

Comparing methods of brand image measurement

Carl Driesener and Jenni Romaniuk

Ehrenberg-Bass Institute of Marketing Science, University of South Australia

This study compared rating, ranking and 'pick-any' measures of brand image associations. The pick-any technique is a free response measure, where respondents are given an attribute as a cue and asked which brands they associate with it. It is a free response in that respondents can link any, all or no brands with each attribute. It only captures the association, however, with no indication of relative strength. The study confirmed past findings that the three measures are highly correlated at brand level (average correlation of 0.90). Further analysis at individual level found that individuals utilised the three measures in a consistent manner, suggesting that the measures are virtually interchangeable. The main exception to this is when respondents rank brands; low ranks may simply be a reflection of unfamiliarity rather than poor performance on the attribute. When examining the time taken to administer each technique, however, there are clear benefits in a pick-any approach, which takes about half the time of the other methods to collect image data.

Introduction

The perceptions that consumers hold about brands, referred to collectively as the brand's image, are an important part of consumer-based brand equity (Keller 1993). They are also used to identify the position of the brand and evaluate advertising effectiveness (Romaniuk & Sharp 2000; Dillon *et al.* 2001; Myers 2003; Romaniuk & Nicholls 2006).

There is a wide range of techniques for measuring the extent of the link between the brand name and relevant concepts of interest. Three techniques that are commonly used involve rating brands on Likert-type rating scales (e.g. how much do you agree that Brand X is *good value?*), ranking measures (which brand ranks highest on *good value*, next highest

Received (in revised form): 12 June 2006

and so on) and pick-any brand-attribute association measures (which brands are good value?).

While there are many different methods of measuring brand image, techniques fall into two main categories: scaling and sorting (Joyce 1963). Scaling techniques determine not only whether there is an association between a brand and an attribute but also the strength of that association. In contrast, sorting techniques merely seek to determine if there is an association. The absence of sorting techniques from most academic literature comparing brand image measures suggests that researchers perceive the scaling brand image techniques as superior, or indeed the only, brand image measures (e.g. Axelrod 1968; Haley & Case 1979; Kalwani & Silk 1982). It has, however, been shown that most variation in brand image data is due to the direction of the association, with only a small component resulting from the measure of the intensity (or strength) of that association (Jacoby & Matell 1971). Therefore, the incremental value of determining the strength of an association is debatable.

The few studies that have compared sorting with scaling techniques have generally concluded that these measures of brand/attribute association provide similar brand-level results (Joyce 1963; Barnard & Ehrenberg 1990). This suggests that both are equally effective in capturing brand-attribute linkage. This past research, however, focused solely on the brand (or aggregate) level. While the brand ranking hierarchies may be comparable, individuals could respond to the alternative techniques less consistently. To date there are no reported studies comparing responses at the individual level for different brand image measures. In addition, prior studies were on the whole limited to fmcg markets (Joyce 1963; Barnard & Ehrenberg 1990; Romaniuk 2005). In one exception, Romaniuk (2005) examined financial services, but only with the two measures of rating and pick-any, and again at the brand level. Thus the generalisability of these results beyond fmcg markets remains unclear.

The objective of this study was to extend the comparison of the different techniques of measuring brand associations to a non-fmcg market and, more importantly, to extend the analysis to examine individual-level responses. Further, given the prior (brand-level) similarity in results, the study explored another consideration of measurement, namely application. It has often been anecdotally noted that the pick-any technique is more efficient to administer, however no quantitative data have been presented (Joyce 1963; Hoffman & Franke 1986; Barnard & Ehrenberg 1990). In this research we measured the time taken to introduce and administer the different techniques and compared them using this criterion.

Brand image measures

The three brand image measurement techniques used in this research were:

1. A Likert rating technique (scaling)
2. A ranking technique (scaling)
3. A pick-any technique (sorting).

When rating brands, respondents are typically asked to respond on a 5- or 7-point scale, which ranges from, for example, strongly agree to strongly disagree (Likert 1932). The score gauges the extent to which the respondent feels the brand is associated with a certain attribute.

The second type of measure (ranking) is where brands are ranked relative to competitors according to their association with an attribute (e.g. a brand ranked 'first' means that the brand is most associated with/strongest on the attribute of the choice set). The key difference between rating and ranking approaches is that while two brands cannot be afforded the same rank, they can be afforded the same rating.

The third measure is the 'pick-any' measure, where respondents are asked which brands, if any, they associate with each attribute (for an example of the technique see Holbrook *et al.* 1982). Respondents can name any, all or none of the brands. As with the other techniques, brand names can be provided in a list or recalled from memory. The key difference between this measure and the others is that the pick-any measure is a response indicating association with the brand, rather than the *degree* of association. With this technique, therefore, multiple brands can be associated with the same attribute, but there is no discrimination between those brands as to the degree of association with the attribute by that individual.

These measures thus differ on three criteria. The first is whether the technique is scaling or sorting (Joyce 1963). Rating and ranking measures require the respondent not only to show whether or not there is an association but also to indicate the strength of that association. These are therefore scaling techniques. In contrast, the pick-any measure requires the respondent only to indicate whether they perceive the brand and attribute to be associated. The second criterion is that for some measures brands are not directly compared, whereas others explicitly require a direct comparison of the brands. Ranking is considered to be a comparative measure, while rating and pick-any approaches are not (Joyce 1963). The final distinction is forced choice vs free choice. With a forced-choice technique, the respondent is required to give an explicit response for every

brand/attribute combination. In comparison, free-choice techniques allow the respondent to choose those brands and attributes to which they would like to provide a response. The pick-any technique is free choice, while rating and (generally) ranking are forced choice (Joyce 1963; Driesener 2002).

Brand-level responses

Despite the differences between the measures, past empirical research comparing them in fmcg markets found they provided similar brand-level assessments for each attribute, with brand-level correlations generally higher than 0.85 observed across the various markets examined (Joyce 1963; Barnard & Ehrenberg 1990; Driesener & Romaniuk 2001; Romaniuk 2005). Thus, regardless of the technique, brands are ranked similarly for each attribute. In further analysis comparing responses between users and non-users, all three techniques were found to be similarly biased (Barnard & Ehrenberg 1990; Driesener & Romaniuk 2001; Romaniuk 2005) in that users of a brand consistently score that brand higher than non-users. Our first hypotheses regarding brand-level correlations and the effect of brand usage are thus:

H1a: The three techniques – pick-any, rating and ranking – place the competing brands in the same relative positions for each attribute.

H1b: For each technique, users of a brand give higher belief scores than do non-users.

These brand-level hypotheses replicate those of Barnard and Ehrenberg (1990).

Individual-level responses

It is reasonably well known that even when brand image scores are stable at the brand level under test–retest conditions, scores vary at an individual level (Castleberry *et al.* 1994; Dall’Olmo Riley *et al.* 1997, 1998). Thus when the same respondent is interviewed twice using the same measure, they do not always give the same answer. In fact, the proportion of respondents that give the same answer on a subsequent occasion is related to the number of respondents that initially gave that response (Sharp 2002; Rungie *et al.*

2005). While the stability of scaling techniques has not been examined to the same extent, there is evidence that these also exhibit considerable individual variation between two interviews (Holmes 1974; Sharp 2002).

Thus it is possible that while there are strong correlations between brand image measurement techniques at the brand level, individually respondents may be using the techniques in quite different ways. For example, someone could rate a brand highly on a rating scale but rank them relatively lowly when asked to rank the brand relative to competitors. Similarly someone might not choose a brand with the pick-any technique but give that brand a high rating when forced to respond with a rating measure.

The lack of understanding about how individuals respond to each of the measures constitutes a major deficiency in our knowledge regarding the use of such techniques. Therefore the following hypotheses are tested in this research:

H2a: The brand score given by an individual respondent using the rating technique will tend to be matched by the rank given using the ranking technique.

H2b: A respondent will tend to give a higher rank to a brand that they also have selected using the pick-any technique.

H2c: A respondent will tend to give a higher score using the rating technique to a brand that they also have selected using the pick-any technique.

Research method

The data were collected from recent car buyers who were asked about their perceptions of car brands in the market. To test the brand-level hypotheses (H1a and H1b), responses were collected using a split-sample approach, where respondents were randomly allocated to one of the three techniques. The total sample size was 369.

Data for the individual-level analysis were then collected from a subset of stage 1 respondents at later points in time in stages 2 and 3. The repeat interviews were conducted at least two weeks apart as part of a larger survey. This time lag reduced the likelihood that people would exhibit memory or learning effects that may contaminate later responses. At the conclusion of data collection, a total of 105 respondents had given valid

responses using at least two of the three possible techniques, enabling the testing of H2a, H2b and H2c. It is important to note that we were not concerned with *which* people used the techniques, only *how* they were used; for experimental work such as this, attrition and non-response bias is not an issue (Barnard *et al.* 1986).

A total of six brands and ten attributes were used in this study. To protect the commercial interests of the sponsor, the brands have been disguised. The ten attributes used were: *good value, advanced, distribution, interesting, local, modern, safe, secure, trust and handles well*.

The brands and attributes used in this research were chosen independently of the researchers by the research sponsor. This lack of input into attribute selection reduced the possibility of researcher bias in this area.

The specific introduction for each technique varied according to the technique and was as follows.

Ranking: We would like to know how you regard the different brands. We will give you a series of statements and we would like you to rank the brands so that the brand ranked 1st is most closely associated with a statement, and the brand ranked 6th is least associated with it. It does not matter whether you have purchased this brand before or not; it is your opinion we are after.

Rating: We would like to know how you regard the different brands. We will give you a series of statements and for each brand we would like to know how strongly you agree or disagree with the statement. It does not matter whether you have purchased this brand before or not; it is your opinion we are after.

Pick any: We would like to know how you regard the different brands. We will give you a series of statements and for each we would like to know which brands you associate with it. You can select as many or as few brands for each statement as you like. It does not matter whether you have purchased this brand before or not; it is your opinion we are after.

Results: brand level

An example of the analysis used to test H1a is shown in Table 1. The brand-level results for each measurement technique for a single attribute

Table 1 Brand-level technique comparison for the attribute *Advanced*

	Pick-any (%)	Rating (mean)	Ranking (mean)
Brand 1	54	2.8	5.2
Brand 3	36	2.6	4.7
Brand 2	27	2.3	4.5
Brand 4	10	1.9	2.7
Brand 6	10	2.0	2.6
Brand 5	7	1.6	2.5
	Pick-any and rating	Rating and ranking	Ranking and pick-any
Correlations	0.95	0.96	0.96

were correlated, which gave three correlations for each attribute (representing the technique pairs). For the sake of clarity and ease of presentation, these scores have been rounded to two effective digits (Ehrenberg 1992).

The results across all attributes (Table 2) showed high, positive correlations (generally over 0.85) between the different techniques. This is consistent with the findings of prior research.

The high overall average correlation, and generally high level of inter-technique correlation, was consistent with that found by Barnard and Ehrenberg (1990), and corroborated their finding that the measures produced very similar results at brand level. It was also comparable to the

Table 2 Brand-level technique correlations for all attributes

	Pick-any and rating	Rating and ranking	Ranking and pick-any	Row average
Advanced	0.95	0.96	0.96	0.95
Distribution	0.94	0.99	0.94	0.96
Secure	0.94	0.96	0.94	0.95
Local	0.95	0.97	0.89	0.94
Handles well	0.93	0.91	0.88	0.91
Good value	0.91	0.95	0.83	0.90
Safe	0.94	0.83	0.93	0.90
Trust	0.98	0.80	0.88	0.89
Modern	0.98	0.71	0.78	0.82
Interesting	0.87	0.77	0.72	0.79
Column average	0.94	0.89	0.88	
Overall average				0.90
Average from Barnard and Ehrenberg (1990)				0.86
Average from Joyce (1963)				0.87

Comparing methods of brand image measurement

average correlation of 0.87 between the pick-any and the semantic differential measurement techniques tested by Joyce (1963). The similarities previously noted in fmcg markets were present in the durables market tested here. This supports H1a.

Table 3 shows the average of the scores given to a brand (any brand) for each attribute by its users and non-users across the three brand image measurement techniques. Users of a brand gave a higher score (or rank, or a greater percentage associated it) than did non-users for that attribute/brand combination on average. To facilitate comparisons between techniques, we have reported the difference as a percentage of the number of scale points of each measure.

The average difference over the ten attributes between users and non-users was greater for the pick-any approach (20%) compared to ranking (17%) and rating (which was the lowest at 6%). While the average difference in the rating technique was somewhat smaller, each of the techniques for all attributes showed a positive difference between users and non-users, thus all of the techniques clearly exhibited this effect. Further, in the ten comparisons of users and non-users that underlie each of the average differences above, in not a single case did non-users score more than users.

The consistent difference between users and non-users (Table 3) supports H1b. Attributes that describe functional qualities about brands tend to have a similar level of agreement between users and non-users, and are referred to as descriptive (Bird & Ehrenberg 1970; Barwise & Ehrenberg 1985; Winchester & Romaniuk 2003). In the current research there was one attribute that met this criterion, namely *Local*, and this had a systematically higher level of agreement between users and non-users across all three measures. This provided further verification of the similarity between the measures.

Overall, this replication of Barnard and Ehrenberg's (1990) research had very similar results. The three techniques ordered the brands in a consistent fashion (H1a), and users of a brand scored that brand higher than did non-users for any given attribute on all of the three techniques (H1b).

Table 3 Average of users vs non-users

	Users	Non-users	Difference	Difference adjusted for scale points (%)
Pick-any	49	29	20	20
Rating	2.7	2.3	0.3	6
Ranking	4.5	3.6	1.0	17

Results: individual level

The first of the individual-level analyses examined the relationship between the rating and ranking techniques (H2a). Correspondence analysis (Greenacre 1984, 1993) was used to test this relationship. In addition to the more commonly produced perceptual maps describing relationships between brands and attributes, this technique can be used to compare scales (Herman 1991; Greenacre 1993; Bendixen & Sandler 1994; Kennedy *et al.* 1995; Riebe *et al.* 1998). A correspondence analysis plot enables a direct comparison of the resultant inter-(data) set distances on the map, with proximity on the map representing similar response patterns (Bendixen 1991; Greenacre 1993). If 'strongly agree' and '1st rank' are in close proximity, this would suggest that many people who ranked a brand 1st on an attribute for the ranking question also gave a 'strongly agree' when asked about the same brand and attribute on the scaling question.

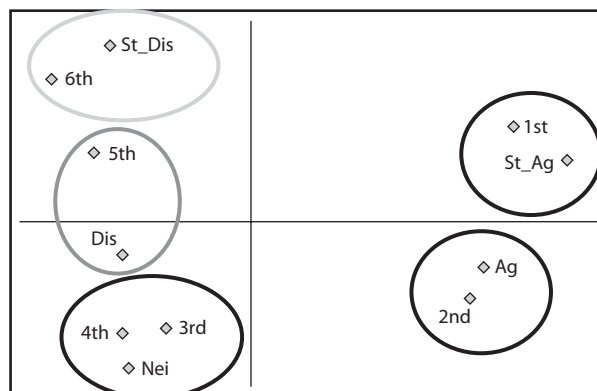
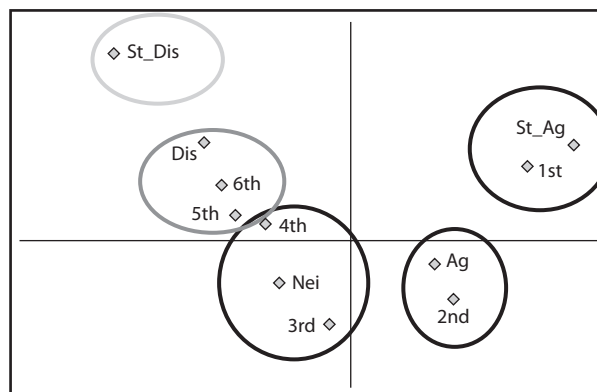
The sample size for this analysis consisted of 62 respondents; however, the analysis of these data was conducted at the attribute level in which all brand responses were combined for each attribute. This aggregation of responses created a 'virtual' sample of 372 as each respondent features six times in the data (once for each brand).

Figure 1 shows the correspondence analysis plots for the attributes *local* and *secure* for all brands. The key difference between the two techniques appeared to be at the lower end of the scales, where many attributes did not show the clear relationship demonstrated in the *local* example.

Table 4 shows the relationship between the ranking and rating techniques for each attribute as revealed by correspondence analysis. The column headings show the rating scale points, with the body of the table showing the ranking closest to the rating point. In several cases, the same ranking point is associated with more than one rating point.

In all cases there was a consistency between the techniques at the higher (positive) end of the score spectrum. The 1st and 2nd rankings were consistently related to 'strongly agree' and 'agree'. There was less agreement with the 3rd ranking, which tended to be associated with 'agree' or 'neither agree nor disagree', while the brand ranked 4th tended to be linked to 'neither agree nor disagree'. Finally 'strongly disagree' was not typically associated with any ranking point. It appears that the ranking method forced the respondent to discriminate between some brands where they might have naturally preferred to give a neutral or mildly (i.e. disagree) negative response. A low score on a ranking scale cannot be interpreted as a rejection of that attribute. Overall, H2a is supported by

Comparing methods of brand image measurement

Local**Secure****Figure 1** Ranking/rating correspondence analyses**Table 4** Rating and ranking scale comparison

	Strongly agree	Agree	Neither	Disagree	Strongly disagree
Local	1st	2nd	3rd, 4th	5th	6th
Secure	1st	2nd	3rd, 4th	5th, 6th	–
Interesting	1st	1st	2nd, 3rd, 4th	5th	6th
Contemporary	1st	2nd, 3rd	4th, 5th	6th	–
Advanced	1st	2nd	3rd, 4th	5th, 6th	5th, 6th
Handles well	1st	2nd	3rd, 4th, 5th, 6th	3rd, 5th, 6th	–
Distribution	1st	2nd, 3rd	4th, 5th	6th	–
Good value	1st	2nd	3rd, 4th, 5th, 6th	3rd, 4th, 5th, 6th	–
Safe	1st	2nd, 3rd	4th, 5th, 6th	4th, 5th, 6th	–
Trust	1st	2nd, 3rd	4th	5th, 6th	–

these data with the caveat that lower ranks do not necessarily correspond to low ratings.

As correspondence analysis requires at least three scale points, a different approach was needed for testing the remaining two hypotheses involving the pick-any measure (Bendixen & Sandler 1994). Our approach was to compare the mean rank and rating from those who selected the brand as associated with the attribute with those who did not. The greater the difference between the two, the greater the individual-level agreement between measures. If respondents' use of the two techniques was unrelated, the difference in the means between the two groups should be randomly divided between positive and negative. Further, it could be expected that 1 in 20 of the difference in means would be statistically significant at the 0.05 level due to chance alone.

Table 5 shows the analysis for pick-any and the rating technique for the attribute *advanced*. The difference between those who picked the brand as associated with the attribute and those who did not was always positive, as indeed it should be if the techniques were used consistently. In addition, the test statistic for the significance test is reported. Identical analyses were conducted to assess the relationship between the pick-any and ranking techniques.

For each of the two comparisons (pick-any with each of the other two techniques) 60 comparisons of means were conducted; these consisted of the ten attributes by six brands (of which six are shown in Table 5). Table 6 shows the average difference and the number (out of 6) of significant differences ($p \leq 0.05$) for each attribute for both the ranking and rating techniques, and was developed from the analyses demonstrated in Table 5.

In all cases, the average difference between those who indicated an association with the pick-any technique and those who did not was positive. This means that those who indicated an association also tended

Table 5 Pick-any and rating comparison of means, for the attribute *Advanced*

	Picked	Not picked	Difference	Statistical significance
Brand 4	3.1	1.9	1.3	0.03
Brand 5	3.0	1.8	1.2	0.01
Brand 6	3.3	2.0	1.2	0.00
Brand 2	3.0	2.2	0.8	0.00
Brand 3	3.2	2.5	0.7	0.23
Brand 1	3.3	2.7	0.5	0.01
Average difference			1.0	
No. significant ($p \leq 0.05$)				5

Comparing methods of brand image measurement

Table 6 Testing H2b and H2c

	Pick-any and ranking (H2b)		Pick-any and rating (H2c)	
	Average difference	No. significant out of 6 ($p \leq 0.05$)	Average difference	No. significant out of 6 ($p \leq 0.05$)
Advanced	1.8	5	1.0	5
Interesting	1.3	5	0.5	3
Modern	1.2	4	0.6	4
Local	0.6	2	0.2	3
Handles well	1.2	4	0.5	4
Distribution	0.9	4	0.6	5
Good value	1.3	4	0.7	5
Secure	0.8	2	0.5	3
Safe	1.2	3	0.5	2
Trust	1.4	5	0.7	5

to give a higher (more positive) score on either the ranking or rating techniques than those who did not. This suggests that respondents tended to be consistent in their use of the techniques in a manner that might be expected if the techniques were measuring the same latent construct. Not only was the relationship in the right direction, but the difference in means between the two groups was also generally significant.

On examining the ranking data further (H2b), there were 38 differences (out of 60) in means that were significant at the 0.05 level, which was considerably more than would be expected by chance alone. The average rank given to a brand by respondents indicating an association using the pick-any technique was 3rd, compared to the average rank of 4th from those who did not indicate an association. Of the 60 tests of difference, only one negative difference was observed. This was due to the attribute (Local) not being associated with the particular brand by any respondent using the pick-any technique, thus the average rank for the picked group was 0, and the average rank of 5th by those who did not indicate an association resulted in a negative difference. These findings suggest that respondents used the ranking and pick-any techniques in a consistent fashion. Respondents who indicated an association between brand and attribute using the pick-any technique tended to give that association a higher rank than those respondents who did not indicate an association. This supported H2b.

The testing of H2c revealed 39 differences (out of 60) in means that were significant at the 0.05 level. The average rating score given by respondents who indicated an association with the pick-any technique was

'agree'. The average rating score for those who did not indicate an association was 'neither agree nor disagree'. Two negative differences were observed (both for the same attribute, Local), again due to very few (none and 2 respondents) associating the brands with that attribute using the pick-any technique. Therefore H2c is supported in that a respondent who indicated a brand was associated with an attribute tended to give that brand a more positive score using the rating technique than those who did not indicate such an association. This relationship between pick-any and rating also supports speculation (e.g. Barnard & Ehrenberg 1990) that failure to indicate an association using the pick-any technique cannot be classed as a rejection of the association, as those not indicating an association tended to give the rating 'neither agree nor disagree'.

Time taken to administer

It is often noted, though so far not tested, that the pick-any technique is quicker to administer than either of the ranking or rating techniques (e.g. Joyce 1963; Hoffman & Franke 1986; Barnard & Ehrenberg 1990). As we used CATI surveying tools in this research, we were able to quantify this anecdotal evidence. Each question (i.e. responses for a single attribute) was assigned its own 'page' in the CATI software. In addition, the initial explanation of the brand image measurement technique was also assigned its own page. Each question was timed automatically by the computer from when the question 'page' was first displayed on the computer screen to when the final response was selected. The time taken, therefore, consisted of the reading of the question (or attribute) and any explanation relevant to the question, and the time taken for the respondent to provide their response(s) and the interviewer to ensure that this was the correct response, if necessary.

All responses for the brand image questions were designed to be point-and-click-type response sets, such that differences were not a by-product of the interviewer having to type in information rather than click on a response. The actual questions did not vary between the differing measurement techniques, though obviously the explanation of each technique did (and therefore a greater or lesser time is required to explain them to the respondent). As the questions themselves did not vary, total time required to answer the brand image questions for the three techniques consisted of (1) the fixed amount of time spent asking the questions, and (2) a varying amount of time answering according to the requirements of the technique itself. Thus, most of the variance in times (if any) between

Comparing methods of brand image measurement

Table 7 Results per attribute (total time, seconds)

	Car market		Financial services market	
	Measurement only	Per attribute	Measurement only	Per attribute
Pick-any	132	15	77	15
Rating	288	32	148	30
Ranking	278	31	135	27

the three techniques in terms of answering the questions can be attributed to using the different brand image measurement techniques.

At 20 seconds, the rating scale was quicker to introduce than either the pick-any or ranking techniques (40 seconds each), which required respondents to record the six brands (time to introduce is not shown in Table 7). However, this advantage disappeared when the measurement time was compared (see Table 7); pick-any (130 seconds for attribute measurement) was 40–45% quicker than the other techniques.

Thus while the rating technique was the fastest to describe/introduce, the pick-any had a quicker question administration time and was thus quicker overall. To test if this considerable difference was an aberration, we replicated the research in the financial services market, with five brands and five attributes. Once again, the rating technique was the fastest to describe/introduce, and the pick-any to administer. The total time, however, to collect the data was 30–34% quicker for the pick-any (115 seconds compared to 160 seconds for rating and 175 for ranking, including introduction). The smaller number of attributes meant the time saving was not as substantive. Overall, the pick-any was quicker to administer.

Dividing the time taken for measurement by the number of attributes revealed that, on average, an attribute took 15 seconds to measure using the pick-any technique, while the other two approaches took twice the time at about 30 seconds on average in both of the tests. This average time taken to administer each attribute test was very consistent between the two studies.

Discussion

The three brand image measures examined in this research (rating brands, ranking brands, and a free-choice, pick-any approach) were all found to provide equivalent results at both brand and individual level. This was an

extension of past research applied to a different market (and indeed type of market), country and by independent researchers. This is an important validation of prior findings. At the brand level, brand hierarchies were similar and users gave more positive results than non-users (both findings consistent with the prior research).

The major extension in this research was the analysis at the individual level. This demonstrated that individuals were consistent in their use of the techniques within the context of variation under retest conditions. Despite the overall similarity of the approaches at individual level, the analysis did uncover some important differences between the measures. It was evident that asking respondents to assign unique rankings to brands created artificial differences between brands in the lower rankings. It appears respondents were more capable of deciding which brands were more associated with an attribute than they were of saying which brands were less associated with an attribute. Allowing respondents to choose which brands are ranked rather than requiring all brands to be ranked could address this. We would recommend anyone employing this approach to do so to increase the validity of ranking results.

The results confirm prior findings, which suggests the three measures are relatively interchangeable; however, the additional measurement of time taken to administer the different approaches revealed a significant difference. The free-choice, pick-any approach was about 50% quicker to administer than the other two approaches, a difference that was replicated over two studies. This suggests that the pick-any approach is more efficient, and the time gains are so great that it may be cost-effective to undertake conversion if another approach is being used.

Limitations and future research

The key limitations of this research relate to the individual-level analysis. While efforts were made to minimise the potential contamination due to repeat interviews at individual level, this is still a possibility. Therefore replication of this research at individual level would be beneficial to extend and provide further confidence in the results. Researchers could also interview respondents post-survey and explore their thoughts when using the different techniques. While there was one attribute that was shown to be descriptive over the three different approaches – and this is a useful validity test – replication across a wider range of descriptive attributes would be useful, particularly if the descriptive attributes were identified prior to the study.

Future research should also encompass a comparison between other types of brand image measures, and analyses of results should include multivariate analyses. This is an important area to ensure this fundamental research stream keeps abreast of the uses to which brand image data are put.

References

- Axelrod, J.N. (1968) Attitude measures that predict purchase. *Journal of Advertising Research*, 8, 1, pp. 3–17.
- Barnard, N.R. & Ehrenberg, A.S.C. (1990) Robust measures of consumer brand beliefs. *Journal of Marketing Research*, 27, November, pp. 477–484.
- Barnard, N.R., Barwise, P.T. & Ehrenberg, A.S.C. (1986) Reinterviews in attitude research: early results. MRS Conference, Brighton: Market Research Society.
- Barwise, T.P. & Ehrenberg, A.S.C. (1985) Consumer beliefs and brand usage. *Journal of the Market Research Society*, 27, 2, April, pp. 81–93.
- Bendixen, M. (1991) *Correspondence Analysis*. Johannesburg: University of Witwatersrand.
- Bendixen, M. & Sandler, M. (1994) *Converting Verbal Scales to Interval Scales Using Correspondence Analysis*. Johannesburg: University of Witwatersrand.
- Bird, M. & Ehrenberg, A.S.C. (1970) Consumer attitudes and brand usage. *Journal of the Market Research Society*, 12, 4, pp. 233–247.
- Castleberry, S.B., Barnard, N.R., Barwise, T.P., Ehrenberg, A.S.C. & Dall’Olmo Riley, F. (1994) Individual attitude variations over time. *Journal of Marketing Management*, 10, 1–3, pp. 153–162.
- Dall’Olmo Riley, F., Ehrenberg, A.S.C. & Barnard, N.R. (1998) Changes in attitudes and behaviour. Presented at the 27th European Marketing Academy Conference, Stockholm School of Economics, Stockholm.
- Dall’Olmo Riley, F., Ehrenberg, A.S.C., Castleberry, S.B., Barwise, T.P. & Barnard, N.R. (1997) The variability of attitudinal repeat-rates. *International Journal of Research in Marketing*, 14, 5, pp. 437–450.
- Dillon, W.R., Maddern, T.J., Kirmani, A. & Mukherjee, S. (2001) Understanding what’s in a brand rating: a model for assessing brand and attribute effects and their relationship to brand equity. *Journal of Marketing Research*, 38, 4, pp. 415–429.
- Driesener, C. (2002) *Comparing and Contrasting Three Brand Image Measures*. Marketing Science Centre, University of South Australia, Adelaide, p. 179.
- Driesener, C. & Romaniuk, J. (2001) *Brand Image? Pick a Measure, any Measure*. Australian & New Zealand Marketing Academy, Department of Commerce, Massey University, Palmerston North.
- Ehrenberg, A.S.C. (1992) Report writing – six simple rules for better business documents. *Admap*, June, pp. 39–42.
- Greenacre, M. (1984) *Theory and Applications of Correspondence Analysis*. London: Academic Press.
- Greenacre, M. (1993) *Correspondence Analysis in Practice*. London: Academic Press.

- Haley, R.I. & Case, P.B. (1979) Testing thirteen attitude scales for agreement and brand discrimination. *Journal of Marketing*, 43, Fall, pp. 20–32.
- Herman, S. (1991) *MCA+*. New Jersey: Bretton-Clark.
- Hoffman, D.L. & Franke, G.R. (1986) Correspondence analysis: graphical representation of categorical data in marketing research. *Journal of Marketing Research*, 23, August, pp. 213–227.
- Holbrook, M.B., Moore, W.L. & Winer, R.S. (1982) Constructing joint spaces from pick-any data: a new tool for consumer analysis. *Journal of Consumer Research*, 9, June, pp. 99–105.
- Holmes, C. (1974) A statistical evaluation of rating scales. *Journal of the Market Research Society*, 16, 2, April, pp. 87–107.
- Jacoby, J. & Matell, M.S. (1971) Three-point Likert scales are good enough. *Journal of Marketing Research*, 8, November, pp. 495–500.
- Joyce, T. (1963) *Techniques of Brand Image Measurement. New Developments in Research*. London: Market Research Society, pp. 45–63.
- Kalwani, M.U. & Silk, A.J. (1982) On the reliability and predictive validity of purchase intention measures. *Marketing Science*, 1, 3, Summer, pp. 243–286.
- Keller, K.L. (1993) Conceptualizing, measuring, and managing customer-based brand equity. *Journal of Marketing*, 57, January, pp. 1–22.
- Kennedy, R., Riquier, C. & Sharp, B. (1995) Putting correspondence analysis to use with categorical data in market research. Academy of Marketing Science World Marketing Congress, Melbourne.
- Likert, R. (1932) A technique for the measurement of attitudes. *Archives of Psychology*, 40, pp. 44–53.
- Myers, C.A. (2003) Managing brand equity: a look at the impact of attributes. *Journal of Product & Brand Management*, 12, 1, pp. 39–51.
- Riebe, E., Lowndes, M., Kennedy, R. & Romaniuk, J. (1998) Verbal versus ordinal scales in the measurement of service quality and switching estimates. Australian and New Zealand Marketing Academy Conference 98, University of Otago, Dunedin, New Zealand.
- Romaniuk, J. (2005) *Comparing Measures of Brand Personality Traits*. Italy: European Marketing Academy.
- Romaniuk, J. & Nicholls, E. (2006) Evaluating advertising effects on brand perceptions: incorporating prior knowledge. *International Journal of Market Research*, 48, 2, pp. 179–192.
- Romaniuk, J. & Sharp, B. (2000) Using known patterns in image data to determine brand positioning. *International Journal of Market Research*, 42, 2, Spring/Summer, pp. 219–230.
- Rungie, C., Laurent, G., Dall’Olmo Riley, F., Morrison, D.G. & Roy, T. (2005) Measuring and modeling the (limited) reliability of free choice attitude questions. *International Journal of Research in Marketing*, 22, 3, pp. 309–318.
- Sharp, A. (2002) Searching for boundary conditions for an empirical generalisation concerning the temporal stability of individuals’ perceptual responses. PhD dissertation, School of Marketing, University of South Australia, Adelaide.
- Winchester, M. & Romaniuk, J. (2003) Evaluative and descriptive response patterns to negative image attributes. *International Journal of Market Research*, 45, 1, pp. 21–34.

About the authors

Carl Driesener is a Senior Research Associate with the Ehrenberg-Bass Institute for Marketing Science. His research interests include perceptual measurement, and modeling buyer behaviour. He is particularly interested in the measureable affects of advertising on behaviour.

Jenni Romaniuk is a Senior Research Fellow with the Ehrenberg-Bass Institute for Marketing Science and heads the Brand Equity research stream. Her research combines interests in branding and consumer cognitive processes with the development of empirical generalisations to provide context and meaning to marketing metrics. Jenni is also Editor of the Journal of Empirical Generalisations in Marketing.

Address correspondence to: Carl Driesener, Ehrenberg-Bass Institute for Marketing Science, University of South Australia, GPO Box 2471, SA 5001, Australia.

Email: Carl.Driesener@marketingsciencecentre.com