Brand attribute lists are commonly used as a research survey tool. Despite their prominence, it is widely acknowledged that they can result in the systematic distortion of data due halo effects, and order bias. This paper sets out to understand the nature and extent of these effects, and examine whether randomisation rectifies the problem. We test a number of alternative formats to the brand association question, using an online web panel survey. Results show that while order bias and the halo effect are not eliminated by any alternative approach, limiting respondents’ answer responses can reduce bias if setup correctly.
Section I

1.1 Introduction

One of the most widely used questions in marketing research is the brand attribute association. A typical example would look something like the following:

Q: Displayed is a list of statements that may or may not be used to describe a laundry detergent. Please indicate the brand(s) you feel are best described by each statement. You may select as many or as few brands as you like for each statement.
   A laundry detergent that…
   • Is a fun brand
   • Gives me value for money
   • My mother used to use
   • Gives me confidence in how I look

The aim of this conventional approach to providing brand equity diagnostics is to establish what brand characteristics are important to people in a market; and measure perceptions of brand performance on a range of potential brand characteristics, including those that are currently deemed to be important.

Strategy then involves maintaining or increasing brand performance on current, key attributes; or identifying brand strengths that may not be important at the moment, but that could be leveraged if their importance could be increased. An additional strategy is to try to identify ‘white space’ i.e. attribute territories that are unoccupied, but that could lead to brand success if they were claimed by the brand.

In all of this there are three key questions: first, what should the content of the potential explanatory factors be? Second, what form should the questions take? And third, what mathematics should form the basis for judgements as to relative brand strength; and for identification of what’s important to people1?

In developing the brand attribute question that goes into a survey, the most common approach worldwide is to:

a) Do some kind of in-depth, usually qualitative research to identify potential explanatory factors
b) Draw up a list of factors or attributes to be used in quantitative research
c) Present the list to respondents in the form of an attribute association

List development is usually informed by past research; and quite often informed by both the client’s and the supplier’s theories about what causes brand attachment. Numerous list frameworks exist e.g. Keller (functional performance, emotional engagement, brand personality, etc.); Y&R (differentiation, relevance, esteem, knowledge); frameworks that come from depth psychology (Censydium, NeedScope); etc. In addition, equity lists often include attributes like ‘quality’, ‘value for money’, ‘trust’, ‘reliability’, and so on.

The Keller items recognize that brand attachment can be based on layers – from the functional and tangible, to the more emotional, intangible, and self-expressive.

1 For the purposes of this paper, we will be focussing on the second question, what form should the question take?
Regardless of the nature of attributes however, is the assumption that brand attribute association questions work to produce robust and valid outputs that clients can rely upon to develop future strategy. As researchers, it is our role to scrutinise the validity of this data, and understand to what extent (if any) the structure of the question and answer format influences the resulting output.

To date, the most widely acknowledged recurring problems that affect answer formats that utilise lists are the halo effect, and position bias (or order effects). By ‘position effects’ we mean the impact that the position in the list of an attribute has on its potential to be picked. By ‘halo effects’ we mean the tendency for respondents to attribute everything to their current, main brand, and not to give much attention to describing other brands. As we have noted, this has a direct bearing on the value of research because it calls into question the very conclusions that we draw, and subsequently report to clients.

There have been a large number of studies that have examined the nature and size of position, or order bias in survey responses. The results of these studies overwhelmingly demonstrate that a respondent’s choice of an answer in a list of alternatives depends on the position of the answer in relation to other alternatives, as well as on the content of that answer (Blunch, 1984). A commonly referenced paper with regards to position bias by Payne (1951) reports that an answer option received six percentage points more on average when placed among the top alternatives, and two percentage points when placed at the bottom alternatives in comparison with the same alternative near the middle of the list. In addressing this problem, researchers have randomised the position that responses occupy within a list in the hope of cancelling out preferences for the various positions. This commonsensical response has been widely accepted and adopted: “This problem can be solved in most cases by rotating the order of alternatives” (Green, Tull, 1978 p.123); “The researcher may be reasonably sure that the possible effects of position have been fairly well cancelled by this approach” (Payne, 1951 p.85).

Despite the confidence shown in randomisation as a technique, researchers have found that it does not eliminate the problem of order bias when tested. Without going into detail about each and every study that has tested the affects of randomisation in lists, suffice it to say that they are mostly uniform in methodology, varying chiefly in the conclusions about the possible size of the bias2 (Blunch, 1984 p.216).

While the work done on understanding position bias has covered popular survey response formats such as multiple choice questions (Blunch 1984), pairwise comparisons (Day 1969), two response type tests amongst children (Mathews (1927), multidimensional scales (Jain and Pinson 1976), paired product tests (Day (1969) negative image attributes (Winchester and Romaniuk 2003), and likert scales (Chan, 1991), no empirical study has been done to examine the effects of position bias in brand association lists. More notably, we are not aware of any comparative tests that have been published using a set of alternate formats for brand association questions to gauge if any approach addresses or eliminates the problem of order bias, or the halo effect.

Based on this gap in the literature, the research we conducted aims at answering the following questions:

1. Does the position of an attribute in a brand association list affect the likelihood of it being picked, and if so, what is the extent of this? This will be tested in two separate questions, one consisting of a shorter list of 8 "personal" attributes, and another consisting of a longer list of 18 "functional" attributes.

2 Some notable studies here are those by Carp (1974), Ring (1975), and Shuman and Presser (1981), Chrzan and Golovashkina (2006)
2. Does randomisation average out the effects of order bias by allowing every attribute to "suffer equally"?
3. What are the underlying causes of position bias?
4. What are the alternative approaches to attribute association questions?
5. Do any alternative approaches reduce or eliminate the effects of order bias or the halo effect?

In order to address these questions an online survey was launched using a standard usage and attitude format in the laundry detergents category. It included such questions as brand awareness (first mention, aided), brands used (past 6 months, last ten purchases, most), etc.

The questionnaire included two attribute associations to measure brand image. The first consisted of eight attributes; the second, seventeen. Respondents were asked to rate all brands of which they were aided aware. On average, respondents were aware of 10.1 brands (from a list of 16).

The sample of 4000 was then split and routed through to various formations of the attribute association question. Quotas were set at 1000 for each of the 4 different cells, each cell answering a different format of the question. While the structure of the question differed between cells, the underlying content did not. The answer layout remained the same (brands across the top, attributes down the left hand side) for all cells, which conforms to the standard association matrix grid commonly used in research surveys. The main difference between the 4 cells was how respondents were asked to answer the question. The question formulations, and relevant cell routings were as follows:

- A standard attribute association battery consisting of both functional and self-expressive attributes
- The same battery but respondents are asked to pick at least two but not more than five for each brand that’s relevant to them (so-called ‘pick and choose’)
- The same battery but respondents are first asked to identify ‘attributes they cannot do without’; and then rate brands only on those attributes
- A similar battery, but respondents are asked to list the most important attributes and then rank each relevant brand in terms of its performance

We used the standard approach to minimizing order effects, that is, we randomized the order of attribute presentation at respondent level. The online script would record the number of associations that an attribute and brand received, as well as the number of associations a position in the list received, regardless of the content of the attribute.

The study was conducted in the USA amongst males and females, aged 18 and above with no other quotas on age. All respondents had to be aware of at least one brand of laundry detergent. The intention in recruiting the sample was to obtain as representative a group as possible of target consumers for laundry detergent brands. Respondents were recruited from a web panel managed by legacy Synovate Viewsnet (now Ipsos).

---

3 For a list of the attributes used, as well as a full breakdown of the question wordings and routing please see Appendix 2
Section II

2.1 Results and discussion: an analysis of the standard approach

Firstly, the results of our analysis show that randomisation aside, order effects are prevalent in both attribute lists we tested. Figure 1 below shows that, as expected, the higher an attribute appears in a list, the more likely it is to be picked. The pattern also demonstrates that the probability that an attribute gets chosen decreases as it moves down the list. This propensity doesn’t appear to be much affected by the length of the list – the drop in associations was worse when there were eight as compared to eighteen attributes.

![Figure 1: The inverse relationship between where an attribute appears in the attribute list, and how likely it is to be ticked.](image)

The main message is that position effects are real and substantial. On average attributes lose up to one-third of their associations depending on where they appear in an attribute list. By far the biggest effect seems to be which attribute appears first: the gap between the first and second positions is bigger than the gap between the second and the last. The effect seems to arise even in fairly short lists, and in our data was of a similar size using 18 or 8 attributes.

The most likely candidate for this is respondent fatigue, or to be more particular, a decrease in due diligence as the task set for the respondent wears on. A respondent might start off enthusiastically but quickly tire and only tick those associations that immediately spring to mind. Another possibility is that respondents become more discerning as they move down the list. Initially they may give the benefit of the doubt to a brand that “almost” possesses an attribute. As they move down the list and encounter more and more attributes, they may tighten up their definition of what size association deserves a tick. This need not be a conscious process. We’ll investigate the merits of these explanations in more detail below.

2.2 The role of brand and attribute: why randomisation doesn’t work

Despite order effects being real and substantial as evidenced above, surely the randomisation of the list would off-set potential losses and gains that result from the attributes appearing in the various positions? In other words, every attribute would suffer the effects equally, and the distortive effects of order bias would even out. In order to look into this in more detail, we need to take into account the role of the brand, and the nature of the attribute.
We first examine whether the size of a brand affects how much an attribute is likely to be associated with it. There’s no reason to expect any effect – after all, it’s the attributes that are changing their position, not the brands. But Figure 2 shows that the potential impact on a brand when attributes switch their position does depend on the size of the brand. The graph shows the average absolute change in the proportion of ticks a brand gets when attributes are changed from their most favourable to least favourable position. Bigger brands – brands with higher penetration – tend to experience much bigger changes than smaller brands.

![Figure 2: Relationship between brand size and how much it is affected, on average, by changes in the attribute list. To calculate the size of the position effect, we take the absolute difference between the proportion of ticks each attribute got in its most favourable position (always first in the list) and the proportion of ticks it got in its least favourable position (interestingly, near but not at the end of the list – see Note #1). Doing this separately for each brand and averaging over all attributes gives the average absolute changes for each brand. The x axis shows the size of the brand calculated by the amount of the sample who ticked attributes for that brand. The left- and right-hand plots show results for the 8 “personal” attributes and 18 “practical” attributes respectively.](image)

The figure above is misleading though. In attribute associations, bigger brands tend to get more ticks than small brands (an example of an effect called “double jeopardy”); and they experience bigger changes because they start out with more ticks. An absolute change of 5%, for example, means much more to a small brand that only starts out getting ticked by 6% of respondents than a big brand that gets ticked by 70% of respondents. So we need to also look at relative change, which we do by scaling the absolute differences in Figure 2 by the maximum number of ticks a brand gets on that attribute. The relative changes, which tell a very different story, are shown in Figure 3.

Both plots in Figure 3 clearly show that small brands are more affected by positional changes in the attribute list than big brands. If an attribute is switched from best to worst position, the biggest brand loses around 25% of its ticks. A much smaller brand, say one placed 10th out of 14, loses twice that amount – it loses half of its ticks. So, respondents are effectively ignoring smaller brands once they get towards the bottom of an attribute list.
Figure 3: Relative changes in the proportion of ticks a brand gets, shown as a function of its size (penetration). See caption to Figure 2 for more details.

Another way to examine this is by looking at the difference in the amount of associations (or ticks) an attribute gains when moving from least favourable, to most favourable place, as a function of type of brand for the respondent. By ‘type’ we mean: was the brand the respondent’s main brand, or a brand bought in the past six months but not main, or a brand of which the respondent was merely aware? Brands with a high penetration (x axis on the graphs above), or larger brands, tend to be brands that have a higher proportion of users, as well as a relatively higher percentage of users which use the brand as a main brand compared with smaller brands, so we expect the pattern of smaller brands being affected more than larger brands when analysing the data by brand type. From Figure 4 below, we can see that this in indeed the case: larger, more used brands experience much less dramatic order effects than smaller brands, seen by the difference in size of the green vs. blue and red bars.

---

4 Brand type is established by using the questions included in the survey: Main brand was respondent’s self-identified “used most often brand” out of all brands they currently use. The remaining brands make up “all use, not most often”, and brands not identified at the usage question fall into the “aware not use group”.

© SAMRA 2013 – Conference Delegate Copy 7
In order to provide an explanation for this, we need to think of brand type, based on usage as surrogate for brand importance – the more a brand is used (particularly brands used most often) the more important it tends to be to a person. If we suppose that there is a relationship between a brand’s importance to a person and that person’s willingness to answer questions about the brand, then the observed pattern becomes explicable: respondents are more diligent when it comes to their most used (‘use’ as a surrogate for ‘importance’) brands. This means that the performance of small brands on attributes appearing in the lower part of a list will tend to be understated. This might be especially problematic if a brand occupies a small niche and those niche attributes are located far down the attribute list. In that case the brand may appear to be much worse off than it actually is. If the small brand is the client, they may be disappointed at their apparent lack of progress, while a big client brand may underestimate the threat posed by its smaller competitors.

2.3 The role of attribute type: how different attributes gain or lose, and the relation to brand type

We’ve seen that respondent diligence declines as the attribute association task wears on, but that it does so differentially as a function of the relationship they have with each brand. The less a brand is used, the more it is affected, which results in smaller brands losing out more than larger, more popular brands. This is a matter of the relationship between brand ‘type’ and diligence. Is there, in addition, a relationship between attribute ‘type’ and diligence? In particular, does the probability of selection vary as a function of an attributes overall incidence?

As in Figure 4, we’ve plotted the impact on a brand when attributes switch their position in a list. We’ve grouped the 4 biggest brands together in one group, and the next 4 biggest in another. We know that the effects for the group of smaller brands should be bigger than the group of bigger brands, and indeed this pattern shows up clearly. What we ask now is whether the size of the effect (that is, the height of the blue bar relative to the red bar) is different for different attributes.

Figure 4: Ranked by most popular (most ticked overall) to least, the x axis shows the percentage of associations gained when an attribute shifts from least favourable to most favourable position in the list. The differences in the extent of the gain is broken down by brand type, green being brands used most often, blue used but not used most often, and red aware of but not used.
Figure 5: Ranked by most popular (most ticked overall) to least, the x axis shows the percentage of associations gained when an attribute shifts from least favourable to most favourable position in the list. The differences in the extent of the gain is broken down by brand size, blue being brands ranked 1-5 in terms of penetration (top 5), and red brands ranked 11-15 (bottom 5).

One might suppose that the more an attribute is chosen for association, the more important it is. However, while there is likely to be a co-incidence between attribute incidence and importance, we can’t be certain. We therefore use a more descriptive phrase, namely, ‘attribute popularity’. Put simply, an attribute’s popularity is reflected by the sheer number of times it’s associated with the brands. This may be because brand managers believe it’s an important attribute (what marketers often call ‘cost of entry’). It may reflect the importance consumers attach to it. Whatever the case, it does reflect ‘popularity’.

One idea is that more common “hygiene” attributes may show bigger effects than “niche” attributes, because they’re more sensitive to respondent fatigue. For this reason, in Figure 5 we’ve plotted the attributes from most common at the top to least common at the bottom, and given the proportion of the sample that ticked each attribute. In the personal attributes list, three of the four largest big-brand biases occur in the four most-common attributes (the exception is Keeps the family looking good), while three of the four smallest biases occur for the four least-common attributes (the exception is Fun). For the practical attributes, the three largest big brand biases occur towards the top of the list for (Works well, Reliable, Cleans well) which can be categorised as entry level, or hygiene factors for any washing powder brand to have. More niche statements like For sensitive skin, and to a lesser extent More washes are least affected by brand size. So while the evidence is mixed, there is some suggestion that big brands gain relatively more on common attributes than on niche ones, and smaller, less used brands lose out on hygiene attributes the lower they appear on the list.

Put another way: all attributes continue to be associated with larger brands which are more likely to be used, or main brands, no matter what the order of attribute presentation is. This is so for both the most and the least popular attributes. Even so, associations with the most popular attributes decline more than they do with the least popular attributes. When it comes to smaller brands of which respondents are more likely to be aware of, but not users of, popular attributes are associated at the start of the process, but not by the end. This pattern holds for less popular attributes, but the relative change in proportion of associations is less pronounced than with popular attributes.
2.4 Results and discussion: a comparison of alternative approaches to the brand association question

If the standard attribute association question format results in the systematic distortion of data due to the reasons we’ve outlined above, despite randomisation, what are the alternatives, if any?

As outlined in the sections above, we tested different approaches to the brand attribute association question which kept the core structure and objective of the question in tact: the respondent was required (as per standard approaches) to associate brands with relevant attributes that appear in a list, with brands appearing across the top in a grid formation. However, the form of how respondents went about answering the question was different for each. To recap, the respondents in the first cell were limited to ticking between two (minimum) to five (maximum) attributes across all brands. Respondents in the second alternative cell were asked to identify which attributes were important to them from the total list shown for the standard, and the two to five cells. These selected attributes were then routed through to the association grid answer format which the respondent then associated with their relevant brands. The same format was used for another cell, but instead of being asked to choose which attributes were important to them; respondents were asked which attributes they could not live without. These selected attributes were then routed through to the association grid answer format.

Before we examine the results of the data analysis, the average timings for the various approaches appear in the table below.

<table>
<thead>
<tr>
<th>List type</th>
<th>Standard</th>
<th>2-5</th>
<th>Direct(^6)</th>
<th>Indirect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal attributes</td>
<td>1:38</td>
<td>2:41</td>
<td>1:16</td>
<td>1:14</td>
</tr>
<tr>
<td>Practical attributes</td>
<td>2:02</td>
<td>3:00</td>
<td>2:02</td>
<td>2:00</td>
</tr>
</tbody>
</table>

From the outset we can see that no approach performs significantly better in terms of reducing questionnaire length compared with the standard approach. What’s interesting is that the 2-5 approach on average takes respondents considerably longer to complete. The could be due to the fact that, given the limits on the associations, or ticks they have available to them, respondents spend longer deciding how these should be distributed. If this is indeed the case, then this should be noticeable in the data output with potentially greater differentiation between attributes and less prominent order bias.

Turning to the data, the results for each approach appear below. For consistency’s sake, we’ve analysed the data in the same way as with the standard approach in order to identify any differences in their respective outputs.

\(^5\) For ease of reference the above approaches will be referred to as 2-5, direct importance and indirect importance.

\(^6\) The timings for the direct and indirect approach are the combined timings for both questions required: first, selecting which attributes are important/or respondents’ can’t live without, and the following brand attribute association question.
We’ll start with the personal attributes on top, followed by the practical attributes below:

Figure 6: The inverse relationship between where an attribute appears in the attribute list, and how likely it is to be ticked is evident across the various alternative approaches.

While there are discernable differences between the approaches, the selection bias towards attributes at the top of the list is common to all. The number of associations of attributes diminishes the further down the list an attribute appears in, consistent with the standard approach analysed earlier.

While the 2-5 approach appears to do the best at evening out the probability of selection for the personal attributes, it does the worst out of all the approaches for the longer functional attributes list. In our opinion, this might be a function of the limitation imposed in relation to the number of attributes in the list. The 2-5 attributes limitation works well with the shorter list, where there are a sufficient number of ticks to distribute amongst the statements and brands. With the longer list, respondents seem to stick to the task at hand at the beginning but reach the maximum of 5 ticks across their respective set of brands before they get to the bottom of the list. This is why the probability of selection falls so sharply the further down the order an attribute appears in, as respondents have no more ticks available to associate brands with. The result is a decline in the probability of selection for the 2-5 approach for the longer list which is far more pronounced than for the other approaches.

The direct and indirect importance approaches prove more difficult to analyse, as the number of attributes that appear for each respondent will depend on how many they identify at the previous importance question. In order to keep analysis consistent with the other approaches, we changed the x axis for "Position in attribute list" to a number from 0 to 1, to be interpreted as the relative position in the rank order (rather than a true "position); 0 being first place, 1 last place where the interval is split into 8 segments (for practical attributes) or 4 segments (for personal attributes).

While the same pattern of position bias is evident for both direct and indirect approaches, the indirect approach experiences the smallest changes in probability of selection across the personal
and functional list of attributes. Despite this being the case, the attributes appearing at the top of the list are still more likely to be chosen with the indirect importance approach, so there is no outright winner that eliminates order bias in response, only one that results in the least amount of data distortion due to order bias.

2.5 The role of brand size and attribute type in alternative approaches

If the alternative approaches experience the same order bias problem as the standard approach, do they perform better in other ways instead? In the initial analysis of the standard approach, we saw that brands do not all “suffer equally” when it comes to changes in proportions of ticks due to the placement of an attribute in a list. Smaller brands take the biggest hit because, although respondents consistently start reducing the number of ticks in general the further down the list they go, they tend to default to larger, more used brands and ignore smaller, less used ones.

The results below replicate the analysis that was used for the standard approach. Personal attributes appear at the top, functional at the bottom:

![Graphs showing changes in proportions of ticks for different attribute types and brand sizes.]

Figure 7: Changes in the proportion of ticks a brand gets, shown as a function of its size (penetration). See caption to Figure 2 for more details.

The results indicate a similar pattern to the standard approach where big brands receive the largest change in proportion of tickets across all approaches. What is particularly interesting about the direct approach however, is that for both functional and personal attribute lists, medium sized brands in terms of penetration experience the largest change in number of associations. This pattern is seen in the indirect evaluation approach, but in a less pronounced way. Overall, we can see that brand size has a bearing on the size of the change (proportion of ticks) which means that attributes do not all suffer equally when they appear either at the top, or bottom of the list. As with the standard approach, the size of the brand and how much it is used plays a role in how much an attribute is ticked.

The relative change in the proportion of ticks for each of the alternative approaches also bears a similarity to the pattern displayed for the standard approach.
The distributions above show that smaller brands are again more vulnerable to significant changes in relative proportion of ticks compared with large brands across all three approaches. The largest brands consistently lose about 25% of their ticks when an attribute is moved from best to worst position. However, the three smallest brands lose over 75% of their ticks. None of the alternative approaches were able to show any improvement in flattening out the big brand bias, although despite the smallest brand being affected in similar significant proportions to the other approaches, the 2-5 approach has the lowest average relative change across all brands for the personal attributes. As we saw in the graphs in the previous section (Figure 6), when limiting the number of ticks to between 2 and 5 for shorter lists of attributes, the order in which attributes appear has the least amount of impact on their likelihood to be ticked. The same is apparent with regards to the big brand bias, where the 2-5 approach results in the least amount of ticks being lost across brands (except for the outlying smallest brand), but for the personal attributes only. While this is not conclusive, it indicates that limiting ticks might be the better performing approach of the variations we tested when it comes to minimising halo effects, or the big brand bias. The question this does pose is whether the limited attributes approach would have performed better for the longer, practical attribute list if the number of maximum ticks allowed had been higher. 

2.6 The role of Attribute type in the alternative approaches

In examining the relationship between attribute popularity and brand type, we replicated the analysis performed on the standard approach again for our alternative approaches. The graphs can be found in the appendix, as they were too large and too many to make for easy display (see Figure 9). Each bar represents the percentage of associations gained when an attribute moves from least favourable, to most favourable position in the list. Attributes are ranked by popularity as before. The three alternatives tested produce different response patterns to the standard approach, where there was an observable relationship between attribute popularity and brand type. Of the approaches...

---

7 In his book, A Theory of Data, Clyde Coombs suggests that there is an optimal number of attributes to limit respondents to when collecting data, based on the length of the list concerned. Based on their empirical tests, this optimal limitation was one third of the total number of attributes (1:3).
below, the indirect and direct approaches show some similarity in that the top three popular attributes on average show less of a change in ticks gained than the bottom three for the personal statements. The relationship between brand type and attribute popularity is evident for the direct approach, where the big brand, or used brand bias occurs most prominently in the top three most popular attributes. For both the direct and indirect approaches, small brands lose out more than large brands for the most popular hygiene factors, which is all consistent with the findings for the standard approach.

However, for the practical attributes, the indirect approach shows no clear relationship between attribute popularity, brand type, and the amount an attribute gains in ticks based on a change in position. Quite interestingly, with the indirect approach, all brand types are affected equally for mid-popularity attribute statements for the practical attributes.

The 2-5 approach on the other hand produces a rather distinct and interesting set of results. The difference in ticks gained for all statements is smaller compared with the standard, direct and indirect approaches. This is particularly noticeable for the personal attributes. There is also a much greater similarity between the percentage of ticks gained across all brand types, with larger brands being associated with popular statements only slightly more than smaller brands, whereas this difference was far more pronounced with the other approaches. The 2-5 approach seems to have a different effect where more niche, or less popular statements are more likely to lose out on being associated with big brands the further down the list they go. This is particularly apparent for the practical attributes list. As we explained before, this is most likely because, regardless of the nature of the attributes, the likelihood of an tick being received at the bottom of the list by any brand is low because respondents have used up their allocation by that point. The least popular attributes do however, still experience a larger amount in ticks gained which means they are slightly more susceptible to order bias than popular attributes for the 2-5 approach. Relative to other approaches however, the 2-5 approach goes the furthest in minimising order effects, in particular the big brand bias or halo effect.

2.7 Comparing outputs for all approaches: biplot analysis

The final task we'd like to perform at this point is a comparison of the biplot analysis charts for each approach. This does not identify whether any of the approaches “performs better” as there is no way to validate this, but it does give us an idea of how much differentiation is apparent for each and how differently brands are associated with attributes. Due to their size, these graphs appear in Appendix 1, Figure 10.

For the personal statements, there is a similarity in output for the standard, and the 2-5 approach. This means that respondents go about the brand association task in a similar way (the question and format are the same, except for the limitation). For these two approaches, the largest brand, Brand 12 is associated with *A leading brand* and *It’s a family favourite my mother has always used* for both. In comparison, Brand 12 is most strongly associated with *Keeps my family looking good* for the direct and indirect approaches.

For all approaches though, Brand 1 and Brand 7, which are the second and third largest brands respectively, are closely associated with *Gives me value for money*. The remaining point of similarity across the approaches is that brands that fall into the middle of list in terms of size tend to be grouped together to the left of the vectors. The way that these are grouped and distributed however falls into two groups, the 2-5 and standard group on the one hand, and the direct and indirect group for the other. There is a slight indication that the direct and indirect approaches provide better differentiation between attributes for the personal statements. Overall, outputs show us that the 2-5 and standard approach tend to be asking the same question, and producing similar
results. The direct and indirect approach change the nature of the question slightly, and due to the smaller number of attributes being routed through to the actual association question, produce different outputs compared to the 2-5 and standard approach.

The practical attribute biplots show that the standard approach produces the smallest amount of differentiation, with the 2-5 approach producing the most, although again, it cannot be said that this is a more valid result, or whether this is better in any way. The greater discrimination is a most likely a function of the fewer amount of ticks across all brands, and gives some indication that respondents might be taking more time to associate an attribute with the relevant brand, instead of allocating greater numbers of ticks to main brands only. We must remember however, that the longer list of practical attributes was affected by order bias more than all other approaches, so despite offering greater discrimination and evidence of more thought being applied to answers, the importance of establishing an optimal upper limit for longer lists is important.

What is interesting is that two niche attributes *Is safe for sensitive skin* and *A brand that is gentle on my clothes* are far more strongly associated with two of the three smallest brands (Brand 5 and 8) in the 2-5 approach for the practical attributes.

The leading brand (Brand 12) is most strongly associated with the majority of statements, particularly the cost of entry, or hygiene factors as expected for all approaches, While Brand 7 (3rd largest) is consistently associated with *Value for money*. We again see similarity between the outputs for the direct and indirect approach due to the nature of their question/answer formats. These provide less discrimination than the 2-5, or standard approaches though, which indicates that when respondents only rate attributes that are important to them, they are more likely to associate these with all brands that they use or are aware of. There are, in other words, no unimportant or more obscure statements that might only be associated with select brands which could drive differentiation amongst the attributes.

While a comparison exercise is difficult for reasons already mentioned, all of the three alternative approaches to the standard approach provide more discrimination between attributes. In terms of their reporting output, we’ve observed a similarity between the 2-5 and standard results, and the indirect and direct approach respectively.

**Conclusion**

It’s apparent from our analysis that there are significant order effects in standard formats of the brand attribution list. In general, an attribute has a significantly greater chance of being associated with a brand when it appears higher up in the order. But we’ve also shown that the effects cannot be mitigated by randomization, because the impacts are differential as a function of both brand and attribute type. Specifically, for the standard approach:

- The more a brand is used, the less likely an attribute is to be affected by the order in which it is presented.
- The more popular an attribute is, the more likely it is that it will be affected by the order in which it is presented.

The 2-5, direct and indirect alternative approaches that we tested were all subject to the same distortions in data, to greater or lesser degrees. No approach in particular can claim to have eliminated order bias, or the halo effect on both the shorter, and the longer attribute lists. From what we have seen, order effects seem to be a manifest problem for any survey measure that makes use of a list of answer options. Taking all the alternatives into account however, there are some encouraging results that are produced by the 2-5 approach for the shorter attribute list. As we
discussed before, the limitation of attributes results in the relative evening out of the big brand bias, although this does not mean that the problem is completely overcome (see Figure 8 and Figure 9 for personal attributes). The big brand bias seems to even out when the upper limit on ticks is not considerably less than the list of statements. As we saw with the longer list of attributes, limiting respondents to 5 meant that attributes towards the bottom experienced the greatest fall in probability of selection compared with the approaches. While Coombs has suggest a ratio of 1:3 between upper limits and the total number of attributes in a list, further research is needed to determine what the optimal setup would be in order to increase discrimination, and keep order effects and brand bias at a minimum. This might entail testing different combinations of limitations, as well as different types of attribute lists in different lengths. With the available data it is not possible to say that the 2-5 approach clearly overcomes the problems that face the standard brand association question, but there are signs that it could go some way in addressing order bias and the halo effect, and potentially provide greater discrimination in strategic analysis that is provided to clients.
References


Appendix 1

Figure 9: Ranked by most popular (most ticked overall) to least, the x axis shows the percentage of associations gained when an attribute shifts from least favourable to most favourable position in the list. The differences in the extent of the gain is broken down by brand type, green being brands used most often, blue used but not used most often, and red aware of but not used.

Personal attributes

![Graph showing the percentage of associations gained for various personal attributes ranked by most popular to least. The graph includes categories like Leading brand, Proud to use, Value for money, Keeps family looking good, Cares for environment, Gives me confidence, Family favourite, and Fun. The graph is divided into three sections: MO (green), Used not MO (blue), and Not used (red).]
Indirect

Direct

% ticks gained

© SAMRA 2013 – Conference Delegate Copy
Figure 10: Bi-plots showing the analysis outputs for each of the approaches.

Personal attributes
Practical attributes
Appendix II: Question wordings used

**Questionnaire 1 - 2 to 5 attributes routed off awareness**

Q1.4a Displayed is a list of statements that may or may not be used to describe each laundry detergent brand. Please indicate the brand or brands that you feel are best described by each of the following statements

Please choose at least two, but not more than five statements per brand.

A laundry detergent brand that…

[PROG NOTE: DISPLAY BRANDS MENTIONED IN Q1c. GRID, BRANDS ON TOP. INCLUDE ‘NONE OF THESE’ IN THE LIST DISPLAYED]
[PROG NOTE: RANDOMIZE ATTRIBUTE LIST. RECORD THE ORDER THE ATTRIBUTES ARE SHOWN TO EACH RESPONDENT]
[PROG NOTE: AT LEAST TWO, BUT NOT MORE THAN FIVE STATEMENTS ARE ALLOWED TO BE CHOSEN]

- Keeps my family looking good
- Helps me feel confident with my appearance
- Is currently a leading brand
- Is a family favourite my mother has always used
- Is a fun brand
- Cares about the environment
- I would be proud to use
- Gives me value for money

Q1.4b Below is a list describing the overall characteristics and features of different laundry detergent brands. Please indicate which brands are best described by each of the following statements.

Please choose at least two, but not more than five statements per brand.

[PROG NOTE: DISPLAY BRANDS MENTIONED IN Q1c. GRID, BRANDS ON TOP. INCLUDE ‘NONE OF THESE’ IN THE LIST DISPLAYED]
[PROG NOTE: RANDOMIZE ATTRIBUTE LIST. RECORD THE ORDER THE ATTRIBUTES ARE SHOWN TO EACH RESPONDENT]
[PROG NOTE: AT LEAST TWO, BUT NOT MORE THAN FIVE STATEMENTS ARE ALLOWED TO BE CHOSEN]

- Provides superior whiteness
- Keeps colours bright
- Leaves clothes smelling fresh
- Is good for a variety of fabrics
- Is an innovative brand
- Is a reliable laundry detergent
- Is a laundry detergent that works well
- Has a wide variety of scents
- A brand that makes my clothes soft
- A brand that gives me more washes
- Has a long lasting scent
A brand that is gentle on my clothes
Is safe for sensitive skin
Is an environmentally friendly brand
Gets clothes really clean
Has a scent I like
A brand that is value for money
My kind of brand

**Questionnaire 2 - All attributes routed off awareness**

Q2.4a Displayed is a list of statements that may or may not be used to describe each laundry detergent brand. Please indicate the brand or brands that you feel are best described by each of the following statements. You may select as many or as few brands as you like for each statement.

A laundry detergent brand that…

[PROG NOTE: DISPLAY BRANDS MENTIONED IN Q1c. GRID, BRANDS ON TOP. INCLUDE ‘NONE OF THESE’ IN THE LIST DISPLAYED]
[PROG NOTE: RANDOMIZE ATTRIBUTE LIST. RECORD THE ORDER THE ATTRIBUTES ARE SHOWN TO EACH RESPONDENT]

**SAME ATTRIBUTES APPEAR**

Q2.4b Below is a list describing the overall characteristics and features of different laundry detergent brands. Please indicate which brands are best described by each of the following statements. You may select as many or as few brands as you like for each statement.

[PROG NOTE: DISPLAY BRANDS MENTIONED IN Q1c. GRID, BRANDS ON TOP. INCLUDE ‘NONE OF THESE’ IN THE LIST DISPLAYED]
[PROG NOTE: RANDOMIZE ATTRIBUTE LIST. RECORD THE ORDER THE ATTRIBUTES ARE SHOWN TO EACH RESPONDENT]

**SAME ATTRIBUTES APPEAR**

**Questionnaire 3 - important attributes routed off awareness**

Q3.4a Displayed is a list of statements that may or may not be used to describe laundry detergent. Please select the statements that you feel are the most important to you when deciding on which laundry detergent to buy.

A laundry detergent that…

**SAME ATTRIBUTES APPEAR**

Q3.4b You have selected a list of statements you feel are most important to you. Now please could you indicate the brand or brands that you feel are best described by each of the statements below. You may select as many or as few brands as you like for each statement.

[PROG NOTE: DISPLAY BRANDS MENTIONED IN Q1c. GRID, BRANDS ON TOP. INCLUDE ‘NONE OF THESE’ IN THE LIST DISPLAYED]
[PROG NOTE: DISPLAY STATEMENTS SELECTED AT Q3.4a]
[PROG NOTE: RANDOMIZE ATTRIBUTE LIST. RECORD THE ORDER THE ATTRIBUTES ARE SHOWN TO EACH RESPONDENT]

Q3.5a Below is a list describing the overall product characteristics and features of laundry detergent brands. When deciding on a laundry detergent brand, which characteristics and features are most important to you?
SAME ATTRIBUTES APPEAR
Q3.5b You have selected a list of statements you feel are most important to you. Now please could you indicate the brand or brands that you feel are described by each of the statements below? You may select as many or as few brands as you like for each statement.

[PROG NOTE: DISPLAY BRANDS MENTIONED IN Q1c. GRID, BRANDS ON TOP. INCLUDE ‘NONE OF THESE’ IN THE LIST DISPLAYED]
[PROG NOTE: DISPLAY STATEMENTS SELECTED AT Q3.5a]
[PROG NOTE: RANDOMIZE ATTRIBUTE LIST. RECORD THE ORDER THE ATTRIBUTES ARE SHOWN TO EACH RESPONDENT]

Questionnaire 4 - indirect importance attributes routed off awareness

Q4.4a Different people look for different things when deciding on which laundry detergent to buy. Thinking about what you look for in a laundry detergent, which of the statements below describe what you could not do without?

A laundry detergent that…

SAME ATTRIBUTES APPEAR
Q4.4b You have selected a list of statements you feel describe what you cannot do without. Please could you now indicate the brand or brands that you feel are best described by each of the statements below. You may select as many or as few brands as you like for each statement.

[PROG NOTE: DISPLAY BRANDS MENTIONED IN Q1c. GRID, BRANDS ON TOP. INCLUDE ‘NONE OF THESE’ IN THE LIST DISPLAYED]
[PROG NOTE: DISPLAY STATEMENTS SELECTED AT Q4.4a]
[PROG NOTE: RANDOMIZE ATTRIBUTE LIST. RECORD THE ORDER THE ATTRIBUTES ARE SHOWN TO EACH RESPONDENT]

Q4.5a Below is a list describing the overall product characteristics and features of laundry detergent brands. Some of these may be more important to you than others. When deciding on a laundry detergent brand, which of the following characteristics and features could you not do without?

Q4.5b You have selected a list of laundry detergent characteristics and features that you feel you cannot do without. Now please could you indicate the brand or brands that you feel are best described by each of the statements below? You may select as many or as few brands as you like for each statement.

[PROG NOTE: DISPLAY BRANDS MENTIONED IN Q1c. GRID, BRANDS ON TOP. INCLUDE ‘NONE OF THESE’ IN THE LIST DISPLAYED]
[PROG NOTE: DISPLAY STATEMENTS SELECTED AT Q4.5a]
[PROG NOTE: RANDOMIZE ATTRIBUTE LIST. RECORD THE ORDER THE ATTRIBUTES ARE SHOWN TO EACH RESPONDENT]