Response order effects – how do people read?

Bobby Duffy, Ipsos MORI

Abstract

This paper outlines the results from an experiment examining response order effects with visually presented lists. In particular it examines the implications of the practical response adopted by most market research agencies – to use normal and reversed show cards. The conclusion is that for most questions the effect is likely to be present, but relatively small, and dependent on the extent of context effects. That is, it appears more important to ensure that the most likely responses are not grouped at either end of the show list. The study also identified that a quarter of respondents do not actually read lists they are presented with in interviews from top to bottom, and significant minorities "jump around" lists looking for eye-catching words or phrases. This clearly has implications for interpreting "primacy" effects and for the design and physical appearance of lists.

Introduction

There is a long history of research into response order effects in surveys. The observed effect is generally that when items are presented to respondents in a list, they are more likely to choose those at the top of the list – called a "primacy effect". In contrast, when options are read out to respondents, they are more likely to choose the last items in the list – a "recency effect".

There have been a large number of experiments that have examined these effects, and a number of theories put forward as to why they might occur. Early theories focused on the impact of memory. This can help to explain recency effects seen in questions where a large number of complex responses are read out to respondents, as recall of early items from short-term memory is likely to be more difficult (Schwarz, Hippler and Noelle-Neumann, 1991**)**. However, this is much less satisfactory in explaining those cases where the same effect is observed with a small number of simple items, and where respondents have a written list of items to refer to.

Alternative theories have therefore focused on the cognitive processes that respondents go through when choosing items (Krosnick and Alwyn, 1987). Here primacy effects for visual material are the result of two key factors. Firstly, options presented early in any list may help establish a cognitive framework or standard of comparison that influences interpretation of later options. These early options may therefore assume special significance with respondents. Early items are also likely to be subject to greater cognitive processing so that by the time respondents consider later items their mind may be cluttered with thoughts about previous items, which may in turn prevent full consideration of these later items.

This also relates to the theory of satisficing (Simon 1957, Krosnick 2000), where people will choose adequate answers rather than optimal. It is much easier in an interview to choose the first acceptable answer than to consider the options deeply and choose an optimal response.

However, there are other effects at work here. In particular, there has been work that has highlighted the importance of asymmetric contrast effects – ie where the impact

of an order effect is mediated by where in the list the most extreme or appropriate responses are placed. Here the theory is that introducing a more extreme or appropriate item results in a wider perspective regarding the options, and this will affect evaluation of each subsequent item. If this factor is in play it could therefore confound the predicted order effects, depending on the position of the most extreme or appropriate items.

The study

This paper looks at data from an experiment by the Ipsos MORI Research Methods Unit using visual lists (showcards) in a personal interview. It does not cover the ground again on the origins of order effects, and instead looks at the impact of the practical response from most market research companies to order effects when showing materials to respondents – ie to reverse the order in which items are shown to half the sample. A more complete solution is to design a large set of showcards with responses in all possible orders, but as is recognised in the literature the costs of this are high and it is not practicable for survey work generally.

We also wanted to examine the interaction of contrast and order effects in a bit more detail – again in light of the practical approach that researchers often take to list-making for surveys. In particular, we hypothesised that many show lists may have more popular items near the top, as these will be the first that are thought of by researchers. If this is the case it would lead to more extreme order effects when showcards are simply reversed. That is, given the theories of satisficing and contrast effects, we would expect to see more marked order effects where this is the case.

Finally, we also wanted to check a key underlying assumption in all work on response order for show lists – that people read the material we show them in a predictable way, i.e. from top to bottom. There are good reasons to think this should be the case given how we generally read. However, an interview is rather different to situations when we are normally reading list material – and respondents may themselves learn different reading approaches through taking part in the interview.

As part of this, we also wanted to check the impact of assigning each response item a letter. This again is a common practice in research agencies, as it can lessen the embarrassment of a respondent reading out potentially personal or sensitive responses, as well as helping to speed up the interview process.

The test used Ipsos MORI's Omnibus survey, which provides a large nationally representative sample (around 2,000 people aged 15+ in Great Britain). This allowed us to look at a number of different conditions for each question included. It is also conducted by CAPI, which means that showcards can be easily randomly assigned for each interview. The profiles of the achieved samples receiving each showcard has been checked on key demographic variables, and there are no significant differences in profile. The questions included cover different subject areas, but each is attitudinal, where respondents are required to select appropriate options from a relatively long list.

The impact of item popularity

The first test looked at the interaction between item order and item popularity. This question on the key issues in influencing voting intention has been asked a large number of times in regular tracking surveys, and while there have been some marked changes in responses over the long-term, there is relative stability in responses in the shorter-term. We could therefore predict the most likely responses with some certainty.

Four different versions of the showcard were used. Table 1 shows the results from two versions where the list order was effectively randomised on popularity, by presenting responses in alphabetical order for the normal card (and reverse alphabetical order for the reversed card). This shows no evidence of any order effect, with no differences significant at the 95% confidence level.

Table 1 – Looking ahead to the next General Election, which of these issues, if any, do you think will be most important to you in helping decide which party you vote for?

	Alphabetical	**Reversed alphabetical**
Base: All respondents	*(488)*	*(435)*
	%	*%*
Animal Welfare	14	16
Constitution/devolution	8	7
Defence	22	24
Education	59	62
Europe	27	26
Health Care	74	77
Housing	26	28
Law and Order	62	60
Managing the economy	35	38
Northern Ireland	8	9
Pensions	45	40
Protecting the natural environment	31	29
Public transport	41	38
Taxation	35	36
Trade Unions	7	5
Unemployment	36	32
Other	3	1
Don't know	3	4

Source: Ipsos MORI

However, our hypothesis is that the impact of order interacts with the popularity of responses – in this case, that response order effects would be much more marked where the most popular responses were all at the top of the list for the standard version and the bottom for the reversed card.

This does seem to be supported by results from the second two versions of the showcard as shown in Table 2, where responses were presented in the order of popularity we predicted from previous studies (and reversed for those receiving the reversed card). In general those who received the reversed card are much more likely to choose items that are near the middle or top of their list, with housing and unemployment standing out particularly. They also make a higher number of selections on average than those who were shown the card in popularity order.

Table 2 – Looking ahead to the next General Election, which of these issues, if any, do you think will be most important to you in helping decide which party you vote for?

	Popularity order	**Reversed popularity order**
Base: All respondents	*(536)*	*(516)*
	%	*%*
Health Care	80	77
Education	59	59
Law and Order	56	56
Pensions	38	**50**
Taxation	37	41
Public transport	35	41
Unemployment	34	**42**
Managing the economy	33	37
Protecting the natural environment	29	34
Europe	28	29
Housing	27	**37**
Defence	17	20
Animal Welfare	13	18
Northern Ireland	10	11
Constitution/devolution	8	9
Trade Unions	7	9
Other	2	2
Don't know	3	2

Source: Ipsos MORI

These effects are consistent with the cognitive theories that have been applied to order effects, and satisficing in particular. Those who receive the reversed showcard will need to read through a number of relatively low salience options before coming to acceptable responses. They can see there are relatively few options left available to them and will be more likely to choose adequate categories that appear relevant enough. When they get to most appropriate items, at the end of their list, some who will have already chosen less relevant items will also select further items, raising the average number of responses.

Subject salience – and showcard lettering

Prior to carrying out this new experimental work, we examined a number of questions asked in previous Ipsos MORI studies, and in general these show very little impact from item order alone. We therefore wanted to replicate a study where order effects had been seen to be significant. This was from a study by Krosnick and Alwin (1987), and used a question asking which of a relatively long list of attributes are most important for children to possess.

Table 3 compares our results with those achieved in the previous study. The first point to note here is that there is some significant variation in responses between the two studies. In a sense this is to be expected, given the number of years between the two studies, and the fact that the Krosnick and Alwin study was conducted in the US, while the Ipsos MORI study was in Great Britain. In particular, the greater apparent emphasis on manners in Britain may have been predicted, although it is less clear why good sense and sound judgement should be more valued in the US.

Table 3 – The qualities listed on this card may be important, but which three would you say are the most desirable for a child to have?

	Krosnick and Alwin – standard order	**Krosnick and Alwin – reversed order**	**Ipsos MORI – standard order**	**Ipsos MORI – reversed order**
Base: All respondents	*(905)*	*(446)*	*(1,014)*	*(969)*
	%	*%*	*%*	*%*
(A) Has good manners	**26**	10	**51**	33
(B) Tries hard to succeed	**19**	15	16	**20**
(C) Is honest	**66**	48	**57**	48
(D) Is neat and tidy	7	7	3	5
(E) Has good sense and sound judgement	39	41	20	**25**
(F) Has self-control	12	14	18	20
(G) He acts like a boy or she acts like a girl	3	3	1	2
(H) Gets along well with other children	14	11	17	18
(I) Obeys his parents well	31	**37**	26	25
(J) Is responsible	34	33	25	26
(K) Is considerate of others	25	**40**	38	37
(L) Is interested in how and why things happen	18	**25**	16	**23**
(M) Is a good student	7	**16**	4	**10**

Source: Ipsos MORI

The order effect appears to be less clear-cut in the new Ipsos MORI version of the question than seen in the earlier study. It is not clear why this should be the case. The US study was run on the GSS, which provides a large, reliable sample, and is in many ways very similar to an Omnibus study – a relatively long interview covering a range of different topics.

In any case, there is still a greater effect with this question than seen in the previous question on political issues. This is likely to be related to the nature of the questions, where the question used here relates to a construct that is rather hard for respondents to define, and where opinions are likely to be rather less well-formed. These type of questions, where the subject has not been considered previously are likely to be more subject to satisficing, given the greater cognitive effort to respond optimally. There is also likely to be some impact from the question wording, where in the political issues question there is no requirement to pick a certain number of categories, while in the Krsonick and Alwin question there is. It is likely that requiring a particular number of selections will encourage satisficing among those who have no strong opinions on the subject.

However, our main purpose for asking this question was to test the impact of another widely adopted practice in commercial research – the lettering (or numbering) of items that appear on showcards. The aims of this are to help speed up the process of interviews, and to avoid respondents having to read out potentially sensitive or embarrassing responses.

Our hypothesis here is that this may encourage respondents to read reversed showcards from bottom to top rather than top to bottom. This could, for example, be because they have a general propensity to read in an alphabetical order, or because they take this as an indication of how the researcher intends they should read the list.

If this factor is at work, we would expect to see less of a primacy effect from showcard reversal where items are labelled. However, the results shown in Table 4 appear to disprove this. There is little difference in the proportions of respondents who received the labelled and non-labelled reversed cards selecting each category, and no sign that response order effects are lessened.

Table 4 – The qualities listed on this card may be important, but which three would you say are the most desirable for a child to have?

	Standard order	Reversed order	Standard order - labelled	Reverse order – labelled
Base: All respondents	*(468)*	*(453)*	*(541)*	*(513)*
	%	*%*	*%*	*%*
(A) Has good manners	**48**	31	**53**	34
(B) Tries hard to succeed	18	19	16	**22**
(C) Is honest	**55**	45	**59**	49
(D) Is neat and tidy	3	4	2	5
(E) Has good sense and sound judgement	21	26	19	**27**
(F) Has self-control	17	20	17	20
(G) He acts like a boy or she acts like a girl	1	2	1	3
(H) Gets along well with other children	19	17	17	18
(I) Obeys his parents well	26	25	25	24
(J) Is responsible	26	26	24	26
(K) Is considerate of others	39	38	38	36
(L) Is interested in how and why things happen	16	**26**	16	**22**
(M) Is a good student	4	**10**	3	**9**

Source: Ipsos MORI

But how do people read?

We also asked respondents directly which order they read the list they were shown in this previous question (Table 5). This does not appear to have been questioned in previous work on order effects – but it is clearly critical to explanations of observed effects. If people read lists differently to the natural assumption, then even the description of a "primacy" effect is misleading.

Table 5 – Thinking about the last question that you just answered from the card, how did you read the list? Did you read it from top to bottom, from bottom to top or in some other way?

	All	Standard order	Reversed order	Standard order – labelled	Reversed order – labelled
Base: All respondents		*(468)*	*(453)*	*(541)*	*(513)*
	%	*%*	*%*	*%*	*%*
Top to bottom	74	**77**	70	**82**	66
Bottom to top	13	11	**16**	6	**19**
Other	11	10	13	9	13
Don't know / can't remember	2	2	1	2	2

Source: Ipsos MORI

Further, there are many good reasons why reading in real-world interviews may not be the same as how we might normally read. In particular, most interviews involve presenting respondents with a large amount of information in a short period, which is likely to lead to hurried reading strategies. There is also likely to be some impact from the presence of an interviewer, where respondents are conscious that they are waiting for a response, and so attempt to assimilate the information as quickly as possible. Some respondents may also learn from previous questions in longer interviews, particularly if there is some bias in the "popularity order" of responses.

The results for the sample as a whole and for each showcard version are shown in the table below. This confirms that most people do indeed read lists from top to bottom. But this is not the case for a quarter of all respondents, and a significant minority read the list from bottom to top, while others apply a variety of techniques.

These other approaches were coded from verbatim responses into a common themes, and over half of these (7% of the total population) say they read the list "randomly", "browsed", "dodged" or "jumped around", which is consistent with satisficing strategies. The other common approaches are to start in the middle then read up, then down – or to read from the middle down then middle up. It is also worth noting that significant minorities look for responses that stand out in some way, to help them get to grips with the information – with the length of the line or eye-catching words being key here. Clearly some will not have perfect recall of how they read the list, and many will not be aware of their reading strategy. However, it is likely that this will only serve to underestimate the proportion that are actually unconsciously influenced by the appearance of particular statements or words.

Table 5 also highlights the impact of reversing showcards on the order in which people read. That is, those who received the reversed list are more likely to read the list from bottom to top than those who received the standard order card. The most likely explanation for this is that respondents who receive reversed showcards have learnt from previous questions that the most likely responses are near the bottom of their lists.

The findings also lend some weight to the case for the mitigating impact of lettering. In particular, those who were shown the lettered reversed list are more likely to read the list from bottom to top than other groups. Similarly the group who are most likely to read top to bottom are those who were shown the standard labelled list.

There are only a few significant differences between respondents of different educational levels (as seen in Table 6). Perhaps predictably those with higher educational levels are slightly less likely to read the lists from top to bottom and more likely to read in another way.

Table 6 – Thinking about the last question that you just answered from the card, how did you read the list? Did you read it from top to bottom, from bottom to top or in some other way?

	No quals	**GCSE**	**A-level**	**Degree**	**Higher degree**
Base: All respondents	*(595)*	*(430)*	*(216)*	*(213)*	*(56)*
		%	*%*	*%*	*%*
Top to bottom	**78**	74	69	76	64
Bottom to top	13	13	15	9	14
Other	**6**	12	14	**16**	**21**

Source: Ipsos MORI

Finally, it is also worth looking at how the response pattern varies for those who received the reversed showcard, depending on the order in which they read the items (Table 7). Firstly, there are in fact not that many statistically significant differences between the findings. However, good manners, which appears at the bottom of the list on this reversed card, is much more likely to be chosen by those who read bottom to top and much less likely to be chosen by those who read the list less systematically. While not significant given relatively small base sizes, there is also some indication that those who read the list in other ways pick out the longer statements, such as E, H and L, as suggested in the verbatim comments.

Table 7 – The qualities listed on this card may be important, but which three would you say are the most desirable for a child to have?

	All reversed	Top to bottom	Bottom to top	Other
Base: All respondents who were shown the reversed card only	*(969)* %	*(661)* %	*(169)* %	*(124)* %
(A) Has good manners	33	32	**48**	**20**
(B) Tries hard to succeed	20	21	18	20
(C) Is honest	48	49	46	44
(D) Is neat and tidy	5	4	7	7
(E) Has good sense and sound judgement	25	25	21	32
(F) Has self-control	20	20	23	22
(G) He acts like a boy or she acts like a girl	2	2	3	5
(H) Gets along well with other children	18	17	18	23
(I) Obeys his parents well	25	24	23	32
(J) Is responsible	26	27	26	21
(K) Is considerate of others	37	37	37	40
(L) Is interested in how and why things happen	23	23	17	27
(M) Is a good student	10	12	9	7
Don't know				

Source: Ipsos MORI

Conclusions

This experiment has confirmed the relatively important impact of primacy effects with visual presentation of material. However, it is clear that the extent of the effect is dependent on the nature of the question, and it appears much less important for questions where respondents are likely to already have given some thought to the issues. In contrast, therefore, there is a particular need to give thought to the possible impact when asking questions on issues that respondents may not have considered in any detail prior to the interview. Overall, however, there does not seem to be a need for more costly and time-consuming measures such as randomising response categories.

Having said that, there is a very clear and important interaction between response order and contrast effects. Of particular practical importance, attention needs to be given to the likely popularity order of lists. While it will not always be possible to predict this accurately, in most cases we are likely to be able to guess – and lists should then be randomised on the basis of this predicted popularity.

There are also still lessons to be learnt about how people read lists in interview situations. A significant minority does not read lists as we would expect, and the unpredictable effects of this could be reduced by simple instructions to respondents to read through the list thoroughly before making their selection. It is also more important than is recognised to ensure that particular items on lists do not stand out, particularly in terms of statement length.

References

Krosnick, J. (2000) *The threat of Satisficing in Surveys: the Shortcuts Respondents take in Answering Questions* in Survey Methods Newsletter Volume 20 No.1

Krosnick , J. & Alwin, D. (1987) An Evaluation of a Cognitive Theory of Response Order Effects in Survey Measurement. *Public Opinion Quarterly* Volume 51 Issue 2

Schwarz, N., Hippler, H-J. & Noelle-Neumann, E. (1991) *A Cognitive Model of Response-Order Effects in Survey Measurement*

Simon, H.A. (1957) *Models of Man*. New York: Wiley