PRESENTING PUBLICATIONS IN ONLINE PRINT AUDIENCE QUESTIONNAIRES

John Carroll, Sandra Collins and Neil Farrer, Ipsos MediaCT

Introduction

It is recognised that respondent engagement is an important factor in maximising the quality of data completion in surveys. This is especially so where the Internet is used as the mode of data collection. Interviewer presence on face-to-face surveys helps to focus the respondent's attention, which is particularly important in print measurement research where the requirement may be to present a lengthy list of titles. Therefore a means of replicating this high level of attentiveness in an unsupervised online environment is needed.

Some questions arising on how one should present publications to respondents include:

- How can we encourage active participation? The technical capabilities of the Internet, e.g. drag and drop, could be exploited to help maintain respondent interest, however, this may alienate less Internet-savvy respondents.
- How can we ensure that it is speedy to complete? Minimising the number of clicks to answer questions could facilitate a timely completion. On the other hand, this could mean less time thinking about past reading occasions and deflate readership estimates.
- How simple can we make the design? Easy to complete surveys will appeal to all types of people. However, this may mean a bland design and switch off respondents.

In this paper, we share the results of an experiment devised and conducted by Ipsos MediaCT designed to try and provide insight to these questions. The objective of this test was to enhance our understanding of what works better and what does not work so well, relative to the alternative versions tested, for presenting a print media list in online self-completion surveys. The results provide a contribution to online questionnaire design considerations in print audience surveys.

Test design

In the following, we describe the four alternative layout designs we tested before giving an overview of the sample and questionnaire used.



Version 1: 'Multiple logos'

In this design, 8 publications' colour mastheads were presented on each screen in a 4 column x 2 row matrix. This number was chosen to minimise screen-load and also to reduce the risk of scrolling. There were 16 screens comprising 2 screens for daily newspapers, 2 screens for Sunday newspapers and 12 screens for magazines. Clearly, there is a number of alternative layout styles one could consider for this particular design, however, that would require a separate test of its own. The requirement overall was to have four different and distinct versions for testing. A progress bar was also included on the screen (as with all four versions).

First, respondents were asked to click on the title(s) they had read in the past 12 months, or to select 'none of these', before proceeding to the next screen. After each of the daily newspaper, Sunday newspaper and magazine sections, there was a review screen, where respondents could check their selection and have the opportunity to make any changes. This was then followed by individual screens asking first recency and then frequency questions for each title selected as read in the past 12 months (see Appendix).

We hypothesise that the benefit of this design is that it presents a number of titles together on the screen in a clear and simple layout, aided by the use of colour mastheads. However, a possible drawback with this design is that only those titles selected were confirmed as being read, i.e. a response was not required to confirm that a particular title had not been read.

Version 2: 'Drag & Drop'

Have 12 m Drag	e you read or looked at any nonths? I publication into the relev	v of the following newspapers fo	click here for negoti budyes decourter technical problem r at least 2 minutes in the past
	(Read/looked at past 12 months)	16 of 16 remaining	(Did not read/look at past 12 months)
		Baily Mail	
		Mon-Sat	
		Go to the next question 🛁	

In this version, the 128 publications' colour mastheads were presented one at a time in the centre of the screen and respondents were asked to click on these and drag into the appropriate 'read' or 'not read' box on either side. The selections remained visible in each of the response boxes while each new title appeared in the centre. There was also a guide to how many titles were left to appear. A review screen appeared after each sequence of daily newspapers, Sunday newspapers and magazines, which again provided the opportunity to amend any responses; it also helped to break up the repetition. This was then followed by individual screens asking first recency and then frequency questions for each title selected as read in the past 12 months. These questions were also answered using the drag and drop method.

We see the possible benefit of this design as its ability to exploit the functionality of online survey software and provide a more engaging and active role for the respondent. However, a potential drawback is that it could alienate some Internet users, particularly those who are infrequent users. Also, the repetitive nature of the task may lead to pattern responses and a higher rate of drop out.

Version 3: 'Click box'



This design, which was similar in layout to the Version 2 design, asked respondents to click directly onto the 'read' or 'not read' boxes rather than dragging each title across. Again, the 128 titles were presented one by one on the screen. When the response box was selected, it changed colour to confirm the selection had been made and then the next title appeared on the screen. The previous selections did not remain visible during the sequence. Respondents did have the opportunity to review and change their selections after each of the three daily newspaper, Sunday newspaper and magazine sections.

For this version, the recency and frequency questions were asked of the individual titles, but presented as a build on the screen. First, the recency question was asked and when the response was made, the frequency question appeared below on the same screen.

A potential benefit with this design is that it may facilitate a timely completion with the respondent having the relatively straightforward task of clicking in one of two boxes as each title appears on the screen. We did consider having the next title appear instantaneously with each response, however, this did start to appear confusing when responses were being made quickly. We therefore allowed for the response box to change colour to confirm the selection and so a delay of around one second was introduced between each title. The potential drawbacks to this design are similar to those with Version 2; the repetitive nature of the task could lead to pattern responses and ill-considered claims.

ave you re	ad or looked at any of the onths?	e following newspapers f	or at least 2 minute	s in the
lease sel	ect "Yes" if you have read	d or looked at this news	paper in the past 12	2 months
"No" if ye	ou have not		n na an ann an 1873 an a 1973 ann an 1873 ann	
The	e Sundav Times	. <u>Anân</u> a	Yes No	
		THE SUNDAY TIMES		_
Su	nday Telegraph	The Sunday Telegraph	Yes No	
Da	ily Star Sunday	STAR	Yes No	
-		USUNDAY		
The	e Observer	TheObserver	Yes No	
The	e Independent on Sunday	THE *** INDEPENDENT	Yes No	
				_
Su	nday Express	SUNDAYSEXPRESS	Yes No	
		an and an draw a lat	Vac Na	
Su	nday Herald	sundayneraid	165 100	

Version 4: 'Traditional list'

In a design perhaps not too dissimilar to a paper questionnaire layout, a number of publications were shown on the screen as a list and the respondent was asked to select a 'yes' or 'no' response to each title they had read or not read in the past 12 months. Publications were shown in both typescript and colour masthead form. Six screens were shown to each respondent. The daily

newspaper section comprised one screen of 16 titles which required the respondent to scroll down the list; one screen of 16 Sunday newspaper titles; and four screens each of 24 magazines. This was followed by separate screens for recency and frequency questions, each presented as grids to the respondents showing only those newspapers and magazines read in the past year. It was decided not to include a review screen for this design as respondents were able to see their full selection on the screen as they went along.

In self-completion surveys, people tend to be familiar with grid designs on paper questionnaires and therefore this version's apparent simple layout could appeal to all types of people. However, a potential drawback is that it may be perceived as being a bland design and grids could appear off-putting in web surveys and lead to 'straightlining' responses.

Sample

The test was conducted using the Ipsos Online Access Panel with the four versions of the questionnaire each asked of matched samples of 500 adults. Fieldwork was conducted between 15 and 24 May 2009. The panellists received a standard survey payment for taking part.

Each version of the survey was randomly assigned and sent to around 6,000 panellists in order to deliver the required 500 usable responses. Quotas were applied for sex interlocked with age, with further minimum quotas set for social grade and geographic region. Figure 1 summarises the panel response by version and also the age and sex sample profile indexed on the target profile.

	VERSION 1	VERSION 2	VERSION 3	VERSION 4
Mailout	6,200	5,850	5,680	6,140
Final sample	504	512	508	503
	8%	9%	9%	8%
Indices -				
Male	97	96	96	97
Female	103	104	104	103
18-34	109	107	109	109
35-54	95	96	95	95
55+	96	95	94	93

Figure 1: Analysis of sample

Indices show unweighted sample profile / target sample profile

The relatively low return rate of around 8% to 9% highlights a more general issue of representativeness of online panel surveys, however, this sample provided the most cost effective source for our small budget test. Putting any issues of non-response bias to one side, it was reassuring to observe that there were no notable differences in return rate by questionnaire version. Although there were some differences between the achieved profile and the target profile within each version, e.g. with a bias towards women and younger adults, the unweighted sample profile was similar across the four samples.

The data were weighted to represent the profile of the online population for age within sex, social grade and geographic region based on data from the British National Readership Survey (NRS).

Questionnaire

Although we are focussing on print readership for the purpose of this paper, the questionnaire included other topics to help us with our understanding of the respondent profile and the readership data, plus to help simulate a typical length of a readership interview. The survey was constructed with the following question topics:

- 1. General lifestyle (10 statements)
- 2. Usage of the Internet
- 3. Visited past year, recency & frequency of visiting 24 websites (11 newspaper and magazine publisher websites, 13 general websites e.g. news, shopping, social)
- 4. Read past year, recency & frequency of reading 16 national daily and 16 Sunday newspapers
- 5. Read past year, recency & frequency of reading 96 magazines (28 general titles, 44 women's titles and 24 men's titles)
- 6. Basic demographic information
- 7. Feedback on the survey (pre-coded plus open-ended)

The questions and media lists were common across all four versions of the questionnaire. The four different layout designs applied solely to the three website and print readership sections (3, 4 and 5 above). There were two rotations of the media list for each version: Forward (websites, daily newspapers, Sunday newspapers, magazines) and Reverse (magazines, Sunday newspapers, daily newspapers, websites).

The basis for the media list selection was to ensure we included a cross-section of publications. Specifically, we included all national and some regional daily and Sunday newspapers. Also, we included the popular, general interest magazines (including TV listings) to maximise opportunities for positive reading claims and therefore for respondents to route through to the recency and frequency questions. We also included both male and female-oriented magazines.

Results

In the results summarised below, we first analyse respondent factors, such as completion rate, time taken and qualitative feedback. We then look at the readership claims.

Completion rate

One aspect of the assessment we want to consider is whether there were any notable differences in the drop out rate for the four versions of the questionnaire. Figure 2 below shows the percentage of respondents who started the survey and went on to complete it.

Figure	2.	Survey	Comp	letion	Rate
rigure	4.	Survey	Comp	enon	nuie

VERSION 1	VERSION 2	VERSION 3	VERSION 4
88%	83%	86%	92%

The two versions with the lowest completion rate were the 'drag and drop' and 'click box' versions, both designed to exploit the functionality of online survey design. However, the drawbacks as noted earlier suggest that the repetitive nature of responding to the individually-presented 128 titles may have contributed to the relatively low compliance rate.

Time taken

Figure 3 below shows the average time taken in minutes and seconds to complete the survey for each version. The times are also shown for both the print readership and website visiting sections i.e. those questions with different layouts, and the remaining questions which had common content and layout across the versions.

	Total	Print Readership	Website Visiting	Other Questions
		(Varied Layout)	(Varied Layout)	(Common Layout)
Version 1	10m 34s	6m 03s	2m 29s	2m 02s
Version 2	11m 28s	7m 04s	2m 18s	2m 06s
Version 3	15m 30s	10m 07s	3m 23s	2m 00s
Version 4	10m 14s	6m 04s	2m 01s	2m 09s

Figure 3: Average Time Taken (minutes and seconds)

Versions 1 and 4 with their relatively straightforward multiple title layouts took the shortest time to complete and, as seen earlier, had the lowest drop out rates. Versions 2 and (especially) 3 with their interactive functionality took notably longer. As noted earlier, we introduced a short delay on Version 3 between each title to minimise any possible confusion, however, this resulted in a completion time for the print readership section that was around 70% longer than for Versions 1 and 4. As may have been expected, the commonly laid out questions were completed in a similar time across the four versions.

Analysis by the two rotations within each of the four versions shows that the reverse rotation, where websites and then magazines were shown first, was consistently longer across all versions and this was particularly the case for Version 1 (see Figure 4). Considerations such as title confusion and order effects fell outside the scope of this test but were possible contributory factors to this observed variation.

Figure 4: Time Taken for Print Readership Questions (Index of reverse rotation / forward rotation)

Version 1	126
Version 2	110
Version 3	110
Version 4	106

A question could be asked as to what we are looking for as a 'measure of success' regarding survey length. The longer the survey length may mean that the design is allowing the respondent to give considered answers to the questions. However, other panel research has shown (Puleston & Sleep, 2008) that for lengthy online surveys there is a decline in the quality of response towards the end of the survey. Our intention here is to design a layout that facilitates a relatively speedy completion. This suggests we should favour Versions 4 and perhaps 1 on this measure, but we will revisit this having examined the readership data.

Respondent feedback

A question was added at the end of the survey asking respondents to rate their experience of completing the questionnaire using a short battery of agree/disagree statements. Overall, respondents were favourable towards all four questionnaire versions with 75% of all responses being positive; only 7% were negative responses. Figure 5 below summarises the mean scores to the agree/disagree eleven-point scale for each statement for each version; the scale ranged from 0 (disagree strongly) to 10 (agree strongly).

	Version 1 (sample=504)	Version 2 (512)	Version 3 (508)	Version 4 (503)
The questions were easy to understand	9.2	9.1	9.3	9.2
I found the questionnaire too long	3.0	2.9	4.0	3.0
I was unsure how to complete some of the questions	1.3	1.5	1.4	1.4
I enjoyed taking part in the survey	7.8	7.8	7.7	7.8
I got a bit bored	3.0	2.9	3.8	3.1
I liked the look and style of the questionnaire	8.0	8.3	8.0	7.9
I would take part in more surveys similar to this	8.7	8.7	8.5	8.5

Figure 5: Mean	agreement to survey	<i>y feedback statements</i>	(10 = agree; 0)	= disagree)
		J		

Overall, the mean scores for each statement were similar across all four versions. However, we noted that Version 3 respondents were relatively more likely to agree with 'I found the questionnaire too long' and 'I got a bit bored'; although it was also noted that they did not appear to enjoy taking part any less. Version 2 scored marginally higher for 'I liked the look and style of the questionnaire' and 'I would take part in more surveys similar to this'. We also noted for Version 2 that it delivered the lowest mean score for 'I found the questionnaire too long', yet it was the second longest of the four versions to complete. This coupled with its relatively low score for boredom and high score for future participation intent, suggests we may have achieved a certain level of respondent engagement with this particular layout on this comparison test.

In summary, on the basis of the above analysis of the respondent factors covering completion rate, time taken and respondent feedback, Versions 1, 2 and 4 appear to be favoured over Version 3, which met with some negative response. At this stage, there is a suggestion that exploiting the functionality of the online survey software does not necessarily translate into respondent engagement - Versions 2 and 3 are examples that show what perhaps can work well and not so well in this regard.

In the following, we take a look at some of the readership results.

Read past year (RPY)

Caution is required when one analyses individual absolute readership data on samples of 500 online survey panellists, so we grouped the titles together and looked at the weighted gross read past year percentage figures to provide us with indications of where any disparities may lie. We also looked at the corresponding figures on the British NRS for comparison and to this end all the figures are based on only those titles common to both surveys -29 out of 32 newspapers; 88 out of 96 magazines (see Figure 6).

	Version 1	Version 2	Version 3	Version 4	GB NRS
	(sample=504)	(512)	(508)	(503)	(25,117)
Newspapers	367	530	464	544	390
- Daily	227	318	297	330	241
- Sunday	139	213	167	214	149
Magazines	520	836	718	796	481
- General	127	222	185	205	119
- Women	298	475	415	448	263
- Men	95	138	118	144	99

Figure 6	. (Fross	Read	Past	Year	(%)
rigure 0	. U	JIUSS	пеши	1 usi	reur	10	,

The British NRS data have been filtered on those with Internet access in the last 12 months (Apr 08 – Mar 09, 34.9m adults 15+) Figures are based on weighted data

Version 1 with the 8 masthead logos per screen delivered the lowest gross RPY percentage results for both the newspapers and particularly the magazines. This was also the version with a multiple title layout which did not require a positive response for non-reading of a title. Versions 2 and 4 recorded the highest levels of claims of the four versions. Comparing these results with the corresponding data from the British NRS, filtered on those accessing the Internet in the last 12 months, suggests that Version 1 produces results closest to the currency estimates. However, sample variation and panel bias may be affecting these comparisons.

For three of the versions, respondents were given the opportunity to review and amend their RPY claims; with Version 4, respondents were able to see all their confirmed claims on the screens as they went along. For Version 1 there were a total 137 claim changes, 109 for Version 2 and 621 for Version 3. On these figures alone, it suggests that the layout of the 'click box' design we used confused respondents, inhibited their recall of past 12 month reading and affected their ability to accurately record their reading behaviour.

Recency and frequency of reading

As noted earlier, there were four very different layouts for the recency and frequency of reading questions, which were designed to complement, where possible, the preceding 'read past year' screen layouts. Given the smaller sample sizes available for analysis of recency and frequency data (responses were filtered on those making an RPY claim for each title), we are limited in our opportunities for a detailed assessment of any differences. We therefore looked at the ratios of average issue readership (AIR) and frequency of reading to read past year (RPY) for newspapers and magazines, again with a comparison to the British NRS and including only those titles common to both surveys (see Figure 7).

	Version 1 (sample=504)	Version 2 (512)	Version 3 (508)	Version 4 (503)	GB NRS (25,117)		
Gross AIR / RPY:							
- Newspapers	0.36	0.25	0.27	0.28	0.32		
- Magazines	0.45	0.40	0.34	0.39	0.39		
Gross Almost Always / RPY:							
- Newspapers	0.25	0.20	0.23	0.21	0.25		
- Magazines	0.19	0.17	0.13	0.15	0.23		
Gross Quite Often /	RPY:						
- Newspapers	0.25	0.26	0.24	0.22	0.19		
- Magazines	0.26	0.28	0.25	0.24	0.19		
Gross Only Occasionally / RPY:							
- Newspapers	0.49	0.55	0.53	0.57	0.56		
- Magazines	0.55	0.55	0.62	0.61	0.58		

Figure 7: Ratios of Average Issue Readership and Frequency of Reading to Read Past Year

The British NRS data have been filtered on those with Internet access in the last 12 months (Apr 08 – Mar 09, 34.9m adults 15+) Figures are based on weighted data As we saw earlier, Version 1 delivered the lowest gross RPY results, however, as may be seen in Figure 7, this version recorded the highest conversion to average issue readership of the four versions. It also yielded the highest conversion to 'Almost Always' frequency of reading, but picked up relatively fewer infrequent readers – a possible direct consequence of the multiple title layout design. We note that the Version 1 design for the recency and frequency questions showed the response options in a vertical list (see Appendix), which may have introduced a primacy effect, but this is conjecture. The remaining three versions recorded similar results to each other for these average issue readership and frequency of reading ratios.

Comparing the four test versions to the British NRS, Version 1 produced similar ratios for AIR / RPY and Almost Always / RPY for newspapers.

Conclusion

This has been an interesting and informative small-scale exercise in questionnaire design. We have learnt that different versions can be favoured on different respondent engagement criteria. Readership results from online questionnaires are sensitive to variations in the details in the questionnaire design. This may have wider implications, as not only online samples, but also online questionnaires could be prone to changing audience estimates when switching the mode to online interviewing. It is apparent that our assessment criteria only cover respondent engagement and quality of readership estimates up to a point. Our conclusions are confined to the four specific designs used in our test. Aspects such as title confusion and order effects cannot be addressed here, but may form the basis of any subsequent testing.

As noted earlier, the objective of this test was to enhance our understanding of what works better and what works less well, relative to the alternative versions tested. We also take the following specific learnings from the four alternative designs used in the test:

- Presenting multiple rather than single publications on a screen improves both respondent engagement and quality of the (read past year) data.
- Exploiting the functionality of the online survey software does not necessarily translate into respondent engagement; however, with careful application there may be a place for using interactive (e.g. drag and drop) techniques if extensive repetition can be avoided.
- Aspects of the design that inhibit a speedy completion, or frequently disrupt the flow of the survey, should be avoided.

There is scope for further testing work in this area, for example using a non-access panel sample and considering further alternative versions. This test has been successful in that it has added to our knowledge but we recognise that it is just another piece in helping to shape the online print audience questionnaire design jigsaw.

References

Puleston, Jon and Sleep, Deborah. "Measuring the value of respondent engagement." ESOMAR Panel Research 2008

Appendix: Recency and Frequency Screens



Version 4:

Go to the next question 🔶

	P1125555555555555		0%	25% 50%	75% 100%		Click shoul tech	there for support Id you encounter inical problems		os		0% 25%	50% 75% 100%		Click here for support should you encounter technical problems	
	(Which of	these best des	scribes how often y	ou read or look at	this publication?		
	١	When did you last read or look at any copy of this publication?								Please select an answer for each title						
	1	Please select an answ	er for each tit	tle								Almost Always (at		Only Occasionally		
		TV ₂ Satellite	Yesterday	Past 7 days	Past 4 weeks	Past 3 months	Longer ago					least 3 issues out of 4)	Quite Often (at least 1 issue out of 4)	(less than 1 issue out of 4)	Not in the past 12 months	
	-	la oct								TV&	Satellite					
	-	Desu									est					
		Bella								B	ella					
			Yesterday	Past / days	Past 4 weeks	Past 3 months	Longer ago					Almost Always (at least 3 issues out of 4)	Quite Often (at least 1 issue out of 4)	Only Occasionally (less than 1 issue out of 4)	Not in the past 12 months	
Go to the next question																

Go to the next question 🔶