Position bias in best-worst scaling surveys: a case study on trust in institutions

Danny Campbell[†] Seda Erdem

Economics Division, Stirling Management School, University of Stirling, Scotland

2014 Agricultural and Applied Economics Association Annual Meeting
Minneapolis, July 27–29

This paper investigates the effect of physical position on 'best' and 'worst' choices in the best-worst scaling technique. Although the best-worst scaling technique has been used widely in many fields, the phenomenon of consumers' adoption of processing strategies while making choices has been largely overlooked. We examine this issue in the context of consumers' perception of trust in institutions to provide information about a new food technology, namely nanotechnology, and its use in food processing. Our results show that around half of the consumers used position as a schematic cue when making choices. We find the position bias is particularly strong when consumers chose their most trustworthy institution compared to their least trustworthy institution. In light of our findings, we recommend researchers in the field to be aware of the possibility of position bias when designing best-worst scaling surveys. We also encourage researchers who have already collected best-worst data to investigate whether their data shows such heuristics.

Keywords: best-worst scaling ⬩ position bias, consumer trust ⬩ multinomial logit model ⬩ latent class logit model
JEL codes: C25 ⬩ D12 ⬩ Q18

Introduction

A cause for concern in stated choice experiments is that respondents exhibit a decision rule or processing strategy while making choices. A number of these processing strategies, such as attribute non-attendance (Hensher et al., 2005; Campbell et al., 2008; Scarpa et al., 2013), attribute-level non-attendance (Erdem, Campbell and Hole, 2014), elimination- and selection-by aspects (Campbell et al., 2012; Erdem, Campbell and Thompson, 2014; Campbell et al., 2014), and ordering-effect (Day et al., 2012; Carlsson et al., 2012) have been studied in the stated preference literature, mainly in discrete choice experiments. This paper is motivated by the question of whether some of these issues discovered in other stated preference methods are also present in the best-worst scaling (BWS), which is also a pref erence elicitation technique and developed by Finn and Louviere (1992) and colleagues.

Although BWS is not a recent development, it is only recently that we have witnessed its widespread application in a number of disciples, including agriculture (e.g., Lusk and Briggeman, 2009; Erdem et al., 2012), environment (e.g., Scarpa et al., 2011), health (e.g., Louviere and Flynn, 2011), and marketing (e.g., Cohen, 2009). The technique involves respondents choosing two items in a subset of a large list in terms of an underlying scale of importance (e.g., best and worst, or most and least important, items). More about the technique and recent examples can be found in (Flynn et al., 2007; Lusk and Briggeman, 2009; Scarpa et al., 2011; Erdem and Rigby, 2013).

In this paper, we examine whether respondents give different weights to the position of items provided to them in a BWS task. Specifically, we explore the behavioral proposition that respondents used item position as a schematic cue when making choices and the extent to which the probability of an alternative being chosen depends not only on its item, but also on its position in the choice task. To date, this has been an unexplored area. Our paper is motivated by the fact that failing to recognize this phenomena has implications for choice predictions and could have serious repercussions for policy recommendations.

[†]Contact details: Economics Division, Stirling Management School, University of Stirling, Stirling FK9 4LA, Scotland. Email: danny.campbell@stir.ac.uk.

To identify the extent of this issue and to address it, this paper recommends the use of position-specific constants alongside models accommodating a number of latent classes, where the classes are typified by preference heterogeneity and/or position effects.

The empirical case study used in the study focuses on a sample of UK consumers' trust in different sources of information regarding the use of a novel technology, namely nanotechnology, in food production. In particular, the case study investigates consumers' trust in agents and organizations in the UK food industry, and their role in providing balanced and accurate information about nanotechnology and its use in food production and packaging. Giving the contentious history of recent food-related technologies, e.g., genetic modification and irradiation, it is crucial to address whom consumers trust the most and the least regarding said information about emerging food technologies, such as nanotechnology and its implementation. Such information may help explain the public's attitude towards accepting the technology, which may then affect its adoption in the industry. The case study makes an important contribution in this area.

Overall, our findings show that: (1) half of our sample do not consider items listed in different positions in BWS task equally for both 'best' and 'worst' choices; (2) the probability of an institution being chosen depends not only on the institution itself, but also on its position in the BWS choice task; (3) the institution positioned at the top of the choice task stands a significantly higher chance of been identified as being the most trustworthy; and, (4) not accommodating for position bias has implications on choice predictions and the model fit.

The remainder of the paper is structured as follows. In the next section we provide a brief background on position effects. We describe our methodology, as well as the modeling approach for exploring these position effects, or biases, and preference heterogeneity in section three, and outline our empirical case study, BWS data and survey design in section four. Our main results are reported in section five, followed by the final section, which concludes the paper.

Position effects

An extensive literature in consumer research and marketing, and psychology has showed that the manner in which people perceive items, people, or goods often depends on their physical ordering. This includes 'edge avoidance' (Rubinstein et al., 1996), 'centrality preferences' (Shaw et al., 2000), 'middle bias' (Attali and Bar-Hillel, 2003), as well as 'center-stage effect' (Valenzuela and Raghubir, 2009). The situations where these position effects, or biases, have been identified are varied, including: the ordering of response alternatives (Attali and Bar-Hillel, 2003); the allocation of shelf-space in supermarkets (Inman et al., 1990; Wright, 2002; Meier and Robinson, 2004; Valenzuela and Raghubir, 2009); the placement of people (McArthur and Post, 1977; Raghubir and Valenzuela, 2006; Rodway et al., 2013); and, items of choice (Valenzuela and Raghubir, 2009; Guney, 2014). In these situations, the effects are exhibited in both the 'horizontal dimension of space' (see Nisbett and Wilson, 1977; Valenzuela et al., 2013) and 'vertical dimension of space' (see Meier and Robinson, 2004; Koppell and Steen, 2004; Schubert, 2005; Dayan and Bar-Hillel, 2011).

In the horizontal dimension of space, it has been repeatedly shown that the arrangement of products from left to right influences consumers' perception of value, their judgments, and, ultimately, their purchase decisions (e.g., see Raghubir and Valenzuela, 2006; Chandon et al., 2009; Valenzuela and Raghubir, 2009). Specifically, findings in Chandon et al. (2009) and Valenzuela et al. (2013) revealed that consumers perceive products positioned in the center of a shelf more popular, premium or promoted products. Similar effects are observed in other contexts. For example, using six different case studies, Raghubir and Valenzuela (2006) ascertained a strong "center-stage" influence. Their research revealed that people often judged the person in a central position as being more important, a better performer or more likely to be successful. They acknowledged that this heuristic may be due to salience effects (i.e., stimulus that makes it stand apart from other similar stimuli due to either its inherent characteristics), attributional effects (i.e., better performers often chose positions that are more salient and more likely to be evaluated more favorably) and social norms (i.e., more prominent people sit in the middle of the table (McArthur and Post, 1977)).

Connotations and associations of vertical space are in widespread metaphoric use in our daily life. For example, we commonly use phrases such as "on top of things", "high points", "thumbs up/down", "hitting rock bottom", and "climbing the corporate ladder" that all signify a vertical schema of the top being better compared to the bottom in a normative sense. In this case, the vertical dimension of space influences perceptions of value. Not surprisingly, several studies in the fields of marketing and psychology have sought to investigate the issue. The overwhelming evidence from these studies is that items or products located at the top (or higher) are perceived to be 'better' or evaluated more positively than those placed at the bottom (or lower) (Meier and Robinson, 2004; Schubert, 2005; Meier et al., 2007; Valenzuela and Raghubir, 2010; Valenzuela et al., 2013). These are also referred to as primacy and recency order effects. For example, Chandon et al. (2009) found that products on the top and middle shelves gain more attention compared to those on the bottom shelf, and, interestingly, discovered that the effects of vertical position (especially the positive inferences associated with high locations) are stronger than any left *versus* right effect. Again, the influence of vertical positioning goes beyond marketing. For instance, research by Meier and Robinson (2004) has demonstrated that 'positive' words are recognized faster when they were placed at the top of the screen, whereas the recognition of 'negative' words is stronger when they were placed at the bottom of the screen.

Vertical position is also often linked with the notion of power and seniority. The findings in Schubert (2005) revealed that group labels are typically perceived as being more powerful when they were placed at the top of the screen relative to the bottom of the screen. This is also exemplified in corporate organizational charts, where the CEO is located at the top of the chart, followed by directors, managers, and other employees in the hierarchy.

Horizontal and vertical position effects do not necessarily work in isolation. For example, Valenzuela et al. (2013) found that retailers place the premium brand on the top, the cheapest brand on the bottom, the most popular brand in the center, products in promotion at the horizontal extremes (like in Inman et al. (1990)), and store brands next to promoted and popular brands in the center.

With the knowledge that position is a commonly employed heuristic, researchers have looked into whether it can be used for 'nudging' people towards healthier decisions. Dayan and Bar-Hillel (2011) is one of such studies that specifically explored the vertical position effect on food menus. In a lab and real-world studies, they repeatedly showed that food items presented at the top and bottom of the restaurant menu were perceived up to twice as popular as when they were placed in the center of the list. The authors indicated that, given this result, placing the healthy food options at the top or bottom of the lists and less healthy ones in their center should result in healthier food choices. Similarly, van Kleef et al. (2012) examined the effect of manipulating the assortment structure and shelf layout of a display including both healthy and unhealthy snacks near the checkout counter of a canteen. Their participants preferred shelf displays including healthy snack assortment located at top shelves, rather than at the bottom shelves. This, perhaps, contributes to the finding of Rozin et al. (2011) who found that making food slightly more difficult to reach by varying its proximity decreases food intake in obese people.

In this paper, we are only focusing on the vertical position effect in a BWS survey. The BWS data is particularly well suited to exploring this position bias due to the nature of BWS tasks asking respondents to identify their 'best' and 'worst' choices among a subset of a large list of items. In the following section, we describe how we identify and accommodate for this position bias.

Methodology

We start this section by providing a brief description of the BWS technique. We follow this by introducing the necessary notation and a basic model for analyzing the BWS data. Then, we expand on this base model to uncover the role of an item's position on its likelihood of being chosen as best and worst and by making provision for preference heterogeneity.

The best-worst scaling method

While people can usually comfortably rank a small list of items, as the list of items that are to be ranked increases, the ranking task, obviously, becomes more cognitively challenging and, importantly, susceptible to a range of anomalous behaviors. The BWS technique avoids this by breaking tasks into more manageable sizes, thereby reducing—if not eliminating—difficulty in ranking the full list of items in terms of their importance (or preferability). Furthermore, as respondents only choose at the extreme (i.e., best/worst or most/least), the process is considered to be "scale-free" and prevents a scale-use bias (Baumgartner and Steenkamp, 2001). For example, in case of the use of a likert-scale for identifying respondents' level of preferences, there may be situations where respondents may only focus on one part of the scale. Moreover, there may be cases where respondents have difficulty in distinguishing the differences between the levels of the scale. For example, the difference between 'strongly agree' and 'agree' may be difficult to identify. This creates an ambiguity in the interpretation of these scale levels across respondents. In BWS, however, such ambiguity is absent, as only extremes are needed to be identified in a subset of items. There is also evidence that people use better judgment when they only need to identify the extremes, rather than preferences with levels (Louviere, 1993; Marley and Louviere, 2005).

The BWS approach has been used, and shown to be suitable, in a number of research areas to assess people's perception of intangible concepts. For example, Erdem and Rigby (2013) examined the general publics' perception of control and worry over various risks, Erdem et al. (2012) looked at consumers' perception of relative responsibility for ensuring food safety, Louviere and Flynn (2011) examined the public's perception and preferences for healthcare reform in Australia, and Auger et al. (2007) investigated the attitudes of consumers towards social and ethical issues, such as recycling and human rights, across six countries.

Basic model and background notation

BWS is an application of the random utility maximisation theory (Manski, 1977; Thurstone, 1927), whereby respondents evaluate all possible pairs of items within the displayed BWS task and choose the pair that reflects their maximum difference in preference. The number of unique pairs, J, is given by $S(S-1)$, where S represents the number of items in the BWS task. Overall utility, U, associated with respondent n's chosen pair, i, in BWS task t is given by the difference in utility between the best and worst items:

$$U_{nit} = \underbrace{\left(\beta x_{b_{nit}}\right)}_{\text{Best}} - \underbrace{\left(\beta x_{w_{nit}}\right)}_{\text{Worst}} + \varepsilon_{nit}, \tag{1}$$

where β is a vector of estimated parameters (subject to $\sum_{k=1}^{K}\beta_k = 0$) relating to the best and worst items, x (indexed by b and w respectively), and ε is an *iid* type I extreme value (EV1) distributed error term, with constant variance of $\pi^2/6$. Given these assumptions, the probability of the sequence of best-worst choices made by individual n can be represented by the MNL model:

$$\Pr\left(y_n|x_n\right) = \prod_{t=1}^{T_n} \frac{\exp\left(\left(\beta x_{b_{nit}}\right) - \left(\beta x_{w_{nit}}\right)\right)}{\sum_{j=1}^{J}\exp\left(\left(\beta x_{b_{njt}}\right) - \left(\beta x_{w_{njt}}\right)\right)}, \tag{2}$$

where y_n gives the sequence of best-worst choices over the T_n BWS tasks for respondent n, i.e., $y_n = \left[i_{n1}, i_{n2}, \ldots, i_{nT_n}\right]$.[1]

Accounting for position bias

The choice probability retrieved from (2) assumes all respondents consider all offered items and the likelihood of best and worst choices are independent from their position. However, it is important to recognize that the probability of choice may depend not only on utility, but also on an item's location. In particular, in line with evidence found in the papers discussed previously, one could postulate the hypothesis that position acts as a schematic cue that leads to systematic biases in respondent's decisions. For example, when choosing the item that provides them with the greatest utility, respondents may be more inclined to choose among the options located closer to the top (or left) of the choice task. In contrast, the item they indicate as being the worst has a tendency to be located closer to the bottom (or right) of the choice task.

Failing to account for this vertical or horizontal position bias could lead to misguided inferences, as the model does not reflect actual choice behavior. A straightforward approach for addressing this phenomenon is to introduce position-specific constants into the utility function, as follows:

$$\Pr\left(y_n|x_n\right) = \prod_{t=1}^{T_n} \frac{\exp\left(\left(\beta x_{b_{nit}} + \gamma_{b_{nit}}\right) - \left(\beta x_{w_{nit}} + \gamma_{w_{nit}}\right)\right)}{\sum_{j=1}^{J}\exp\left(\left(\beta x_{b_{njt}} + \gamma_{b_{njt}}\right) - \left(\beta x_{w_{njt}} + \gamma_{w_{njt}}\right)\right)}, \tag{3}$$

where the γ terms denote the position-specific constants, which capture the average effect on utility of all factors that are not included in the model.[2] In cases where there are no systematic differences due to item position, we should expect $\gamma = 0$. However, in situations where item position has a bearing on choice

[1] We note that accounting for the panel effect is immaterial in the MNL model due to the independence of choice probabilities. We, nevertheless, present the MNL model in this manner to introduce the necessary terms as early on as possible so that differences in models are clearer as we progress through this section.

[2] Note that these position-specific constants are analogous to the alternative-specific constants that are routinely used in discrete choice modeling. However, in our case, the alternative-specific constants are the effectively the difference between the relevant pair of position-specific constants for the best and worst choices.

outcomes (either negative or positive), we can expect to find $\gamma \neq 0$. Note that the γ terms are indexed by either b or w to distinguish the role of position on best and worst choices, respectively. For identification purposes, the values of γ_b and γ_w are subject to the constraint $\sum_{s=1}^{S} \gamma_{b_s} = 0$ and $\sum_{s=1}^{S} \gamma_{w_s} = 0$ respectively.

The introduction of position-specific constants represents a first step in uncovering the systematic impact of item position on best and worst choices. Nevertheless, a concern remains that the results could be biased by a subset of respondents who entirely overlooked the items, but made their choices purely on the basis of the item's position. In the same vain, there may be another subset of respondents who consistently disregarded the position and made choices that were solely driven by the items themselves. Suggesting the adoption of these different processing strategies is equivalent to identifying three separate classes of choice behavior among respondents:

1. a class in which the choices reflect the preferences of the items in the BWS survey;
2. a class where both preferences and position influenced choice outcomes; and, finally,
3. a class in which the choices are entirely a result of schematic cues based on the item's position within the choice task.

Respectively, the utility functions associated with these three classes can be described by:

$$V_{1_{nit}} = \left(\beta x_{b_{nit}}\right) - \left(\beta x_{w_{nit}}\right), \tag{4a}$$
$$V_{2_{nit}} = \left(\beta x_{b_{nit}} + \gamma_{b_{nit}}\right) - \left(\beta x_{w_{nit}} + \gamma_{w_{nit}}\right), \tag{4b}$$
$$V_{3_{nit}} = \left(\gamma_{b_{nit}}\right) - \left(\gamma_{w_{nit}}\right), \tag{4c}$$

where V_c represents the observable part of utility associated with class c. While deciding the number of processing strategies to accommodate is an empirical consideration, the actual choice process used by respondents remains latent. To get around this, on the basis of observed choice behavior, probabilistic conditions can be imposed on the utility expressions in (4). In doing so, the presence of processing strategies can be established up to a probability, with the full probability per respondent allocated across all C classes. Under this framework the probability of best-worst choice can be represented as follows:

$$\Pr\left(y_n | x_n\right) = \sum_{c=1}^{C} \pi_c \prod_{t=1}^{T_n} \frac{\exp\left(V_{c_{nit}}\right)}{\sum_{j=1}^{J} \exp\left(V_{c_{njt}}\right)}, \tag{5}$$

where π_c denotes the (unconditional) probabilities associated with observing the utility function relating to class c (i.e., the likelihood of competing processing strategies being their actual strategy).

Accounting for heterogeneous preferences and processing strategies

The model described in (5) accommodates respondents with different utility functions and, to avoid confounding between heterogeneity in preferences and processing, equality constraints (see Scarpa et al., 2009) are imposed on β in (4a) and (4b), as well as for γ (4b) and (4c). The model is based on the assumption that all respondents have the same preferences and/or are equally influenced by position. For a variety reasons, most empirical evidence reveals heterogeneity rather than homogeneity across respondents. Accordingly, we treat each of the β and γ parameters as finitely distributed random terms, now denoted with the subscript c (i.e., β_c and γ_c respectively), to represent classes with separate preferences and processing strategies. For example, if we assume two latent segments on the basis of preferences and position effects six utility functions, and hence classes, initially come to mind:

$$V_{1_{nit}} = \left(\beta_1 x_{b_{nit}}\right) - \left(\beta_1 x_{w_{nit}}\right); \tag{6a}$$
$$V_{2_{nit}} = \left(\beta_1 x_{b_{nit}} + \gamma_{1_{b_{nit}}}\right) - \left(\beta_1 x_{w_{nit}} + \gamma_{1_{w_{nit}}}\right); \tag{6b}$$
$$V_{3_{nit}} = \left(\gamma_{1_{b_{nit}}}\right) - \left(\gamma_{1_{w_{nit}}}\right); \tag{6c}$$
$$V_{4_{nit}} = \left(\beta_2 x_{b_{nit}}\right) - \left(\beta_2 x_{w_{nit}}\right); \tag{6d}$$
$$V_{5_{nit}} = \left(\beta_2 x_{b_{nit}} + \gamma_{2_{b_{nit}}}\right) - \left(\beta_2 x_{w_{nit}} + \gamma_{2_{w_{nit}}}\right); \tag{6e}$$
$$V_{6_{nit}} = \left(\gamma_{2_{b_{nit}}}\right) - \left(\gamma_{2_{w_{nit}}}\right). \tag{6f}$$

In addition to the above, we should recognize that for some respondents estimated as having β_1, their position effects may be best described used γ_2 and, similarly, the possibility that the position effects characterized by γ_1 may be have been exhibited by respondents with the item coefficients β_2. Therefore, it is important to recognize two further classes:

$$V_{7_{nit}} = \left(\beta_1 x_{b_{nit}} + \gamma_{2_{b_{nit}}}\right) - \left(\beta_1 x_{w_{nit}} + \gamma_{2_{w_{nit}}}\right); \tag{6g}$$

$$V_{8_{nit}} = \left(\beta_2 x_{b_{nit}} + \gamma_{1_{b_{nit}}}\right) - \left(\beta_2 x_{w_{nit}} + \gamma_{1_{w_{nit}}}\right). \tag{6h}$$

Including (6g) and (6h) means that we are in a better position to jointly identify the marginal utilities and position influences. This is important since it goes some way to alleviate the risk of confounding (see Campbell et al., 2012; Hensher et al., 2012; Hess et al., 2012, for a discussion on this issue). Once again, the probabilities associated with the above eight representative utility functions, as well as the segment-specific vector of βs and γs, can be derived using (5), where $C = 8$.

Data

The BWS data is obtained from an empirical case study that investigates consumers' perception of trust in institutions about providing accurate and balance information regarding the use of a novel technology, namely nanotechnology, and its use in food packaging and production. Overall, we used 16 institutions, ranging from government institutions to the media, friends and family. Table 1 shows the institutions included in the BWS survey.

Survey design plays an important role in obtaining reliable responses. In our survey, each respondent was presented with five institutions at each of eight BWS choice tasks. For each choice task, they were asked to indicate the 'most' and 'least' trustworthy institutions among presented subset of institutions. Figure 1 illustrates a typical BWS task presented to the respondents. Given that the total number of items

Table 1: Institutions and agents included in the best-worst scaling study

Item	Institution/agent	Coding
Government institutions		
1	Department for Environment, Food and Rural Affairs	DEFRA
2	Food Standards Agency	FSA
3	Department of Health	DH
Scientists		
4	Food industry scientists	FoodIndSci
5	University scientists	UniSci
Non-government organizations		
6	Consumer organizations (e.g., Which?, National Consumer Federation etc.)	ConsumOrg
7	Environmental groups (e.g., Greenpeace, Friends of the Earth etc.)	EnvGrps
Food handlers		
8	Food manufacturers/processors	Manufact
9	Farmers/growers	Farmers
10	Supermarkets	Supermkt
11	High street butchers	Butchers
Friends and family		
12	Friends and family	Friends
Media		
13	TV/radio: news programmes	News
14	TV/radio: food and cooking programmes	FoodProg
15	Newspapers	NewsPaps
16	Food magazines (e.g., Good Food magazine, Sainsbury's and Tesco's magazines etc.)	Magazines

Who do you trust to tell you about nanotechnology?

Consider the five organisations/people shown below. Please indicate which of the five you:

- Trust **MOST** to provide accurate and balanced information about nanotechnology and its use in food production
- Trust **LEAST** to provide accurate and balanced information about nanotechnology and its use in food production

Trust **most** on nanotechnology		Trust **least** on nanotechnology
○	Newspapers	○
○	Environmental groups	○
○	Farmers/growers	○
○	Food industry scientists	○
○	Food Standards Agency	○

Figure 1: Typical best-best scaling task

used in the survey is quite large[3], we felt that it is plausible to use five items in each choice task. This decision was also driven by feedback from the pilot study and evidence that showing more than five items to respondents may result in confusion and fatigue (e.g., see Cohen and Orme, 2004), which may, in turn, result in unreliable responses.

The experimental design comprised of 300 versions (i.e., blocks) to avoid any context and ordering based biases. In each BWS choice task, different combinations of five institutions were shown to respondents. The combinations of five institutions in these choice tasks satisfy the optimal design characteristics: frequency balance; orthogonality; positional balance; and, connectivity among tasks. That is, the one-way frequencies reveal that the survey design was perfectly balanced as each item in the survey was displayed 750 times[4] across all versions of the surveys. The two-way frequencies show that the survey had a nearly orthogonal main-effects design, in which each item appeared 200 times on average with every other item, with a standard deviation of 0.51. The positional frequencies show that each item, on average, appeared 150 times at each position (i.e., first, second, third, fourth, and fifth) with a standard deviation of 0.45. After ensuring a balanced and nearly orthogonal survey design, tasks were randomized, and a participant was randomly assigned to a version.

The web-based surveys were conducted with a sample of 616 consumers in the UK in 2010. With each respondent answering 8 BWS tasks, we obtained a total of 4,298 observations for model estimation. Just over half of the respondents were female (51%), approximately one quarter were in full-time employment, had education until at least 18 years-old, and fell in the 30–45 age group (35%). The average annual household income was about £25K–£30K. A comparison with the 2011 UK census data shows that the respondents in our study were similar to the general UK population with respect to age, gender, and employment status.

Results

In an attempt to tease out the impact of item position on the BWS choices, we begin this section with a rudimentary examination of the choices made by individuals. Following this, we report results from our econometric models and post-estimation analysis.

Examination of choices

As a first step in assessing the role that item position had on choices in the BWS exercise, we test the H_0 that, other things being held constant, there is no association between the set of observed counts of best and worst choices in each position and their expected counts. If there is no ordering effect, each position should be chosen an equal number of times as the best and worst option. Given that there were five items per choice task (i.e., $S = 5$) and our dataset consists of 4,928 choice observations this would equate to an expected breakdown of 985.6 (i.e., 20 percent) best and worst choices per position and 246.4 (i.e., 5 percent)

[3] Other trust studies typically included less than ten institutions.

[4] As five items are presented in each set, there were overall 40 items shown in every version (i.e., 5 items times 8 tasks). As there are 16 institutions in total, each institution appeared approximately 2.5 times in each version. Across all 300 versions, each institution appears 750 times.

for each combination of best-worst choices. Table 2 compares the actual and expected distribution of choices.

Looking firstly at the positional spread of best choices, there appears to be some deviations between the observed and expected distributions. Notably, all else being equal, there is seemingly clear evidence that institutions located closer to the top of the BWS choice task have a higher likelihood of being chosen as best (in this case, most trustworthy). Indeed, further inspection of the best choices reveals that, for the most part, the observed choices for each position follow a monotonically decreasing pattern as one moves from the top of the choice task to the bottom, which is in accordance with findings in Valenzuela et al. (2013). Moreover, with a χ^2 test statistic of 52.006, against the critical value of 9.488 ($\chi^2_{0.05,4}$), we can reject the H_0 that, *ceteris paribus*, the institutions identified as being most trustworthy were not subject to a position bias.

It is interesting to note that as we move our attention to the distribution of worst choices(in this case, least trustworthy) we find the opposite finding—other things being constant, institutions located closer to the bottom of the choice task were more likely to be identified as being least trustworthy. While the pattern is not as striking, the χ^2 test statistic of 15.458 (against the same critical value of 9.488 ($\chi^2_{0.05,4}$)), does, nevertheless, point towards a significant ordering effect.

Taking this analysis further, we also compare the 20 best-worst choice combinations. In this case, we, again, find that the χ^2 test statistic of 100.239 exceeds the critical value of 30.144 ($\chi^2_{0.05,19}$). This, therefore, provides further compelling evidence to support the rejection of the H_0 in favor of the H_1, where position plays an influential role on the choices made by respondents.

Table 2: Observed choices versus expected choices

			Observed		Expected	
Best (s)	Worst (s)	Pair (j)	Count	Percent	Count	Percent
1			1136	23.05	985.60	20.00
2			1078	21.88	985.60	20.00
3			909	18.45	985.60	20.00
4			917	18.61	985.60	20.00
5			888	18.02	985.60	20.00
	1		1008	20.45	985.60	20.00
	2		930	18.87	985.60	20.00
	3		910	18.47	985.60	20.00
	4		1038	21.06	985.60	20.00
	5		1042	21.14	985.60	20.00
1	2	1	252	5.11	246.40	5.00
1	3	2	297	6.03	246.40	5.00
1	4	3	304	6.17	246.40	5.00
1	5	4	283	5.74	246.40	5.00
2	1	5	260	5.28	246.40	5.00
2	3	6	219	4.44	246.40	5.00
2	4	7	314	6.37	246.40	5.00
2	5	8	285	5.78	246.40	5.00
3	1	9	261	5.30	246.40	5.00
3	2	10	210	4.26	246.40	5.00
3	4	11	215	4.36	246.40	5.00
3	5	12	223	4.53	246.40	5.00
4	1	13	244	4.95	246.40	5.00
4	2	14	239	4.85	246.40	5.00
4	3	15	183	3.71	246.40	5.00
4	5	16	251	5.09	246.40	5.00
5	1	17	243	4.93	246.40	5.00
5	2	18	229	4.65	246.40	5.00
5	3	19	211	4.28	246.40	5.00
5	4	20	205	4.16	246.40	5.00

Estimation results

While 'eyeballing' the breakdown of choices and the non-parametric test statistics might provide the first clue of position bias, they cannot rule out experimental design artifacts relating to the location of the trust items. For this reason, we turn to the results of the models described in the methodology section. We present these results in table 3. All models are estimated using maximum likelihood estimation.[5] For each model, we separately report the estimated trust coefficients for the institutions used in the BWS survey and position-specific constants—where, for normalization, the final item (food magazines) and position $s = 5$ are arbitrarily set to the base level (i.e., the negative sum of their respective coefficients) and, therefore, omitted from the table. Class membership probabilities, where applicable, along with model fit and diagnostic statistics are also provided in the table.

As a point of reference our analysis starts with the MNL model with position-specific constants (labeled Model 1), as specified in (3).[6] Looking firstly at the results of this model, we observe that, on average, respondents are more likely to trust communication on nanotechnology and its use in food production from government institutions and scientists compared to non-government organizations, food handlers, friends and family and the media. To ease the interpretation, we provide the ratio-scaled probabilities, $\overset{*}{\Pr}(x)$, in table 4.[7] These scores provide a more intuitive interpretation. As they show, for instance that, under Model 1 the information coming from the Food Standard Agency is, on average, considered to be more than seven times more trustworthy compared to information provided in newspapers (i.e., 13.95/1.97).

The position-specific constants retrieved under our first model give an important insight into position effects. Firstly, we draw attention to the fact that they are non-zero and, importantly, in most instances the deviations from zero are statistically significant—meaning that we cannot accept the H_0 that there are no systematic differences due to item position. Moreover, the position-specific constants differ between the best and the worst choices also signifies that the schematic cues stemming from an item's position are not the same for best and worst choices. Notably, the values of the position-specific constants for the best choices (γ_b) are of a higher magnitude compared to those obtained for the worst choices (γ_w)—implying that, other things remaining constant, the position bias is stronger for the best choices compared to the worst choices. Interestingly, in accordance with the pattern of observed best choices in table 2, there is a reduction in the position-specific constants as one moves from the top to the bottom item position. The position-specific constants for the worst choices indicate a somewhat different pattern. While position bias does not seem to have played as strong a role, the estimates do, nonetheless, imply that, compared to the uppermost and lowermost items, respondents were slightly less inclined to choose items located in the center when making these choices, a similar finding was shown in Dayan and Bar-Hillel (2011). To facilitate the interpretation, we calculate the probability of each combination of pairs being chosen using only the position-specific constants. The retrieved, *ceteris paribus*, best and worst position probabilities are reported in table 5. From these calculations, the position effects predicted under Model 1 are more clear to see.

Model 2 is a latent class logit model. However, in this case, each latent class is described by the set of specific heuristics described in (4), rather than a set of marginal utilities that is more common in latent class models. Firstly, we remark the large increase in the model fit. We do acknowledge that this improvement is

[5] In the case of the models that retrieve class probabilities, we are mindful of their vulnerability to local maxima of the sample-likelihood function. Thus, in an attempt to reduce the possibility of reaching a local, rather than a global, maximum, we started the estimation iterations from a variety of random starting points. Specifically, we do this by estimating these models many times, but each time using a different vector of starting values, which are chosen randomly. We also note that all models were coded and estimated in Ox version 6.2 (see Doornik, 2009, for further details).

[6] For the sake of brevity, we do not report the MNL model without position-specific constants nor the MNL with only position-specific constants. With log-likelihood values of -12,495.56 and -14,728.66 respectively (versus a null log-likelihood of -14,762.97), these were both found to be inferior to our reference model in table 3, which is associated with -12,455.82 log-likelihood units.

[7] We recognize that the trust coefficients in table 3, which are on an interval scale and consist of both negative and positive values, make interpretation difficult. For this reason, similar to Erdem and Rigby (2013), we convert the raw trust coefficients, which are zero-centered, to ratio-scaled probabilities, which we denote using $\overset{*}{\Pr}(x)$. For item k, the conversion to a 0–100 point ratio scale is achieved as follows:

$$\overset{*}{\Pr}(x_k) = \left(\frac{\exp(\beta_k)}{\exp(\beta_k) + S - 1} \Big/ \sum_{k=1}^{K} \frac{\exp(\beta_k)}{\exp(\beta_k) + S - 1} \right) \times 100,$$

where S, as previously defined, is the number of items shown per choice task (in our case $S = 5$). We acknowledge that the conversion to ratio-scaled probabilities does not factor out the scaling of the parameter estimates that is related to the scale factor of the unobserved Gumbel error component. In each of our models (and latent classes) these scale parameters are normalized in estimation (essentially to 1.0), which we admit thereby prevents any meaningful comparison of parameter estimates between models (and classes). Notwithstanding this limitation, we feel that the ratio-scaled probabilities do, nevertheless, provide a valuable insight into how position bias and the manner in which it is addressed has an impact on the model outputs.

Table 3: Estimation results

	Model 1		Model 2		Model 3				Model 4			
LL	-12,455.82		-11,818.46		-11,687.46				-11,478.93			
K	23		25		47				53			
$\bar{\rho}^2$	0.155		0.198		0.205				0.219			
AIC	24,957.64		23,686.91		23,468.92				23,063.87			
BIC	25,107.20		23,849.48		23,774.55				23,408.51			
Trust coefficients												
	$\hat{\beta}_1$	$\|t\text{-rat.}\|$	$\hat{\beta}_1$	$\|t\text{-rat.}\|$	$\hat{\beta}_1$	$\|t\text{-rat.}\|$	$\hat{\beta}_2$	$\|t\text{-rat.}\|$	$\hat{\beta}_1$	$\|t\text{-rat.}\|$	$\hat{\beta}_2$	$\|t\text{-rat.}\|$
Government institutions												
DEFRA	0.89	20.71	1.68	25.34	1.87	24.53	-0.05	0.61	1.95	23.58	0.95	6.91
FSA	1.41	31.52	2.41	34.12	2.64	31.63	0.25	3.17	2.78	32.59	1.33	9.08
DH	0.98	22.53	1.79	26.55	1.96	25.45	0.05	0.68	2.07	25.07	0.85	5.99
Scientists												
FoodIndSci	0.68	15.92	1.29	19.65	1.53	19.48	-0.13	1.81	1.82	22.68	-0.47	2.91
UniSci	0.85	19.84	1.42	21.63	1.51	21.09	0.35	4.80	1.52	18.68	1.33	9.77
Non-government organizations												
ConsumOrg	0.77	17.93	1.09	17.50	1.01	14.68	0.76	9.90	0.89	10.75	1.97	14.19
EnvGrps	-0.11	2.70	-0.38	6.12	-0.63	8.53	0.44	5.89	-0.89	11.65	1.40	9.60
Food handlers												
Manufact	-0.76	17.90	-1.15	19.12	-1.09	16.36	-0.61	8.06	-0.87	10.50	-2.51	16.04
Farmers	-0.14	3.40	-0.39	6.44	-0.46	7.07	0.15	2.01	-0.42	5.43	-0.18	1.12
Supermkt	-0.88	20.55	-1.37	23.39	-1.39	21.46	-0.46	6.26	-1.31	18.15	-1.89	13.86
Butchers	-0.35	8.33	-0.78	12.79	-0.93	13.98	0.21	2.83	-0.95	12.70	-0.24	1.38
Friends and family												
Friends	-0.73	17.17	-1.39	22.83	-1.56	23.53	0.16	2.15	-1.66	23.37	-0.53	3.29
Media												
News	-0.51	12.09	-0.88	14.96	-0.93	14.21	-0.19	2.58	-1.07	15.16	-0.18	1.50
FoodProg	-0.52	12.27	-0.84	14.42	-0.91	14.40	-0.25	3.58	-1.02	14.76	-0.32	2.72
NewsPaps	-1.18	27.27	-1.79	29.60	-1.87	28.17	-0.56	7.66	-2.02	26.46	-1.13	9.24
Position-specific constants												
	$\hat{\gamma}_1$	$\|t\text{-rat.}\|$	$\hat{\gamma}_1$	$\|t\text{-rat.}\|$	$\hat{\gamma}_1$	$\|t\text{-rat.}\|$	$\hat{\gamma}_2$	$\|t\text{-rat.}\|$	$\hat{\gamma}_1$	$\|t\text{-rat.}\|$	$\hat{\gamma}_2$	$\|t\text{-rat.}\|$
Best												
γ_{b_1}	0.19	6.23	0.36	7.01	0.14	3.02	0.31	5.98	1.14	5.51	0.25	3.12
γ_{b_2}	0.10	3.22	0.21	4.13	0.04	0.87	0.19	3.82	1.07	5.38	0.10	1.28
γ_{b_3}	-0.10	3.02	-0.14	2.56	-0.08	1.81	-0.11	1.93	0.22	1.03	-0.13	1.91
γ_{b_4}	-0.07	2.32	-0.13	2.21	-0.05	1.00	-0.10	1.84	-0.92	2.22	-0.03	0.48
Worst												
γ_{w_1}	-0.07	2.12	-0.16	2.95	-0.06	1.55	-0.12	2.13	0.16	0.75	-0.23	3.12
γ_{w_2}	0.05	1.68	0.17	2.93	-0.01	0.16	0.18	3.14	0.16	0.69	0.14	1.84
γ_{w_3}	0.11	3.38	0.16	3.02	0.12	2.89	0.08	1.47	-0.31	1.92	0.29	3.80
γ_{w_4}	-0.05	1.54	-0.15	2.91	0.04	1.06	-0.17	3.42	-0.48	4.11	-0.05	0.78
Unconditional class membership probabilities												
	$\hat{\pi}$	$\|t\text{-rat.}\|$	$\hat{\pi}$	$\|t\text{-rat.}\|$	$\hat{\pi}$	$\|t\text{-rat.}\|$			$\hat{\pi}$	$\|t\text{-rat.}\|$		
π_{β_1}			0.53	7.58					0.36	4.34		
$\pi_{\beta_1\gamma_1}$	1.00	fixed	0.19	2.98	0.66	18.97			0.01	1.02		
π_{γ_1}			0.27	12.09					0.06	3.76		
π_{β_2}									0.14	3.14		
$\pi_{\beta_2\gamma_2}$					0.34	12.84			0.06	1.48		
π_{γ_2}									0.16	7.74		
$\pi_{\beta_1\gamma_2}$									0.21	2.62		
$\pi_{\beta_2\gamma_1}$									0.00	0.00		

in part due to the fact that the panel nature of the data is being accounted for. However, the magnitude of this increase does, nevertheless, provide clear evidence in favor of simultaneously accounting for the three information processing strategies over the assumption of processing homogeneity. Indeed, at the expense of just two additional parameters, we witness an improvement of over 600 log-likelihood units. Importantly, with an unconditional class membership probability of $\pi_{\beta_1} = 0.53$, choices made by the majority of respondent were not sensitive to any such position bias. While this is a somewhat reassuring finding, it also draws attention to the alarming fact that for almost half of the respondents the item's position influenced its likelihood of being chosen. In addition, upon further inspection, we observe that

Table 4: Ratio-scaled probabilities

	Model 1	Model 2		Model 3			Model 4		
x	$\overset{*}{\Pr}(x)$	$\overset{*}{\Pr}(x\|\beta_1)$	$\mathbb{E}\left(\overset{*}{\Pr}(x)\right)$	$\overset{*}{\Pr}(x\|\beta_1)$	$\overset{*}{\Pr}(x\|\beta_2)$	$\mathbb{E}\left(\overset{*}{\Pr}(x)\right)$	$\overset{*}{\Pr}(x\|\beta_1)$	$\overset{*}{\Pr}(x\|\beta_2)$	$\mathbb{E}\left(\overset{*}{\Pr}(x)\right)$
Government institutions									
DEFRA	10.46	13.51	7.23	14.19	5.84	11.37	14.33	9.63	10.19
FSA	13.95	17.32	9.26	17.83	7.36	14.29	18.04	11.91	12.78
DH	11.00	14.10	7.54	14.67	6.31	11.85	14.99	9.06	10.45
Scientists									
FoodIndSci	9.12	11.22	6.01	12.25	5.45	9.96	13.67	3.30	8.54
UniSci	10.20	11.98	6.41	12.18	7.96	10.75	12.02	11.92	9.31
Non-government organizations									
ConsumOrg	9.66	10.05	5.38	9.35	10.54	9.75	8.52	15.72	8.05
EnvGrps	5.03	3.43	1.86	2.70	8.48	4.65	2.10	12.35	3.67
Food handlers									
Manufact	2.89	1.72	0.95	1.78	3.61	2.40	2.13	0.49	1.33
Farmers	4.90	3.41	1.85	3.11	6.81	4.36	3.18	4.22	2.68
Supermkt	2.60	1.41	0.78	1.34	4.12	2.28	1.43	0.90	1.00
Butchers	4.12	2.42	1.32	2.06	7.15	3.78	1.98	4.02	1.95
Friends and family									
Friends	2.97	1.38	0.76	1.14	6.90	3.09	1.02	3.15	1.22
Media									
News	3.59	2.20	1.20	2.05	5.22	3.12	1.77	4.23	1.87
FoodProg	3.57	2.30	1.25	2.09	4.95	3.05	1.86	3.76	1.83
NewsPaps	1.97	0.94	0.53	0.85	3.80	1.85	0.72	1.84	0.78
Magazines	3.96	2.62	1.43	2.41	5.49	3.45	2.24	3.50	1.99

over one-quarter (i.e., $\pi_{\gamma_1} = 0.27$) of respondents are predicted as having made their choices solely on the basis of item position. Scrutinizing the position-specific constants (along with the derived probabilities in table 5) reveals a similar pattern of position bias to that which emerged from Model 1. We note, however, that in the case of Model 2, the estimated position-specific constants only relate to the subset associated with the utility functions given by (4a) and (4c). Within these two classes, as presented in table 5, the top two positions alone account for over half of the best choices (0.27 for the first, 0.25 for the second position). In contrast, the respective figure for the bottom two positions is approximately 30 percent. Although, again, position is found to be less influential in the worst choices, we find further supporting evidence of a inclination towards the top and bottom positions.

Model 3 assumes two latent classes, which differ in terms of the trust coefficients and position-specific constants (i.e., the utility functions represented by (6b) and (6e)). This model attains a superior fit as compared to Models 1 and 2. Although estimating separate trust coefficients and position-specific constants for each class comes at a very high parametric cost, the $\bar{\rho}^2$, as well as both information criteria, confirm

Table 5: Position probabilities

	Model 1	Model 2		Model 3			Model 4		
s	$\Pr(s)$	$\Pr(s\|\gamma_1)$	$\mathbb{E}(\Pr(s))$	$\Pr(s\|\gamma_1)$	$\Pr(s\|\gamma_2)$	$\mathbb{E}(\Pr(s))$	$\Pr(s\|\gamma_1)$	$\Pr(s\|\gamma_2)$	$\mathbb{E}(\Pr(s))$
Best									
1	0.24	0.27	0.23	0.23	0.26	0.24	0.41	0.24	0.23
2	0.22	0.25	0.22	0.21	0.25	0.22	0.38	0.23	0.22
3	0.18	0.18	0.19	0.19	0.18	0.19	0.14	0.19	0.19
4	0.18	0.16	0.18	0.19	0.17	0.18	0.04	0.19	0.19
5	0.17	0.14	0.17	0.19	0.15	0.17	0.03	0.16	0.17
Worst									
1	0.20	0.21	0.20	0.20	0.20	0.20	0.12	0.23	0.21
2	0.18	0.16	0.18	0.20	0.16	0.19	0.12	0.17	0.18
3	0.18	0.18	0.19	0.18	0.19	0.18	0.26	0.15	0.18
4	0.21	0.24	0.22	0.19	0.24	0.21	0.35	0.21	0.21
5	0.22	0.22	0.21	0.22	0.21	0.22	0.14	0.24	0.21

this finding even after accounting for the loss of parsimony. By jointly inspecting the trust coefficients and ratio-scaled probabilities, we find that the main differences between the first and second latent classes (which are associated with unconditional class membership probabilities of $\pi_{\beta_1\gamma_1} = 0.66$ and $\pi_{\beta_2\gamma_2} = 0.34$ respectively) are the level of trust placed on communication from government institutions and scientists. Other things being equal, whereas the first class considers information originating from these institutions to be highly trustworthy (ratio-scaled probabilities in the range 12–18), the second class considers the information relatively less reliable (ratio-scaled probabilities in the range 5–8). The second class appear to deem communications on nanotechnology relatively more trustworthy when it originates from friends and family as well as media sources.

Focusing on the position-specific constants for the best choices, we, again, discover that the average effect on utility reduces as we move from the top item position to the bottom item position. Interestingly, this same position bias is manifested in both latent classes, albeit it appears to be more perceptible in the second class. We also find the position effect in worst choices. We remark that the effect is quite similar in both classes and broadly consistent with that uncovered in the previous models.

Our final model is a further latent class logit model, combining features of Models 2 and 3. It simultaneously accounts for position effect and trust heterogeneity by allowing for all eight utility expressions in (6). As expected, Model 4 is associated with the best model fit and, importantly, this is corroborated by all of the diagnostic statistics which account for the increase in estimated parameters. Looking firstly at the trust coefficients and ratio-scaled probabilities, we find that they correspond reasonably well to those retrieved under Model 3. Relatively speaking, classes associated with $\hat{\beta}_1$ (with an aggregate unconditional class membership probability of $\pi_{\beta_1} + \pi_{\beta_1\gamma_1} + \pi_{\beta_1\gamma_2} = 0.58$), once more, perceive government institutions and scientists more trustworthy than other institutions, as compared those estimated with $\hat{\beta}_2$ (with an aggregate unconditional class membership probability of $\pi_{\beta_2} + \pi_{\beta_2\gamma_2} + \pi_{\beta_2\gamma_1} = 0.20$), who, again, favor information from friends and family, non-government organizations and media sources. Indeed, comparing these groups of consumers, the first group considers information from food industry scientists to be, on average, 19 times more trustworthy compared information provided in newspapers (i.e., 18.04/0.72), whereas the second groups deem the information to be less than two times as trustworthy (i.e., 3.30/1.84).

Of central interest in this paper is the effect of position on consumers' choices in the BWS survey. From the results of our best fitting model, we find that approximately half (i.e., $\pi_{\beta_1\gamma_1} + \pi_{\gamma_1} + \pi_{\beta_2\gamma_2} + \pi_{\gamma_2} + \pi_{\beta_1\gamma_2} + \pi_{\beta_2\gamma_1} = 0.50$) of the respondents used position, to some extent, as a schematic cue. Inspecting position probabilities obtained from the position-specific constants for the best choices, we see that the same pattern is emerging. Irrespective of the item itself, respondents are systematically more inclined to choose it if it is located at the top of the BWS task and this tendency reduces as the item approaches the bottom position. Startlingly, over 40 percent of the respondents associated with the first set of position-specific constants are predicted to chose the top item, no matter what it is. Moreover, this proportion drops to almost zero for the bottom position. While we add a cautionary note that this behavior only applies to a subset of approximately 7 percent of respondents, it is clearly non-trivial. Interestingly, the two sets of probabilities established for the worst positions show contrasting patterns. Respondents estimated as having $\hat{\gamma}_1$ appear to be subject to a strong centrality bias (Shaw et al., 2000; Attali and Bar-Hillel, 2003), or center stage effect (Valenzuela and Raghubir, 2009), whereas a top-bottom effect (Meier and Robinson, 2004; Dayan and Bar-Hillel, 2011) is found for those with $\hat{\gamma}_2$. Taking the effects of item position on best and worst choices together, the results stemming from Model 4 provide compelling evidence of the extent to which an item's position influences its likelihood of being identified as the most and least trustworthy. This is an important finding and gives an important insight into the decision making heuristics adopted in BWS.

Scenario analysis

To further tease out the effects of position bias, we explore choice probabilities for a specific choice task. This analysis uses the estimates reported in table 3 to assess choice predictions under each model (and latent class) specification discussed earlier. For this analysis, in order to clearly demonstrate the repercussions of the position bias, we deliberately place the items that were consistently found to be the least and most trustworthy, namely newspapers and the Food Standards Agency, at the top and bottom positions respectively. For the intermediate positions, we place environmental groups, farmers/growers and food industry scientists, sequentially in positions 2–4 (as portrayed in figure 1). Results from this post-estimation analysis are given in table 6. For ease of comparison, we also report the expected values, which accounts for the unconditional class membership probabilities.

As expected, under Model 1, we observe the largest prediction for best choice to be the Food Standards Agency (c. 51 percent), and newspapers having the smallest probability of choice (c. 2 percent). The

Table 6: Scenario probabilities

	Best					Worst				
	NewsPaps	EnvGrps	Farmers	FoodIndSci	FSA	NewsPaps	EnvGrps	Farmers	FoodIndSci	FSA
Model 1										
$\Pr(s)$	0.02	0.12	0.10	0.25	0.51	0.58	0.16	0.16	0.07	0.02
Model 2										
$\Pr(s\|\beta_1)$	0.00	0.04	0.04	0.22	0.70	0.66	0.16	0.16	0.02	0.00
$\Pr(s\|\beta_1,\gamma_1)$	0.01	0.06	0.04	0.25	0.65	0.72	0.12	0.13	0.03	0.00
$\Pr(s\|\gamma_1)$	0.27	0.25	0.18	0.16	0.14	0.21	0.16	0.18	0.24	0.22
$\mathbb{E}(\Pr(s))$	0.08	0.10	0.08	0.21	0.54	0.55	0.15	0.16	0.08	0.06
Model 3										
$\Pr(s\|\beta_1,\gamma_1)$	0.00	0.02	0.03	0.23	0.71	0.67	0.18	0.13	0.02	0.00
$\Pr(s\|\beta_2,\gamma_2)$	0.11	0.38	0.20	0.13	0.18	0.38	0.08	0.14	0.26	0.14
$\mathbb{E}(\Pr(s))$	0.04	0.14	0.08	0.20	0.54	0.57	0.15	0.14	0.10	0.05
Model 4										
$\Pr(s\|\beta_1)$	0.00	0.01	0.02	0.26	0.70	0.65	0.21	0.13	0.01	0.00
$\Pr(s\|\beta_1,\gamma_1)$	0.02	0.12	0.09	0.31	0.46	0.61	0.18	0.19	0.02	0.00
$\Pr(s\|\gamma_1)$	0.41	0.38	0.14	0.04	0.03	0.12	0.12	0.26	0.35	0.14
$\Pr(s\|\beta_2)$	0.02	0.44	0.08	0.05	0.41	0.51	0.02	0.19	0.26	0.03
$\Pr(s\|\beta_2,\gamma_2)$	0.02	0.50	0.07	0.05	0.35	0.58	0.02	0.13	0.25	0.03
$\Pr(s\|\gamma_2)$	0.24	0.23	0.19	0.19	0.16	0.23	0.17	0.15	0.21	0.24
$\Pr(s\|\beta_1,\gamma_2)$	0.00	0.02	0.03	0.29	0.66	0.74	0.16	0.09	0.01	0.00
$\Pr(s\|\beta_2,\gamma_1)$	0.05	0.83	0.06	0.01	0.06	0.36	0.01	0.23	0.38	0.02
$\mathbb{E}(\Pr(s))$	0.07	0.16	0.07	0.20	0.50	0.54	0.15	0.14	0.11	0.05

predictions for the worst choice is essentially the mirror image of the best choice (c. 2 percent for FSA and c. 58 percent for newspapers).

Results arising from Model 2 clearly shows how the predictions differ depending on the processing strategy adopted by respondents. For respondents whose choice were in no way influenced by position, approximately 70 percent are predicted as identifying the Food Standards Agency as being the most trustworthy to provide accurate information on nanotechnology. This is in contrast to the prediction of only 14 percent who made their choices exclusively on the basis of item position. Relatedly, whereas the respective prediction for newspapers is effectively zero in the first two latent classes, it jumps to almost 30 percent in the case of the third latent class, which is comprised of those who made choices based on item position only. Although there is, again, a reversal of the predictions as we move to the worst choice, position bias plays a somewhat lesser role.

The separate predictions based on Model 3 are a consequence of differing levels of trust placed on the items and position-specific constants. For this model, we draw particular attention to the marked difference between predictions for the best and worst choices across the two classes.

Interpreting the predictions attained in Models 2 and 3 demonstrates the difficultly in deciphering whether these differences are an artifact of heterogeneous levels of trust or position bias. To some extent, Model 4 overcomes this issue of confounding, since both of these influences are isolated. For instance, comparing the two groups of respondents who made choices independently of position, we see that for one group the Food Standards Agency is deemed most trustworthy (i.e., $\Pr(s|\beta_1) = 0.70$), while the other group it is environmental groups (i.e., $\Pr(s|\beta_2) = 0.44$). For both these groups the probability of choosing newspapers as being most trustworthy is effectively zero. However, as already established, for the respondents who completely disregarded the items and choose purely on the basis of position the respective prediction is either approximately 41 or 24 percent, depending on which position-specific constants they are connected with (i.e., $\Pr(s|\gamma_1)$ or $\Pr(s|\gamma_2)$ respectively). Similarly, the item predicted as least trustworthy differs across the eight latent classes. Classes which retrieve position-specific constants and, thus, accommodate position bias predict a substantially larger share of respondents who select the Food Standards Agency as providing the least accurate information. Related to this, in these classes the respective predictions for newspapers are much reduced. We note here that the centrality bias identified in classes 2, 3 and 8 has also led to relatively higher predictions of worst choices for farmers and food industry scientists.

Discussion and conclusion

In this paper, we present results from a best-worst scaling (BWS) study investigating consumers' perceived level of trust in different sources of information regarding the use of a new technology, namely nanotechnology, in food production. As part of the analysis, we explore the behavioral proposition that respondents used position as a schematic cue when making choices. To empirically explore this issue, we use position-specific constants and a series of latent class logit models, where the classes differ according to: (1) the extent to which location confers a systematic advantage or disadvantage of being chosen; and/or (2) perceptions of trust.

Hitherto, position effects have been overlooked in the analysis of BWS data. From this study, we report several important methodological insights. Firstly, a simple 'eyeballing' of observed choices and the use of a straightforward non-parametric test can help signal the extent of position effects in a given BWS dataset. However, the use of latent class logit models can further shed light into the issue. In our case study, we find that the choices made by around half of our sample were subject to a position effect. Furthermore, comparing the results from four different models, we consistently find evidence that the probability of an institution being chosen depends not only on the institution itself, but also on its position in the BWS choice task. In accordance with findings in the marketing and psychology literature, in all models, we find that the institution positioned at the top of the choice task stands a significantly higher chance of been identified as being the most trustworthy. While we do find a position bias associated with the worst choice, it is not as strong as compared to the best choice. We also find that the position effects differ between at least two subgroups of consumers.

From a modeling perspective, the consequences of overlooking position bias are clear. Substantial gains in model fit can be achieved and much richer insight into choice behavior and decision rules can be obtained. Importantly, failing to account for this can result in erroneous trust coefficients and ratio scaled probabilities and limit their validity when used for generating policy recommendations. Researchers engaged in the BWS method should be weary of this phenomenon. This should be especially considered at the experimental design stage, where it is possible to factor in that some respondents have an increased tendency of selecting the item positioned at the top when making their best choices and, perhaps, the bottom item when they make their worst choices. The number of items to include per best-worst task is another important consideration. While five items per BWS choice task has been found to be acceptable, we should be cognizant of the fact that this may have been a factor which led to respondents considering only a portion of the information available. With fewer items per task, there may be the potential to reduce these position effects.

Methodological aspects aside, we find, on average, that consumers tend to perceive information about nanotechnology and its use in food production to be most accurate and balanced when it originates from government institutions and scientists compared to non-government organizations, food handlers, friends and family and the media. Our results reveal that consumers can be clearly segmented into at least two separate subgroups on the basis of their trust perceptions—one who considers government institutions and scientists to be most trustworthy and another who appears to mistrust these organizations, but instead perceives non-government organizations, friends and family and the media as being relatively more trustworthy. This insight provides valuable information for those who are engaged in communicating food safety. This is especially important as communication with consumers about emerging food safety concerns may help explain consumers' attitude towards accepting this new technology, which may then affect its adoption in the industry. Our results help ensure communication can be achieved and contribute to more effective and successful awareness campaigns.

Some potential limitations of this study must be acknowledged. Firstly, while we wanted to bring position bias to the fore, we appreciate that there are a number of other decision-making heuristics and processing strategies that we did not address in this paper. This would be particularly important if one aims to explore meaningful differences among heuristics. Secondly, while our latent class segmentation of trust perceptions and position effects afford a readily identification of heterogeneity, we recognize that it would have been possible to further uncover within class continuous variation and/or to increase the number of classes. We also note that socio-demographic variables could, of course, be included as covariates to help establish profiles of respondents. However, both of these would entail considerably more computational effort. Thirdly, for this analysis we do not implement nor compare our results against the models typically used in choice set generation analysis, such as the independent availability logit model (see Swait and Ben-Akiva, 1987; Swait, 2001, for a description), which might be better suited at retrieving which positions were taken into account by respondents. While initial effort was given in this area, with $J = 20$ best-worst pairs in our case, the choice set generation was too complex and computationally burdensome. We leave this challenge for further research, and suggest that it would be more feasible in case studies with fewer items per choice task. Fourthly, while we recognize the value in identifying the reasons explaining the adoption of these position effects, in this paper, we focus only on the identification of such heuristics and how to accommodate them. And finally, we focused only on position effects relating to the vertical dimension of space. An obvious extension to this paper would be to test whether or not similar results would be attained from BWS data based on horizontally arranged choice tasks.

Notwithstanding these potential limitations, our findings provide compelling evidence for further research in this area. While specific to this dataset, we show the repercussions of failing to recognize position effects in the analysis of BWS surveys and provide a practical empirical solution for it. We, therefore, encourage researchers who have already collected best-worst data to investigate whether their data shows such heuristics. Although we explore the issue of position bias in BWS, our approach can easily be adapted to explore other behavioral heuristics that may also be at play in BWS, as well as in other stated and revealed preference studies.

References

Attali, Y. and Bar-Hillel, M. (2003). Guess where: The position of correct answers in multiple-choice test items as a psychometric variable, *Journal of Educational Measurement* **40**(2): 109–128.

Auger, P., Devinney, T. and Louviere, J. (2007). Using best?worst scaling methodology to investigate consumer ethical beliefs across countries, *Journal of Business Ethics* **70**(3): 299–326.

Baumgartner, H. and Steenkamp, J.-B. E. (2001). Response styles in marketing research: A cross-national investigation, *Journal of Marketing Research* **38**(May): 143–156.

Campbell, D., Hensher, D. A. and Scarpa, R. (2012). Cost thresholds, cut-offs and sensitivities in stated choice analysis: identification and implications, *Resource and Energy Economics* **34**(3): 396–411.

Campbell, D., Hensher, D. A. and Scarpa, R. (2014). Bounding wtp distributions to reflect the 'actual' consideration set, *Journal of Choice Modelling* .

Campbell, D., Hutchinson, W. G. and Scarpa, R. (2008). Incorporating discontinuous preferences into the analysis of discrete choice experiments, *Environmental and Resource Economics* **41**: 401–417.

Carlsson, F., Mørkbak, M. and Olsen, S. (2012). The first time is the hardest: A test of ordering effects in choice experiments, *Journal of Choice Modelling* **5**(2): 19–37.

Chandon, P., Hutchinson, J. W., Bradlow, E. T. and Young, S. H. (2009). Does in-store marketing work? effects of the number and position of shelf facings on brand attention and evaluation at the point of purchase, *Journal of Marketing* **73**(6): 1–17.

Cohen, E. (2009). Applying best-worst scaling to wine marketing, *International Journal of Wine Business Research* **21**(1): 8–23.

Cohen, S. and Orme, B. (2004). Asking survey respondents about their preferences created new scaling decisions, *Marketing Research* .

Day, B., Bateman, I. J., Carson, R. T., Dupont, D., Louviere, J. J., Morimoto, S., Scarpa, R. and Wang, P. (2012). Ordering effects and choice set awareness in repeat-response stated preference studies, *Journal of Environmental Economics and Management* **63**(1): 73 – 91.

Dayan, E. and Bar-Hillel, M. (2011). Nudge to nobesity II: Menu positions influence food orders, *Discussion paper series*, The Center for the Study of Rationality, Hebrew University, Jerusalem.

Doornik, J. A. (2009). *An Object-Oriented Matrix Programming Language Ox 6*, Timberlake Consultants Ltd.

Erdem, S., Campbell, D. and Hole, A. R. (2014). Accounting for attribute-level non-attendance in a health choice experiment: does it matter?, *Health economics* pp. n–n.

Erdem, S., Campbell, D. and Thompson, C. (2014). Elimination and selection by aspects in health choice experiments: Prioritisation health service innovations, *Journal of Health Economics (forthcoming)* pp. n–n.

Erdem, S. and Rigby, D. (2013). Investigating heterogeneity in the characterization of risks using best worst scaling, *Risk Analysis* **33**(9): 1728–48.

Erdem, S., Rigby, D. and Wossink, A. (2012). Using best?worst scaling to explore perceptions of relative responsibility for ensuring food safety, *Food Policy* **37**(6): 661–670.

Finn, A. and Louviere, J. (1992). Determining the appropriate response to evidence of public concern: The case of food safety, *Journal of Public Policy & Marketing* **11**(2): 12–25.

Flynn, T. N., Louviere, J. J., Peters, T. J. and Coast, J. (2007). Best–worst scaling: what it can do for health care research and how to do it, *Journal of health economics* **26**(1): 171–189.

Guney, B. (2014). A theory of iterative choice in lists, *Journal of Mathematical Economics* **53**(0): 26 – 32.
URL: *http://www.sciencedirect.com/science/article/pii/S0304406814000664*

Hensher, D. A., Collins, A. T. and Greene, W. H. (2012). Accounting for attribute non-attendance and common-metric aggregation in a probabilistic decision process mixed multinomial logit model: a warning on potential confounding, *Transportation* **40**(5): 1003–1020.

Hensher, D. A., Rose, J. M. and Greene, W. H. (2005). The implications on willingness to pay of respondents ignoring specific attributes, *Transportation* **32**: 203–222.

Hess, S., Stathopoulos, A., Campbell, D., O'Neill, V. and Caussade, S. (2012). It's not that I don't care, I just don't care very much: confounding between attribute non-attendance and taste heterogeneity, *Transportation* **40**(3): 583–607.

Inman, J. J., McAlister, L. and Hoyer, W. D. (1990). Promotion signal: Proxy for a price cut?, *Journal of Consumer Research* **17**(1): pp. 74–81.

Koppell, J. G. and Steen, J. A. (2004). The effects of ballot position on election outcomes, *The Journal of Politics* **66**: 267–281.

Louviere, J. J. (1993). The best-worst or maximum difference measurement model: Applications to behavioral research in marketing, *The American Marketing Association's 1993 Behavioral Research Conference*.

Louviere, J. J. and Flynn, T. N. (2011). Using best-worst scaling choice experiments to measure public perceptions and preferences for healthcare reform in australia, *The Patient: Patient-Centered Outcomes Research* **3**(4): 275–283.

Lusk, J. L. and Briggeman, B. C. (2009). Food values, *American Journal of Agricultural Economics* **91**: 184–196.

Manski, C. F. (1977). The structure of random utility models, *Theory and Decision* **8**: 229–254.

Marley, A. A. J. and Louviere, J. J. (2005). Some probabilistic models of best, worst, and best-worst choices, *Journal of Mathematical Psychology* **49**(6): 464–480.

McArthur, L. and Post, D. (1977). Figural emphasis and person perception, *Journal of Experimental Social Psychology* **13**(6): 520–535.

Meier, B., Hauser, D., Robinson, M., Friesen, C. and Schjeldahl, K. (2007). What's "up" with god? vertical space as a representation of the divine, *Journal of Personality and Social Psychology* **93**(5): 699–710.

Meier, B. P. and Robinson, M. D. (2004). Why the sunny side is up: Associations between affect and vertical position, *Psychological Science* **15**(4): 243–247.

Nisbett, R. E. and Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes, *Psychological Review* **84**(3): 231–259.

Raghubir, P. and Valenzuela, A. (2006). Center-of-inattention: Position biases in decision-making, *Organizational Behavior and Human Decision Processes* **99**(1): 66 – 80.

Rodway, P., Schepman, A. and Lambert, J. (2013). The influence of position and context on facial attractiveness, *Acta Psychologica* **144**(3): 522 – 529.

Rozin, P., Scott, S., Dingley, M., Urbanek, J. K., Jiang, H. and Kaltenbach, M. (2011). Nudge to nobesity I: Minor changes in accessibility decrease food intake, *Judgment and Decision Making* **6**(4): 323–332.

Rubinstein, A., Tversky, A. and Heller, D. (1996). *Understanding strategic interaction—Essays in honor of Reinhard Selten*, Springer-Verlag, chapter Naive strategies in competitive games, pp. 394–402.

Scarpa, R., Gilbride, T. J., Campbell, D. and Hensher, D. A. (2009). Modelling attribute non-attendance in choice experiments for rural landscape valuation, *European Review of Agricultural Economics* **36**(2): 151–174.

Scarpa, R., Notaro, S., Louviere, J. and Raffaelli, R. (2011). Exploring scale effects of best/worst rank ordered choice data to estimate benefits of tourism in Alpine grazing commons, *American Journal of Agricultural Economics* pp. 1–16.

Scarpa, R., Zanoli, R., Bruschi, V. and Naspetti, S. (2013). Inferred and stated attribute non-attendance in food choice experiments, *American Journal of Agricultural Economics* **95**(1): 165–180.

Schubert, T. (2005). Your highness: Vertical positions as perceptual symbols of power, *Journal of Personality and Social Psychology* **89**(1): 1–21.

Shaw, J. I., Bergen, J. E., Brown, C. A. and Gallagher, M. E. (2000). Centrality preferences in choices among similar options, *The Journal of General Psychology* **127**(2): 157–164.

Swait, J. (2001). A non-compensatory choice model incorporating attribute cut-offs, *Transportation Research Part B: Methodological* **35**(10): 903–928.

Swait, J. and Ben-Akiva, M. (1987). Incorporating random constraints in discrete models of choice set generation, *Transportation Research Part B: Methodological* **21**(2): 91–102.

Thurstone, L. L. (1927). A law of comparative judgment, *Psychological Review* **34**: 273–286.

Valenzuela, A. and Raghubir, P. (2009). Position-based beliefs: The center-stage effect, *Journal of Consumer Psychology* **19**(2): 185 – 196.

Valenzuela, A. and Raghubir, P. (2010). Are consumers aware of top-bottom but not of left-right inferences? implications for shelf space positions.

Valenzuela, A., Raghubir, P. and Mitakakis, C. (2013). Shelf space schemas: Myth or reality?, *Journal of Business Research* **66**(7): 881 – 888.

van Kleef, E., Otten, K. and van Trijp, H. C. (2012). Healthy snacks at the checkout counter: A lab and field study on the impact of shelf arrangement and assortment structure on consumer choices, *BMC public health* **12**(1): 1072.

Wright, P. (2002). Marketplace metacognition and social intelligence, *Journal of Consumer Research* **28**(4): pp. 677–682.