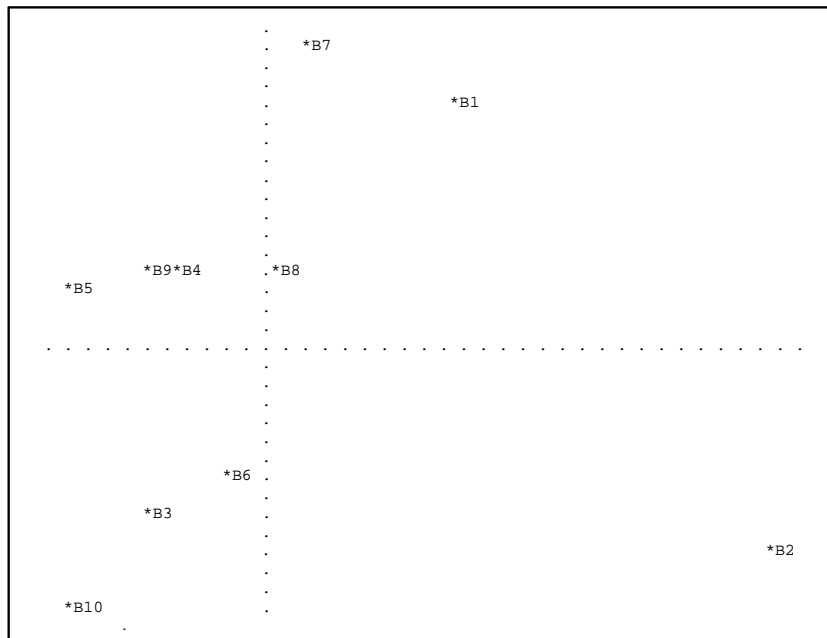
Relative brand dominance between $B2$ and $A8_+$: *can use all over the body* is shown, in other words, the proportion of people, who gave $B2$ in response to $A8_+$ is higher than the mean proportion given for this attribute. On the right hand side, where almost all positive poles are located, we find a relative brand dominance relationship between $B7$ and the three following attributes: $A1_+$: *prevents body odour all day*, $A2_+$: *keeps me dry all day* and $A6_+$: *leaves me feeling confident*. $B1$'s location indicates relative dominance between this brand and the following attributes: $A6_+$: *leaves me feeling confident* and $A10_+$: *is quick to apply*. Finally, $B8$'s image improves if we ask attribute-by-attribute instead of brand-by-brand since it passes from the part of the map where negative poles are concentrated to the one where positive ones are located. To be more concrete, in the brand-by-brand analysis it was perceived as a brand characterized for not having a long lasting fragrance while now, when people have the listed attributes and have to associate them with the evoked brands, if it is true that it does not show strong dominance relations, we can say that it is associated with the attributes of being portable and quick to apply.

Once we go to the left hand side we find that $B10$, which in the previous analysis was situated on the right hand side, close to $A5_+$: *has a pleasant fragrance* and $A4_+$: *has a long lasting fragrance*, now shows high association with $A1_-$: *does not prevent body odour all day*, $A2_-$:

does not keep me dry all day and $A10_-$: *is not quick to apply.* Thus, its image becomes worse if you ask attribute-by-attribute instead of brand-by-brand. $B3$ and $B6$ show association with respect to $A1_-$, $A2_-$ and $A10_-$ and $A6_-$: *does not leave me feeling confident,* which is nothing else but a consequence of missing basic attributes. For these brands, there are no relevant differences, similarly for $B5$, which is related with $A8_-$: *cannot be used all over the body.* $B9$ and $B4$, which are quite close, do not show high correlations with these two principal axes but their positions could imply a relation with $A11_-$: *cost a little more than most others.* While in the previous analysis $B7$ and $B9$ where quite close, now they are quite far apart, with $B9$ in a worse position now.

7 CA of matched matrices: a description of asymmetry

In the above we analyzed the brand-by-brand data separately from the attribute-by-attribute data and compared the results of the two analyses. Greenacre (2000) has already considered the joint visualization of two matrices, with common rows and columns. The analysis consists in applying the singular value decomposition to the two matrices in a particular block format, leading to an analysis of the sum and the difference components. In our case, the difference components will capture the asymmetries existing between the two matrices $\mathbf{B}$ and $\mathbf{A}$.

In the simplest case if $\mathbf{B}$ and $\mathbf{A}$ are two $n \times m$ matrices, then the SVD of the sum, for the matrices $\mathbf{B} + \mathbf{A}$ and the difference $\mathbf{B} - \mathbf{A}$ can be recovered in the SVD of the block matrix:

$$\left[\begin{array}{cc} \mathbf{B} & \mathbf{A} \\ \mathbf{A} & \mathbf{B} \end{array} \right] \tag{3}$$

Greenacre (2000) shows that if the SVD of $\mathbf{B} + \mathbf{A}$ and $\mathbf{B} - \mathbf{A}$ are respectively:

$$\mathbf{B} + \mathbf{A} = \mathbf{U}\mathbf{D}_\alpha\mathbf{V}^T \qquad\qquad \mathbf{B} - \mathbf{A} = \mathbf{X}\mathbf{D}_\beta\mathbf{Y}^T \tag{4}$$

then the SVD of the $2n \times 2m$ block matrix is:

$$\left[\begin{array}{cc} \mathbf{B} & \mathbf{A} \\ \mathbf{A} & \mathbf{B} \end{array} \right] = \frac{1}{\sqrt{2}} \left[\begin{array}{cc} \mathbf{U} & \mathbf{X} \\ \mathbf{U} & -\mathbf{X} \end{array} \right] \left[\begin{array}{cc} \mathbf{D}_\alpha & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_\beta \end{array} \right] \frac{1}{\sqrt{2}} \left[\begin{array}{cc} \mathbf{V} & \mathbf{Y} \\ \mathbf{V} & -\mathbf{Y} \end{array} \right]^T \tag{5}$$

Thus the solutions corresponding to the difference component will appear as repeated vectors with a change in sign in the singular vectors.

In our case, to describe asymmetries and compare them with the previously defined dominance relations, we are going to take the doubled

percentage tables with respect to the evoked brands in each table. Thus the data matrix to be analyzed is:

$$\left[\begin{array}{cccc} \mathbf{B}_+ & \mathbf{B}_- & \mathbf{A}_+ & \mathbf{A}_- \\ \mathbf{A}_+ & \mathbf{A}_- & \mathbf{B}_+ & \mathbf{B}_- \end{array} \right] \tag{6}$$

where subindexes $+$ and $-$ distinguish between matrices containing the positive and the negative versions of the attributes. Applying CA to this block matrix, the row (brand) masses will be constant and the column (attribute) masses will be constant if we sum the positive and the negative poles. Notice that the doubling of the percentages solves the problem mentioned by Greenacre (2000) of varying marginal totals when applying CA to matched matrices of frequencies. The difference matrix analyzed is $[\mathbf{B}_+ - \mathbf{A}_+ \quad \mathbf{B}_- - \mathbf{A}_-]$. Since the elements of $\mathbf{B}_+ + \mathbf{B}_-$ and $\mathbf{A}_+ + \mathbf{A}_-$ are all equal to 100, this difference matrix has the property that $\mathbf{B}_+ - \mathbf{A}_+ = -(\mathbf{B}_- - \mathbf{A}_-)$. Hence we need to display only the differences between the "positive" parts. The difference between the "negative" parts will have coordinates exactly the negative of the "positive" ones.

8 Results

Figures 7 and 8 display the asymmetries existing between the two ways of data collection. The principal axes of the differences are obtained from axes 3 and 5 of the analysis of the block matrix (see the signs of the coordinates in the results given in the Appendix).

Figure 7: CA of differences between doubled matrices. Attributes in principal coordinates
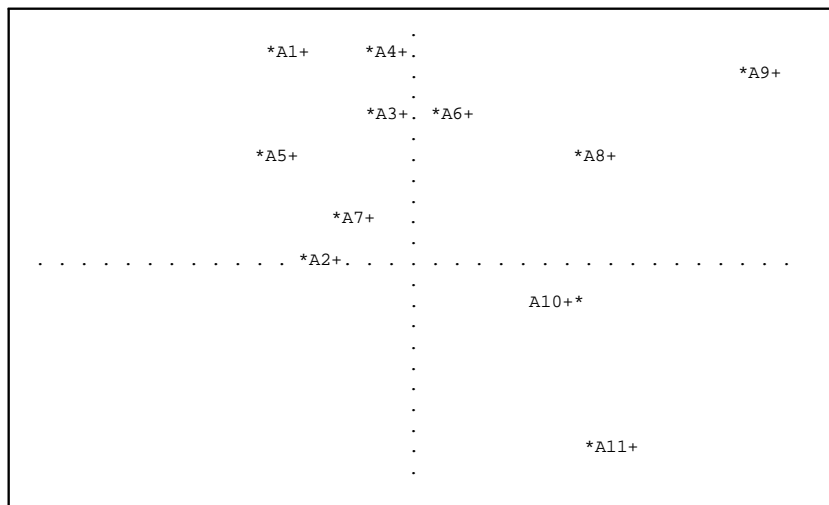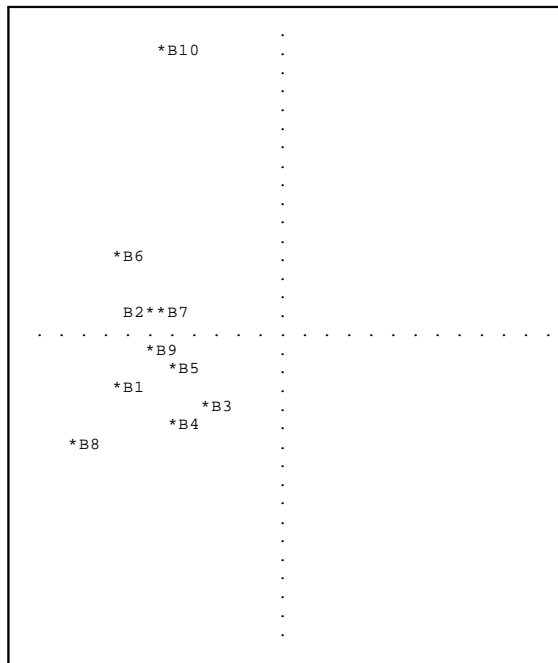
Figure 8: CA of differences between doubled matrices obtained from the analysis of matched matrices. Brands in principal coordinates

All attributes in the map reflect positive differences between B and A. All the brands are situated on the left hand side of the map.

The attribute which seems that has contributed most to the first principal axis of differences is $A9_+$: *is portable*. The attributes which seem have contributed more to the second principal axes are $A1$: *prevents body odour all day*, $A4$: *has a long lasting fragrance*, $A9$: *is portable* and $A11$: *costs a little less than most others*. $B8$ shows association with the counterpart of $A9$, in other words, it is situated just opposite to $A9$ in diagonal, indicating that it is more associated with the attribute of being portable when we ask attribute-by-attribute instead of doing brand-by-brand. The same happens with $A11$ where its counterpart position with respect to it indicates that it also is perceived as a more cheaper brand if the study is done asking attribute-by-attribute instead of brand-by-brand. We follow with $B10$. Its relative position close to $A1$, $A3$ and $A4$ indicates that its is more associated with these attributes, which are *prevents body odour all day*, *does not irritate my skin* and *has a long lasting fragrance*, respectively, when the marketing researcher ask brand-by-brand instead of doing attribute-by-attribute. However, at the same time, it is associated as a more expensive brand when the analysis is done attribute-by-attribute (opposite to $A11$: *costs a little less than most others*).

$B6$ shows a stronger association with the attributes of *having a pleas-*

ant fragrance ($A5$) and *prevents body odour all day* ($A1$) if we ask brand-by-brand instead of doing attribute-by-attribute.

$B1$, $B3$ and $B4$ also occupy the counterpart position with respect $A9$ indicate, as was the case of $B8$, showing that their image as a *portable brand* is stronger in the attribute-by-attribute analysis.

Finally, $B9$ increases its association with respect to *keeping me dry all day* ($A2$) in the brand-by-brand analysis with respect to the attribute-by-attribute one.

9   Conclusions and discussion

We have shown how CA can be an alternative tool to describe dominance relationships, since it offers a map, which shows in an understandable way strong associations between brands and attributes, and so possible dominance relations with respect to a "mean" brand for that particular sample. For this particular case, we expressed the data relative to the evoked brand totals in order to study attribute dominance and brand dominance for the two ways of data collection; brand-by-brand (B) and attribute-by-attribute (A) respectively.

Furthermore, we show a particular way of coding data in CA, called matched matrices as a tool for measuring potential asymmetries between brand-by-brand (B) and the attribute-by-attribute (A) analysis. The results of the two analyses were quite similar, but we could still describe two types of asymmetries. On the one hand, the attributes characterized for being more practical and not such basic properties of a deodorant, such as *to be a portable deodorant*, *quick to apply*, *costs a little less than most other brands* and *can use all over the body*, received more answers for the attribute-by-attribute analysis than the brand-by-brand one. Our interpretation was that these attributes are not as relevant as the others since they are not core attributes normally looked for in a deodorant brand. When one establishes an association brand-by-brand, you normally associate brands with other benefits or attributes compared with the attribute-by-attribute study.

Another possible interpretation of such results could be the following. This increase in answers in the A analysis with respect to the B one is given by the more "practical" attributes which happen to be together at the end of the list. It may be possible that in the brand-by-brand data collection, where 11 attributes have to be associated, that the final attributes on the list receive less associations compared with the attribute-by-attribute analysis. It would be interesting to ask the marketing research company who elaborated the data samples how they established this particular order in the listed attributes, and whether they purposely put the less core attributes at the end.

We would like to point out as a direction for future research the use of compositional data of biplots as an alternative to CA of matched matrices in the description of asymmetries, where, instead of interval differences, a map is constructed of the logarithms of ratios from the two data sets. A recent reference is Aitchison & Greenacre (2001). This analysis could represent an improvement with respect to the previous one if we consider ratios to be better than differences to measure asymmetries.

10   Acknowledgements

References

[1] Aaker, D.A. (1991), Managing brand equity, New York, The Free Press.

[2] Aitchson J. & Greenacre, M.J. (2001), Biplots of compositional data. Submitted for publication, *Working Paper*, n°557, Department of Economics and Business, Universitat Pompeu Fabra, Barcelona.

[3] Ashcraft, M.H. (1978), Property dominance and typicality effects in property statement verification, *Journal of Verbal Learning and Verbal Behavior*, **17**, 155-164.

[4] Barnard, Neil R. & Ehrenberg, Andrew S.C. (1990), Robust Measures of Consumer Brand Beliefs, *Journal of Marketing Research*, **27**, 477-484.

[5] Benzécri, J-P & collaborators (1973), *L'Analyse des Données. Tome 2: l'Analyse des Correspondances*, Dunod, Paris.

[6] Bird, M., Channon, C. & Ehrenbert, A.S.C. (1970), Brand image and brand usage, *Journal of Marketing Research*, **7**, 307-314.

[7] Blasius, J. & Greenacre, M.J. (1994), *Correspondence Analysis in the Social Science.*

[8] Driesener, C. (2001), Brand image? Pick a measure, any measure, Currently in review for Australian and New Zealand Marketing Academic's Conference.

[9] Farquhar, P. H. (1989), Managing brand equity, *Marketing Research*, **1**, 24-33.

[10] Farquhar, P. H. & Herr, P. M. (1992), The dual structure of brand association, in *Brand Equity & Advertising: Advertising's Role in Building Strong Brands*, eds. David A. Aaker & Alexander L. Biel, Lawrence Erlbaum Associates Publishers, Hillsdale, New Jersey, 263-277.

[11] Farquhar, P. H, Herr, P.M., & Fazio, R. H. (1990), A relational model for category extensions of brands, *Advances in Consumer Research*, **17**, 856-860.

[12] Farquhar, P. H., Han J. Y., Herr P. M., & Ijiri Y. (1992), Strategies for leveraging master brands, *Marketing Research*, **4**, 32-43.

[13] Greenacre, M.J. (1984), *Theory and Applications of Correspondance Analsyis*, Academic Press, London.

[14] Greenacre, M.J. (1986), SimCA: a program to perform simple correspondence analysis. *Psychometrika,* **51**, 172-173.

[15] Greenacre, M.J. (1993), *Correspondence Analysis in Practice*, Academic Press, London.

[16] Greenacre, M.J. (2000), Analysis of matched matrices, *Working Paper*, n°539, Department of Economics and Business, Universitat

Pompeu Fabra, Barcelona.

[17] Greenacre, M.J. & Hastie, T. (1987), The geometric interpretation of correspondence analysis, *Journal of the*

[18] Hoffman, D.L. & de Leeuw, J. (1992), Interpreting multiple correspondence analysis as a multidimensional scaling method, *Marketing Letters*, **3**, 259-272.

[19] Hoffman, D.L. & Franke, G.R. (1986), Correspondence analysis: graphical representation of categorical data in marketing research, *Journal of Marketing Research*, **23**, 213-217.

[20] Hughes, G.D. (1969), Some confounding effects of forced choice scales, *Journal of Marketing Research*, **6**, 223-6.

[21] Joyce, T. (1963), Techniques of brand image measurement, in *New Developments in Research*. London: Market Research Society, 45-63.

[22] Lebart, L., Morineau, A., & Warwick, K, (1984), Descriptive Multivariate Statistical Analysis, John Wiley, Chichester, U.K.

[23] Loftus, E. (1973), Category dominance, instance dominance, and categorization time, *Journal of Experimental Psychology*, **1**, 70-74.

[24] Morrison, D. G. (1979), Purchase intentions and purchase behavior, *Journal of Marketing*, **43**, 65-74.

[25] Shahim, V. & Greenacre, M.J. (1988), Brand maps, ideal points and market groups, Proceedings of the Annual Convention of the South African Market Research Association.