



Global – GMG

Response to
Survey Quality Moderator Report

Philip Malivoire

March 2013

1. Executive summary

- I have been asked to review the Survey Quality Moderator Report published in January 2013 in respect of the Existing Customers and Lost Opportunities surveys carried out in the Global / RSL merger inquiry. As the Survey Moderator's report focused on the Existing Customers survey, my report also only focuses on the Existing Customers survey.
- I have been asked to confirm whether the Survey Moderator's concerns about the Existing Customers survey undermine the robustness of the Existing Customers survey's key conclusions about what media alternatives would have most likely been used by Global and RSL's existing customers if Global and RSL's stations had not been available when they booked a recent campaign.
- The core of my review consisted of listening to 124 interviews whilst tracking the answers that had been reported on a spreadsheet provided to me by RBB in order to compare the answers given by the respondent with the answers reported on the spreadsheet. I listened to 37 interviews in their entirety and 87 from question 9 to question 12.
- My view is that the vast majority of respondents understood the questions sufficiently and that the answers can be used with confidence to infer advertisers' next best alternatives to the parties' stations. In short, I am satisfied that the specified additional / different media and stations do accurately represent the alternative media choices given by the respondent.
- In order to reach this conclusion properly, I believe it has been necessary to carry out the extent of work described above. My view is that a review of five interviews is not sufficient to understand the impact of issues relating to the quality of interviewing.

2. Background to report

In May and June 2012, YouGov was commissioned by Slaughter and May to conduct two market research surveys in order to obtain an understanding of the media options that advertisers considered were effective alternatives for Global and RSL stations. YouGov surveyed two different target audiences: 'Existing Customers', i.e. those who have used a Global or RSL station in the past 18 months and 'Lost Opportunities' i.e. advertisers that had considered using a Global or RSL station but ultimately declined campaign proposals.

At the Competition Commission's request, a Survey Quality Moderator Report was produced in January 2013. However, the Survey Moderator's report focussed on the Existing Customers survey only. As such, this report also only focusses on the Existing Customers Survey.

The key aim of the Existing Customers survey was to identify which media options are the "next best alternative" to the merging parties' stations – i.e. which media alternatives an advertiser would be most likely to use instead of those stations – as obtained at question 11 of the Existing Customers survey. As question 11 is central to this report, I set out question 11 below:

Q11a: I would now like you to think back to the time when you were planning how to allocate your media budget for the last campaign when you used [insert [Global/GMG] stations identified at Q8]. If you had been unable to use [insert

[Global/GMG] stations identified at Q8] in that period (for example, if there had been no airtime slots available at all), what would you have done with the budget you would have spent on [that station/those station(s)] [multiple responses allowed, prompted questions]

- Increase my advertising spend on other media that I also used for that campaign
- Increase my advertising spend on different media to those that I used for that campaign
- I wouldn't have spent the money at all (i.e. I would have reduced my overall campaign spend)
- I don't know

If 11a = Increase my advertising spend on other media that I also used for that campaign...

Q11b(i) You said you would increase my advertising spend on other media that you also used for that campaign. What other advertising media would you have been most likely to spend that money on [single response, prompted question] Show media list as at Q9

If 11b(i) = Other radio...

Q11b(ii) You said you would have spent the money you were going to spend on [insert [Global/GMG] stations identified at Q8] on other radio. Which other radio station or stations would you have been most likely to spend the money on? [show list of alternatives depending on area, include don't know]

If 11a = Increase my advertising spend on different media to those that I for that campaign

Q11c(i) You said you would increase my advertising spend on different media to those that you used for that campaign. What other advertising media would you have been most likely to spend that money on [single response, prompted question] Show media list as at Q9

If 11c(i) = Other radio

Q11c(ii) You said you would have spent the money you were going to spend on [insert [Global/GMG] stations identified at Q8] on other radio. Which other radio station or stations would you have been most likely to spend the money on? [show list of alternatives depending on area, include don't know]

Given that question 11 was the key question for the interview, the Survey Moderator focussed on assessing whether it had been implemented correctly – i.e. consistently and in line with the survey script. Following a quality check, the Survey Moderator identified inconsistencies in the way that question 11a had been asked and that one interviewer had undertaken many interviews and was coding the answers differently (as further detailed in section 6 below). The Survey Moderator concluded that “the research conducted is of poor quality and that the Competition Commission should be very cautious about the use of this research data in decision-making”.

I have been asked to review the Survey Moderator's comments on the Existing Customers survey and to provide an independent review of two specific elements of the Survey Moderator's report (as further detailed in section 4 below):

- Did respondents understand the context of question 11 and did the framing of questions or answers impair respondents' ability to answer the question?
- Did an imbalance in interviewer spread introduce bias which affected respondent's media choices?

Moreover, I have been asked to address the critical question as to whether either of these issues undermine the value of the evidence on media alternatives derived from question 11 as a whole – that is, whether question 11 as a whole still allows the Existing Customers survey to identify the “next best

media alternatives” that advertisers would have been most likely to use instead of Global’s and RSL’s radio stations if those stations had not been available when they booked a recent campaign.

3. Scope of report

I have been asked to assess whether the responses to question 11 of the Existing Customers survey accurately identify the next best media alternatives for advertisers that have recently used the parties’ stations for campaigns. In this context, I have been asked to answer the following specific questions::

- Did respondents understand the context of question 11 and did the framing of questions or answers impair respondents’ ability to answer the question?
 - *Was the context of Q11 as a whole clear and were respondents clear as to the precise question they were being asked to respond to – that is, what media alternative they would have been most likely to use instead if the parties’ stations had not been available when they booked their last campaign. Is there any way that the questions as asked could have been misinterpreted by the respondent and if so in what way?*
 - *Did the framing of the question appear to have led to any uncertainty amongst respondents when giving their answers? Do respondents in particular answer clearly to these questions and with confidence?*
 - *Is it reasonable for the survey moderator to have drawn the conclusion that the responses to this question are not robust without listening to the interviews themselves (or a significant sample of interview?). What in your view would be the minimum number of interviews that a moderator would need to listen to in order to draw these conclusions?*
- Did an imbalance in interviewer spread introduce bias which affected respondent’s media choices?
 - *What are the reasons for this coding discrepancy of the main interviewer? Is this a result of the question being asked differently or the answers being coded differently (i.e. a coding error)?*
 - *If the discrepancy was a result of a coding error does recoding the question bring these responses back in line with other respondents?*
 - *What factors explain the shorter length of interviews of the main interviewer and did this appear to have affected the quality of interviews?*

I address each of these questions separately, in section 5 of this report.

4. My approach

I have reviewed the Survey Moderator's report as well as other documentation provided to me by Slaughter and May. However, the core of my review has taken the form of listening to 124 interviews: two CDs were provided to me by YouGov containing recordings of a large proportion of the completed interviews and a spreadsheet containing respondent level data was provided to me by RBB. I was able to use this information to compare the recorded answers with the reported answers. This forms an important part of my comments in section 5.2 below.

Overall, I listened to all or part of 124 interviews. Among the different interviewers, I listened to:

- 102 interviews of "main interviewer 2"; of these I listened to 15 in their entirety and the remaining 87 from question 9 onwards. This 102 included all 85 occasions where she coded 1 and 2 at question 11a
- 3 of "main interviewer 1"; all in their entirety.
- 9 of "main interviewer 3"; all in their entirety
- 10 of other interviewers; all in their entirety

As can be seen from this distribution, I concentrated the available time investigating the specific issues raised by the Survey Moderator's report: the potential impact of main interviewer 2 on the overall survey results.

5. Responses to specific issues

5.1. Did respondents understand the context of Q11 and did the framing of questions or answers impair respondents' ability to answer the question?

Introduction

As noted above, the Survey Moderator identified a number of potential concerns with the way the key "diversion question" – i.e. question 11a – was asked that may have affected the interpretation of responses. In particular the survey moderator noted:

- That the question was not always read with full context;
- That the responses were not always separately prompted.

The Survey Moderator noted *"Inaccurate and inconsistent interviewing is of concern as it means that you cannot be sure precisely how questions are being administered and therefore how to interpret the answers."* As only 3 in 10 interviews had been properly framed (asked in full) and only 2 in 10 interviews had responses read separately the survey moderator concluded that the responses to Q11 as a whole were not reliable – i.e. that they may not reflect the actual media options that respondents would use instead if the parties' stations had not been available for their last campaign.

I have been asked to respond to the Survey Moderator's report under three headings which I discuss in turn below.

Was the context of Q11 as a whole clear and were respondents clear as to the precise question they were being asked to respond to – that is, what media alternative they would have been most likely to use instead if the parties’ stations had not been available when they booked their last campaign. Is there any way that the questions as asked could have been misinterpreted by the respondent and if so in what way?

I note the issues raised by the Survey Moderator and confirm that some interviewers have not followed their instructions accurately. Nevertheless, I have addressed the important issue of whether any such shortcomings had a substantive impact upon respondents’ ability to understand the context of the question, the scenario they are being asked to consider and their ability to provide a realistic answer to the key question of which media is the next best alternative to a Global or RSL station.

The main criticism is that question 11a is was not read out in full in many cases; specifically, some interviewers (mainly main interviewer 2) did not read out the first sentence of question 11a - i.e. “I would now like you to think back to the time when you were planning how to allocate your media budget for the last campaign when you used [insert [Global/GMG] stations identified at Q8]”). I have considered whether this has had a meaningful impact upon the responses obtained. I do not believe that this has been the case. The rationale for including this introductory sentence is clearly to ensure that advertisers are focussed on the recent campaign where they used the parties’ stations and to place themselves at the planning stage of that campaign. I believe that the context of the “last campaign” has already been set out at questions 9 and 10 and the specific premise of not being able to use the particular station is, in the interviews I listened to, always set out. Thus, at very least, the vast majority of respondents will have been asked:

- Q9. For the last campaign when you advertised on [station(s) identified at Q8], which other media, if any, did you also consider using at the time?
- Q10. And what media did you end up using in addition to [station(s) identified at Q8] for that campaign, if any?
- Q11. If you had been unable to use [station(s) identified at Q8] in that period (for example, if there had been no airtime slots available at all), what would you have done with the budget you would have spent on [station(s) identified at Q8]

I have listened to between 100-105 interviews where the first sentence of question 11a has not been read out and, in at least 90% of cases, I do not believe that respondents have had any difficulty in understanding both the context and nature of the question they are being asked. That is, they were clear that they were being asked what they would do instead of using the parties’ stations for the recent campaign where they used the parties’ stations if they had not been available.

In short, based on my experience of listening to 124 interviews, I believe that, in at least 90% of cases, the respondent:

- a) understood the context of the questions (the most recent campaign where they had used a Global or RSL station),

- b) understood the scenario (if that particular radio station had been unavailable), and
- c) understood they were being asked to say what media, if any, they would have used if the Global or RSL station had not been available

Did the framing of the question appear to have led to any uncertainty amongst respondents when giving their answers? Do respondents in particular answer these questions clearly and with confidence?

I believe that at least 90% of respondents have answered question 11 as a whole successfully. Respondents are confident in explaining what they would have done if the radio stations in question were unavailable to them. Clearly different people answer with different degrees of confidence; some are clearly more experienced in selecting advertising media than others. Although the framing of question 11a specifically may have over-complicated the issue with its long options, this has not introduced undue difficulty as many respondents ignore the specific codes offered to them and tell the interviewer what they would have done; i.e. "I would have used another radio station". At this point the interviewer needs to allocate an answer to one or more of the question 11a codes before recording the detailed answer.

Is it reasonable for the survey moderator to have drawn the conclusion that the responses to this question are not robust without listening to the interviews themselves (or a significant sample of interview?). What in your view would be the minimum number of interviews that a moderator would need to listen to in order to draw these conclusions?

I note that the Survey Moderator reached her conclusions on the basis of a single day visit to the offices of YouGov and IndieField where she listened to five interviews and spoke with the fieldwork manager and three interviewers. She also obtained analyses of the following:

- Number of interviews completed by each interviewer
- The average length of interview for each interviewer
- Quality information to identify the extent of adherence to instructions
- Some data analysed by the different interviewers

My conclusions have been drawn following listening to a large number of interviews. As stated above, I have listened to 124 interviews; in part (always including questions 9-12) or in their entirety. I do not believe that there is a specific minimum number, but if the key issue is whether or not the quality of interviewing has impacted upon respondents' ability to understand and answer the questions, which it should be, then a considerably more comprehensive review of the survey than five interviews is required.

Moreover, in undertaking a comprehensive review I believe it is not sufficient to merely listen to the interviews. I have also taken the precaution of listening to the interviews while observing the recorded data from that interview. This enabled me to validate whether the recorded answer is the same as the

reported answer on the spreadsheet. I will describe below a particular inconsistency between what was said and what was recorded. Given the specific comments about main interviewer 2, this was necessary to understand the apparent differences between this interviewer and the others. Was the difference due to the way she asked the question or, as it turned out to be, due to the way she recorded the answers to the question?

Thus, although the Survey Moderator notes that *"Inaccurate and inconsistent interviewing is of concern as it means that you cannot be sure precisely how questions are being administered and therefore how to interpret the answers"*, the opportunity to listen to a large number of recordings means that I can be sure how the questions were administered and can make an informed judgement as to whether the answers can be interpreted reliably. I believe they can.

I believe that my more comprehensive review has enabled me to look beyond the more straightforward issue of whether an interviewer has or has not followed instructions accurately to the more important issue of what impact, if any, a shortfall in quality has had on the results that the survey has obtained.

5.2. Did an imbalance in interviewer spread introduce bias which affected respondent's media choices?

The moderator raised the concern that a small number of interviewers carried out a large number of interviews. The Survey Moderator further noted that a single interviewer carried out over half the interviews and that this interviewer *"asked qn 11 in a different way to all of the other interviewers in that on that qn he/she coded one code significantly more (41% cf 7%)."* I understand the discrepancy is that this interviewer coded more 1 and 2 responses and fewer 1 or 2 responses. Finally the Survey Moderator noted that the interview length for this interviewer was considerably shorter than for other interviewers.

I have been asked to respond to the Survey Moderator's Report under three headings which I discuss in turn below.

What are the reasons for this coding discrepancy? Is this a result of the question being asked differently or the answers being coded differently (i.e. a coding error)?

The interviewer in question has been identified as "main interviewer 2" and the fact that she recorded 41% of her respondents giving codes 1 and 2 at question 11a, compared with 7% of other interviewers was, in my view, a potentially serious criticism of the survey. It was her key (indeed, sole) example of a high level of interviewer bias.

It is worth stating at this stage that, question 11a does not obtain the key information; that of the alternative media that the respondent would have used. Rather, it is a question that introduces the scenario and leads to more detailed questions that obtain the key information. As such the coding of this question does not necessarily impact upon the media alternatives identified by the survey. Nevertheless, if the failure to follow instructions accurately and consistently does have a material impact upon the data obtained or if it reflects actual differences in the responses driven by alternative methods of implementing the survey, then clearly it undermines the work completed by an interviewer

or interviewers. That is, the different results obtained by this interviewer may be an indication of bias, even if this is an introductory question rather than one upon which firm conclusions would need to be based.

However, upon listening to the interviews, it became clear to me very quickly that the reason for the discrepancy in the number of answers of that type was not as a result of a larger proportion of her respondents giving that combination of answers but rather the interviewer incorrectly coding the question. I listened to all 85 interviews where the interviewer had recorded both codes 1 and 2. Of these 85 interviews, I believe that she should have coded the recorded answer as follows:

- Code 1 only – 47 times
- Code 2 only – 29 times
- Codes 1 and 2 – 9 times

In the majority of cases, as I indicated above, the respondent did not answer giving one or other of the options read out to them but told the interviewer what they would have done and the interviewer coded the options available to her as she was best able. As such this reported difference does not necessarily represent a difference in best media alternatives reported by this interviewer but rather a difference in her interpretation of how to code an interim question, before obtaining the key information.

If the discrepancy was a result of a coding error does recoding the question bring these few responses back in line with other respondents?

As a result of listening to these interviews I have re-coded her answers to question 11a where I believe she had inaccurately coded the answers and following my re-coding she does, in fact, have a lower proportion of respondents giving this combination of answers. Below is a table showing the question 11a results before and after my re-coding of main interviewer 2:

	Main Interviewer 1	Main Interviewer 2, before re-code	Main Interviewer 2, after re-code	Main Interviewer 3	Other Interviewers
(1) Increase my advertising spend on the other media that I also used for that campaign	38%	21%	44%	24%	29%
(2) Increase my advertising spend on different media to those that I used for that campaign	16%	7%	21%	15%	26%

(1) and (2) together	7%	41%	4%	6%	7%
(3) I wouldn't have spent the money at all (i.e. I would have reduced my overall campaign spend)	30%	27%	27%	47%	33%
I don't know	3%	2%	2%	1%	2%
Other combinations	6%	2%	2%	7%	2%
<i>Total</i>	<i>69</i>	<i>205</i>	<i>205</i>	<i>72</i>	<i>135</i>

As I mentioned above, although question 11a has been coded incorrectly in a number of cases (most often but not exclusively by main interviewer 2) I have also looked at the other alternative media that have been generated by question 11 as a whole and I am satisfied that the specified additional / different media and stations do accurately represent the alternative media choices given by the respondent. There is no requirement to re-code the media alternatives obtained at subsequent parts of question 11.

Overall, having listened to many interviews I am confident that the main interviewer 2 has captured the next best alternative(s) to the station they used in at least 90% of cases.

As such, I believe that the shortcomings of question 11a are not material in terms of correctly identifying those media.

What factors explain the shorter length of interviews and did this appear to have affected the quality of interviews?

In terms of the length of interview the key difference between the main interviewer 2 and some, but not all, other interviewers is that having read out the list of different media at question 1 and she did not thereafter read out the list at subsequent questions. She offered to do so if the respondent appeared to be unsure but in the majority of cases, she took the unprompted answer(s) given to her at questions 4, 7, 9, 10, 11bi, 11ci. I have noticed that at question 9, where the option "other radio" appears for the first time, she often prompted "did you consider another radio station?" Not reading out the options several times has obviously saved considerable time.

However, as I have explained above, I do not believe that this had a material impact upon the key responses to the survey.

Over the course of listening to interviewers who completed only a small number of interviews, it is apparent that they take more time reading out the lists and the long options towards the end of the

questionnaire. The difference in average lengths of interview comes as no surprise and are, as the Survey Moderator notes, to be expected to some extent.

I have been asked to consider whether the differences in interview length have affected the quality of interviews. On the more substantive point, I do not believe that the differing lengths of interview have materially affected the ability of respondents to answer the key points of the interview.

6. Conclusion

Having listened to 124 interviews, and for the reasons given above, my view is that the vast majority of respondents understood the survey questions sufficiently and that the answers can therefore be used with confidence to infer advertisers' next best alternatives to the parties' stations. In short, I am satisfied that the specified additional / different media and stations do accurately represent the alternative media choices given by the respondent.