

IATROGENIC CONFUSION IN READERSHIP SURVEYS

Claude Heimann, Totum Audits & Surveys, Toronto

Iatrogenic a. *(Of disease) caused by the process of diagnosis or treatment. [f. Gk. iatros physician + -o- + -genic]*

Introduction

This paper deals with the possibility of respondent confusion being caused by the instruments we use to measure readership. In particular, it seeks to determine whether the questions we ask and the aids to recall we provide in our surveys may, in fact, influence the answers we get from respondents in unintended ways.

The conclusions of our research were that such iatrogenic effects may well be caused by the process we choose, and that the results of our research may be affected by them. Because our research was mostly qualitative, we cannot estimate the extent of any effects thus caused. We have also not yet completed the research process, so that we cannot tell whether any proposed solutions will work.

One clear lesson from this series of experiments is that the pretesting of survey instruments is a vital part of the process in audience research. Some qualitative work to determine the respondents' comprehension of the questions asked, and to establish how they use the recall aids, can be helpful in understanding the results we get.

Background

In early 1990 the Canadian Print Measurement Bureau PMB began considering a switch from the Through-The-Book (TTB) method of measuring the audiences of magazines to the Recent Reading (RR) method. The reasons for this suggestion and some of the research associated with it were discussed at the Hong Kong Symposium.

The issue of 'title confusion' was of particular relevance in Canada, for a variety of reasons discussed below. Because of the concern over this potential problem, the Board of Directors of PMB were unwilling to consider a switch unless it could be shown that 'title confusion' either was a less important issue than many members perceived it to be, or that it could be substantially overcome by appropriate choices of the survey materials.

The following will provide some further background to the situation in Canada.

i. **US magazines circulating in Canada**

Winston Churchill, speaking to the United States Congress, said, "Britain and the United States are two countries separated by an ocean and a language." Canada and the United States are separated by neither. We share the longest undefended border in the world and our languages are similar (though the accents may differ). Despite the efforts of the federal government of Canada to protect Canadian cultural industries, this has only been possible in respect to the advertising content of the media.

It is virtually impossible to restrict access by Canadian residents to US television, radio, newspapers and magazines. As a result, many US magazines enjoy circulations in Canada which rival those of even the largest Canadian publications. In one case, as Table 1 shows, an American magazine dwarfs its Canadian counterparts. Since advertising in most US publications is affected by Canadian tax legislation, many of these magazines do not sell space in Canada, and are therefore not willing to pay the costs of measurement by PMB. The potential for 'title confusion' is much greater in Canada than in most other countries as a result.

Table 1 Magazine Circulations in Canada

MAGAZINE	ORIGIN	CIRCULATION '000s
Geographic		
Canadian Geographic	Canada	257
Equinox	Canada	184
National Geographic	USA	781
Shelter		
City & Country Home	Canada	61
Better Homes & Gardens	USA	54
Country Living	USA	61
Ladies Home Journal	USA	98
Parenting		
Today's Parents	Canada	131
Parents Magazine	USA	73
Fashion		
Flare	Canada	213
Toronto Life Fashion	Canada	123
Glamour Magazine	USA	114
Vogue	USA	101

ii. Previous qualitative testing had shown the potential for 'title confusion'

A qualitative study undertaken for PMB in March 1979 by The Creative Research Group found as follows:

"People still tend to react generically to any TV listing, and there is an enormous range of error in responding for these publications. For example, 'TV Times', the one which comes with the Hamilton Spectator, was responded to affirmatively by people who never read the Hamilton Spectator. Only upon debriefing did it become apparent that they were responding to 'any old TV listing'.

The term 'TV Guide' was occasionally used to describe Star Week. It did not seem to matter whether a screening card or a stripped-down issue was in front of the respondent when he or she made such a slip. The point is that to certain readers it simply does not matter what a publication is called, provided it contains some TV listings."

What was particularly interesting in this case was the apparent inability of skeletonized issues to overcome the problem. It was not within the objectives of this qualitative study to examine the reasons for this, but it may be hypothesized that, once all the repetitive material has been removed, so little remains of the TV listing publications, that one is virtually indistinguishable from another. Whatever the reasons, this study presented us with a clear instance of 'title confusion'.

iii. Controlled and newspaper distribution

While not confined to Canada, controlled circulation and distribution as a newspaper supplement are perhaps more prevalent in this country than in most others. In Canada, controlled circulation means free distribution of magazines to entire neighbourhoods, which qualify on the basis of their geo-demographic characteristics. Some of our largest magazines are circulated this way eg. Homemakers Magazine (circulation 1,303,000).

Some of the largest business & financial publications eg. Report on Business Magazine (306,000) and Financial Post Magazine (202,000), as well as several TV listings magazines eg. TV Times (2,080,000) are distributed as newspaper supplements. The latter claims the largest circulation of any Canadian publication. The publishers of these magazines were concerned that, since recipients had not specifically subscribed to or bought the magazine, some readers may not recognize them on the basis of the logo alone. They pointed to

the possibility that reading claims which should have been made for their publications may have been made for some other publication with a similar name or similar content, unless examples of these publications were shown. This was thought to particularly apply to those household members who did not buy the newspaper, or clear the mail on the day the publication arrived,

It was to meet these concerns that a specific recall aid (the CAN card) was devised for use in the Canadian RR tests. Details of this card and the others used in our tests follow.

The Tests

A number of tests were conducted by PMB as part of the decision-making process for their proposed switch to RR. These included studies to determine how to handle irregularly produced publications and which combination of questioning process and screening cards to use, in order to best preserve existing (TTB) relationships between competitive magazines. The latter study was described by Hans Vorster during the Hong Kong symposium, and concluded with the suggestion that a combination of a horizontal, aided questioning sequence using the CAN card be adopted.

This determination was critical to the next part of our programme - the investigation of confusion, since, initially at any rate, we had defined confusion as *'mistakenly believing that title card A represents publication B and that such confusion can lead to inaccurate readership claims'*. Clearly we needed to first decide which title cards were to be used before we could proceed with these tests.

Screening (Title) Cards

Although the recommendation was to use the so-called CAN card, a number of other cards were used at various stages in our testing. These cards are illustrated as appendices to this paper, and are described below.

CAN Card	An 8.5 by 11 inch card showing the logo in the centre surrounded by miniaturized full colour depictions of the previous year's covers and a brief statement about the publishing frequency and methods of distribution.
MRI Card	A 3 by 4 inch black and white card showing the logo and similar frequency and distribution statements.
TTB Card	An 10 by 12.5 inch card with two previous covers pasted onto a black background, and a statement of the publishing frequency and methods of distribution pasted onto the bottom right corner. This card was used by PMB to screen respondents prior to applying the TTB questions.

Test 1 - Small Scale Quantitative

This study called for an evaluation of 40 individual titles, to measure the extent of confusion produced by the CAN card. Respondents were first asked to complete a sort of the cards into one of three piles - Read in Past 3 Months, Read in Past Year, Not Read in Past Year. Having completed the sort, respondents were exposed to a full range of recent copies of actual magazines, both those represented by the CAN cards they sorted as well as by other similar titles from the USA and elsewhere, but which were not measured by PMB. Each respondent was then asked the following questions:

"Now I'd like you to look at this title card again. When you looked at this title card before, which of these publications (POINT TO ACTUAL PUBLICATIONS) came to mind?"

"Earlier, when you placed this particular title card in one of three readership piles, were you, in fact, thinking about (NAME PUBLICATION ON TITLE CARD)? That is, were you specifically thinking about whether you read or looked into any issue of (NAME PUBLICATION AND SHOW ACTUAL PUBLICATION)?"

To avoid fatigue bias, it was felt that no respondent should be questioned about all 40 publications. The study was therefore designed such that all 40 were measured, but each respondent was asked about no more than 20 titles. This was achieved by grouping the publications by category - general interest, newsweeklies, women's, TV, business/finance, shelter and city, and dividing the sample into 5 cells, each of which was questioned about two or three of the categories.

A total of 100 adults 18 years of age or older was interviewed by Canadian Facts at their Consumer Opinion Centres in Toronto. Twenty respondents (10 males and 10 females) were recruited into each of the five cells. This design led to the following number of observations per class:

Table 2 Observations By Publication Class

Publication Class	# of Observations
General Interest	560
Newsweeklies	80
TV	160
Women's	280
Business/Finance	240
City	160
Shelter	120

The results of this study were contained in a number of fairly detailed tables. A summary of these results in relation to the publication respondents claimed to be thinking of when sorting the CAN cards is shown in Table 3.

Table 3 'Confused' Reading Claims

		% of Total Observations By Publication Class Among Those Thinking Of Another Publication When They Claimed To				
			Read In Past 3 Months	Read In Past Year	Did Not Read In Past Year	Total
Publication Class	Observations					
Newsweeklies	80	%	10	10	4	24
Women's	280	%	5	6	25	36
TV	160	%	10	6	23	39
Business	240	%	3	8	34	45
Gen. Interest	560	%	3	4	35	42
Shelter	120	%	8	11	34	53
City	160	%	4	1	32	37
TOTAL	1600	%	5	6	30	41

These results indicated that relatively high levels of confusion were observed among all classes of publications when grouped data were examined. The bulk of this confusion, however, was found among individuals who claimed not to have read in the past year. 10% or less of all observations within each class of publications had some potential for affecting reading claims, because they were made by 'confused' respondents who also claimed to have read within the past 3 months.

In the case of the newsweeklies, much of the confusion was related to Canadian and American editions of one magazine. Since these two editions are identical in most respects (advertising being the major difference), title confusion, as we defined it, is overstated in these results.

Thus, the major concern with confusion was from the TV publications, which we knew existed from the earlier research, and which we had earlier shown to be unaffected by the choice of survey materials. This survey, while useful in confirming the existence of confusion, still left us with some unanswered questions:

- i Since the survey involved only a three month screen, how much lower might the effect have been had we measured it on the issue period?

- ii. What is the source of the confusion, and can something be done about it? Is the confusion generic as in the case of TV listings, or is it induced by the card and questioning sequence we chose?

To answer the first question properly would require a very large sample quantitative study. Since this study had suggested that the level of confusion would be less than 10% in the worst case, the expenditure on such a survey was not deemed to be warranted.

On the other hand, some quantitative work to attempt to answer the second question held the promise of showing ways to reduce some of the confusion, while at the same time providing insights into other aspects of our chosen approach. This was the direction supported by PMB, therefore.

Test 2 - Intensive Interviews

The statement of objectives and the definition of 'title confusion' were somewhat changed for this part of our investigation. The following was the revised wording. Note that the researchers were also invited to look more generally at the chosen technique, and to suggest amendments. This led to some interesting conclusions on the psychology of the interview, which are not discussed in this paper, since they do not fall within its mandate.

"To identify any causes of title confusion caused by PMB's chosen Recent Reading question sequence, wording or visual aids, and to determine whether any modification to the chosen technique would help to minimize this confusion.

For the purpose of this study, 'title confusion' means claiming readership of publication A when publication B was the one that was, in fact, read *in such a way as to affect the average issue audience claims*. Publication B may or may not be a member of the PMB publication set.

While these objectives deal specifically with iatrogenic confusion, ie. confusion introduced by the technique itself, the test may allow us to examine any other aspects of the chosen methodology, and to suggest amendments to reduce potential problems."

The study took the form of a series of intensive interviews with one respondent at a time.. Four sets of these interviews were conducted as follows:

- o 8 pilot interviews
- o 20 interviews using the CAN card
- o 20 interviews using the MRI card
- o 11 'fine tuning' interviews.

In each series, the respondent was taken through the proposed Recent Reading questions, which were somewhat modified between the pilot interviews and the series using the CAN card. Responses to this questionnaire were recorded as if it were a regular PMB interview. After this process had been completed, the interviewer went back over the respondent's answers to check for any evidence of confusion. This was done by exposing the respondent to one or more of the following:

- o A newsstand-type display of recent issues of magazines, similar to those used in the quantitative test. This display was set up in an adjoining room, and had the publications grouped by type eg. womens, newswEEKlies, TV listings etc.
- o The TTB card
- o The MRI card or the CAN card (depending on which was used in the RR sequence)

Where indications of confusion were found, the interviewer was able to refer back to the questionnaire response, to see if the respondent had claimed to have read the publication recently. Irrespective of whether a reading claim may have been affected, however, the interviewer probed for indications of what may have caused the confusion. Respondents were asked what, in their opinion, might have been done differently to help minimize their lack of certainty, and some of the suggestions were carried forward into the final 'fine-tuning' series of interviews.

Pilot Interviews

The first four pilot interviews were conducted by regular PMB interviewers using the standard TTB interview with the qualitative interviewers watching the process. The next four were conducted by the qualitative interviewers following the (RR) procedures described above. These pilot interviews served the following key purposes. They:

- o allowed for detailed observation of the interviewing procedure and process conducted by interviewers who would administer the questionnaires in the real world.
- o helped develop working hypotheses about potential areas of refinement to reduce confusion.
- o identified any refinements needed to the approach before heading into the first formal set of depth interviews.

While these interviews were mostly concerned with 'educating' the depth interviewers and refining the qualitative process, some indications of confusion were already observed at this stage. The most important of these were:

- o The CAN cards failed to show size and texture cues which respondents found helpful in deciding whether or not they had read or looked into some publications during the past year. This was observed on more than one occasion for a controlled distribution women's magazine, which is 'Digest' sized. When shown on the CAN card, the covers appeared the same size as standard magazines, and one respondent, who had claimed not to have read the magazine in the past year, reversed that claim when she saw the actual copy on the rack. She had mistaken this publication for one of larger size, she said.
- o On a number of occasions respondents seemed to rely on recognizing specific covers on the CAN cards to determine both whether they had read or looked into the publication in the past year and with what frequency. This clearly falls foul of the RR model.

As a result of these observations the interviewers were instructed to probe specifically for either of these effects in future interviews. In addition to modifying the interviewing procedures to lessen some of the mechanical problems encountered, the committee authorized the conducting of a series of depth interviews using the MRI card, both as a counterpoint to the CAN card interviews, and to see if use of these cards might help overcome either of these problems.

CAN Card and MRI Card Interviews

The twenty interviews using the CAN card stimuli were conducted using a Recent Reading questionnaire which had been modified to take into account the administrative suggestions revealed in the pilot test. After most of the interviews respondents were shown the newsstand-type display and their responses were probed against the publications appearing thereon. With some respondents, however, the TTB cards, with their two cover paste-ups, were shown instead.

Respondents were closely observed for reference to specific covers, especially when responding to the frequency of reading question. While most respondents were shown a scale to use in answering the frequency question, in some cases it was asked on an open-ended basis. When respondents appeared to be using the covers shown on the CAN card to answer the question, they were probed as to how exactly they were referring to them. When respondents were not obviously using the card to determine their responses, the interviewer asked about the influence of the cover reproductions.

Because the CAN card was so different, the interviews using it seemed to concentrate very heavily on the influence of the recall aid, and other factors appeared to get short shrift. The MRI card interviews were included to offset this, and to concentrate more on the following:

- o the screening, frequency and recency questions
- o the format of the "when (and where) last" question
- o aided and unaided probes on when last

During these interviews, respondents were encouraged to feel free to change their responses if they spontaneously discovered title confusion. In responding to the recency question, they were encouraged to

also think about 'where last'. This was done because, as in many previous studies of the RR approach, respondents were shown to be unclear as to 'when last', unless the exposure was fairly recent. It was hoped that, in thinking about the location of their most recent reading event, respondents might have been able to fix the date more clearly.

Diagnostic Discussion of the CAN Card and MRI Card Interviews

- o Examples of title confusion were observed - mostly among **less frequent and less recent readers** of the publication.
- o In one unique instance the title confusion was a major concern. This was for a financial newspaper, whose weekday and weekend editions were constantly confused with each other, despite the fact that, at the time of the study, one was tabloid size and the other broadsheet. Even the depictions on the CAN card failed to overcome this problem.
- o This absence of size cues was in evidence for other publications, particularly the controlled distribution womens magazine mentioned earlier, and may have led to some (negative) title confusion.
- o There was some evidence of 'similar name' confusion, particularly among television guides, geographic publications, shelter magazines and business publications. This may lead to both over-reporting or under-reporting in a readership study, but the effect appeared not to be caused by different stimuli. The appearance of covers on the CAN card or the TTB card did not appear to reduce this effect, particularly if not all the publications were part of the PMB set eg. Canadian Geographic (in PMB), National Geographic (not in PMB).
- o The array of real life covers on the CAN cards seemed to invite some **issue specific responses**. This again appeared to affect both the recency and the frequency questions.
- o The TTB card, while appearing to invite fewer **issue specific responses**, particularly in reference to frequency of reading, still showed some tendency to under-represent readership where people did not recognize one or both of the covers, even of a publication they actually read with regularity. These cards were also viewed as less professional than were the CAN cards.
- o The MRI cards produced title confusion for the same publications as did the CAN cards, possibly more often, although this could not be accurately measured. Unlike the CAN cards or TTB cards, however, they did not invite any issue specific claims.
- o Cover visuals did seem to aid recall for most respondents, especially for less frequent readers of a publication. This led to a suggestion that a card featuring an unpublished cover be tested. This test remains to be done.
- o The appearance of publishing frequency and distribution information on the recall aid appeared to be helpful for less frequently read publications. In some instances they were overlooked, however, because of the small print size used.
- o Asking respondents to think about where they last read appeared to be helpful in pin-pointing when they last read.
- o A prompt card defining terms such as "anywhere", "your own copy or someone else's" etc. seemed to promote understanding.
- o Specifically mentioning that the respondent may change their responses seemed also to encourage accuracy.

These observations led to a number of 'fine-tuning' recommendations, which were incorporated into the final set of interviews. In particular it was suggested that publications be grouped by topic rather than by publication frequency; that non-PMB publications which might be confused with monitored publications be included, and that respondents be invited to spread each set out to aid side-by-side clarification; that a 'definition of terms' prompt card be used, that the 12-month screen be dropped, if only for the test, and that (unaided) recency be asked before frequency.

In the preamble to the reading questions, respondents were told that we were as interested in what they did not read as we were in what they did read, and they were encouraged to ponder each response and change their replies where necessary for accuracy. They were also told that some publications had been included in the set simply to help avoid confusion, and that not all publications would be included in later questioning. This was done to encourage respondents to consider each publication as carefully as possible.

'Fine Tuning' Interviews

The sequence used in the final set of interviews was as follows:

- o administration of the revised Recent Reading questionnaire using the MRI cards plus some U.S. publications. (It was impossible to group publications if the CAN cards were used, both because of their size, and because back-copies of the U.S. magazines were not available.)
- o partial administration of the frequency, place of reading and other PMB qualitative questions for those publications read in the past 12 months.
- o brief diagnostic exploration of the participants' reactions to the question sequencing and wording.
- o asking about subjective preferences between the MRI and CAN cards.

These interviews showed that the changes we had proposed made little difference to any title confusion which exists. They did confirm, however, that with the exception of the financial paper mentioned earlier, and the television listings publications, most of the title confusion found would be unlikely to influence audience estimates very much. Respondents were clear in their confidence that frequently read or never read publications could be reasonably elicited from any of the proposed stimulus materials.

When, as was the case in these final 10 interviews, participants were provided with an array of title cards covering a category in which they were involved, they appear to quickly scan the publications and sort them (at least mentally) into the following categories:

- o publications which I am familiar with and read some with some frequency
- o publications which I am familiar with, but read infrequently or not at all
- o unfamiliar publications
- o publications I am uncertain about

In other words, they seem to have done an informal filter, based first on familiarity and then on certainty of reading.

On the other hand, when shown an array of title cards in a category in which they had little or no involvement, there was a tendency on the part of some respondents to dismiss the total category, without much scrutiny. This could, of course, lead to under-claiming. Presenting the title cards one by one focused the respondent more clearly on each publication.

The addition of non-PMB publications added little to respondent confidence in reporting on title readership and, at times, complicated the task.

The 'fine-tuning' interviews did not employ a screening question. While this would seem to be ideal, based on the prior evidence that filters screen out legitimate readers, the depth interviewers felt that eliminating the screen made for too little structure, and, in requiring the respondent to address the most taxing ("When last?") question for every publication, might have introduced fatigue. Thus, it is possible that the gains in screening in more readers may be lost in lesser concentration on publications appearing late in the sequence.

Finally, in questioning respondents about their preferences for the different cards, some other factors emerged:

- o Pictures of covers helped respondents feel more comfortable about making claims for less frequently read publications, but only because they were recognized as having been seen before. This appeared to be an indication that using a previously unprinted cover may not have the desired effect.

- o Semantic confusion, caused by common words in magazine titles such as 'home', 'financial', 'TV' and 'business', while it exists, appears not to affect audience levels as much as might be thought. Little can be done to help respondents in this respect.
- o Some magazine logos do seem to be penalized by being shown out of context and in black & white.
- o We again saw examples of underclaiming of some publications because of the absence of size cues.

Conclusions

1. This series of studies confirmed that title confusion may exist. When this occurs, it was shown to be:
 - o mostly caused by semantic factors eg. similar names, formats etc.
 - o a more serious problem among irregular and 'out of issue period' readers, and therefore a less serious problem for measuring audiences than may have been thought
 - o not much affected by the choice of recall aids, particularly in the most frequently encountered instances eg. TV listings publications, for which even showing a skeletonized issue appears not to reduce the confusion
 - o not reduced much by the inclusion of non-measured, but frequently confused publications in the survey set.
2. Iatrogenic confusion (ie. confusion introduced by the technique itself) was also shown to be possible in readership research. In our surveys this was observed in the choice of recall aids as well as in the questioning process. In particular, our inquiry showed:

Recall Aids

- o While cover visuals do appear to aid recall, they also may lead respondents to answer for the specific issue(s) illustrated, despite requests not to. This, of course, violates the RR model.
- o The use of multiple cover illustrations appears to cause some respondents to use them as an aid to their frequency responses.
- o Title confusion may be introduced by the absence of size cues ie. illustrating all covers the same size despite 'real life' differences appeared to mislead some respondents.

Questioning Process

- o When the recall aids for similar publications were presented as a group, respondents who read the class only occasionally appeared to dismiss the group without careful scrutiny, possibly screening out some legitimate reading.
3. On the other hand, our studies indicated that iatrogenic effects could be used to improve readership surveys in the following ways:
 - o Response accuracy seemed to be promoted by stressing to the respondents that 'we are as interested in what you do not read as we are in what you do read', and by specifically giving them 'permission' to change their minds during the readership questioning sequence, especially if a card sort is involved.
 - o Respondents do not always listen intently to interviewer instructions. Involving them more actively in the process appeared to help sharpen their focus.
 - o 'When last' responses in some instances seemed to be made easier for respondents, if they were asked simultaneously to think about where the last exposure occurred.

BIBLIOGRAPHY

- | | |
|--|--|
| Vorster, Hans | "On the Effect of Screening Cards and Questioning Technique
on Readership Scores"
Worldwide Readership Research Symposium V
Hong Kong, 1991 |
| Brown, Michael | Dear Reader
Chapter 7
What's Best to Jog Readers' Memories?
Copyright 1990 |
| Creative Research Group | Report on a Qualitative Study
Prepared for PMB Print Measurement Bureau
March 1979 |
| A.B.C. Audit Bureau
of Circulations | Circulations of US Magazines in Canada |
-