

LEARNING HOW TO ASSESS THE QUALITY OF THE READING CONTACT, AND INTRODUCE OTHER READERSHIP SURVEY IMPROVEMENTS, BY AN INTEGRATED, MULTI-NATIONAL DEVELOPMENT STUDY

Alan Smith, Consultant to FIPP and Survey Research Group

Summary

Advertisers are not interested in information on reading contacts, as such, but rather in the number and type of such contacts which will result in an advertisement being seen. Since it must be accepted, as a matter of common experience, that not all reading contacts are of equal intensity it follows that it is desirable for readership studies to incorporate questions designed to distinguish between contacts in terms of their likelihood of generating ad exposure.

Without doubt, the ideal way to establish how best to meet this requirement would be to study how samples of actual advertisements were treated by different types of readers during the course of actual, natural reading events, but such work would inevitably be expensive. However, such a study could be extended to examine other ways in which the design of readership studies needs to be improved, with such an extension well justifying the cost.

The desirability of making these other improvements has been revealed by fairly recent work which has clearly demonstrated the need for two additional improvements in the planning of print campaigns. These are that it should be possible to take into account the time of exposure, and also the reach and frequency achievements of individual creative treatments. At the same time as introducing such improvements it would also be possible to tackle at least some aspects of the replicated/parallel sources of confusion.

The paper concludes with an outline of the type of study which would be required to enable us to move forward in these directions, and puts forward a proposal for such work to be planned, executed and interpreted on an integrated, multi-national basis. The benefits from such an approach are also described.

SECTION 1 REQUIREMENTS

The Quality of a Reading Event

As far as media is concerned, readership surveys are conducted solely for the purposes of facilitating the creation of a marketplace for print advertising space. However, advertisers are not interested in space, as such, but in the opportunity which that space gives to communicate with those people who are in the market for their product or service. That is, advertisers buy opportunities to communicate, or, to use the jargon phrase with which you may be more familiar, opportunities to see.

The current position is that approximately half the readership surveys around the world include questions which are intended to go beyond a simple measure of a reading contact. The most commonly used question of this type asks how the respondent came into contact with the copy (by subscription, casual purchase, purchase by someone else in the same household, etc.), but other approaches include asking for an estimate of the proportion of pages read, or for the time spent reading.

Until the very recent past all these approaches had one common characteristic, no one knew what they meant. However, this position has been considerably improved by the work of the German AG.MA organisation, as described in a paper by Rolf Speetzen at this Symposium(1). He was kind enough to send me a copy in advance.

Though this AG.MA work has been designed and executed in a very thorough way, I have two basic worries about it. The first one is that they have decided to drop any attempt to assess the incidence of multiple page exposures (MPX), because their preliminary results did not look reliable enough. I believe this to be an unfortunate decision. The ability of certain types of publication, and certain types of reading event, to offer high MPX levels is a genuine strength, which should not be written off too readily. Maybe it will not prove possible to handle the MPX concept within a readership survey, but at this stage I believe strongly that it

should be retained as an objective. My second worry is that by working solely with averages for a publication as a whole, AG.MA is almost certainly under-estimating the level of meaningful print ad exposures provided by many reading events.

Let me explain. In many cases the readers of a publication are very familiar with its make-up, and as a consequence are able to select, for most attention, those sections or pages which are of particular interest to them. This ability of readers to be selective in their use of a publication is one of the main strengths of the print medium, a strength which can be fully exploited by advertisers able to ensure that their advertisements are given appropriate positions. Beer advertisements in the football section, business travel advertisements in the business section, and so on. As a result of this self-selection process any study based on average page traffic levels for the whole of a publication will under-estimate true performance levels. A matter which should be of considerable concern to all publishers.

It is predictable that this consideration will become more important in the future as publications, and especially newspapers, become more and more sectionalised.

This risk of under-estimation means that there are good reasons for the conduct of a development project which takes into account the way in which readers read, and the way in which advertisements get placed within a publication. However, the case for an expensive, multi-national programme of development research has been greatly reinforced recently, by new evidence which has underlined the need for other areas of development.

Lessons from Tracking Print Advertising Campaigns

In December 1992 the research company Millward Brown published a series of papers under the heading 'People, Brands and Advertising'(2). In my view it sets out the most cohesive and persuasive view on how advertising works that has yet been published. It should be required reading for everyone at this Symposium, and even more so, for all involved with the print medium.

The work stems from the efforts of Millward Brown since the mid-1970's to understand the results obtained from the many personal interview campaign tracking studies which they have, and are, carrying out. In the main this work has been concentrated on television, but to an increasing extent Millward Brown have become involved with print campaigns, and it is their two main print findings which are of particular interest to us.

Both of these findings were covered by Gordon Pincott of Millward Brown at the 1991 Symposium in Hong Kong(3), but the evidence is much stronger now. The first one relates to the need to be able to plan a schedule on the basis of the time when a reading event takes place, rather than the date when the issue is published, and the second one is concerned with the evidence for print creative wear-out.

As an aside, may I mention that both these issues were covered in my own contribution to the Hong Kong Symposium(4), dealing with my vision of the ideal readership survey.

The case for taking the time of exposure into account was first fully developed by Neil Shepherd-Smith at the Barcelona Symposium in 1988(5), and is covered again in a contribution to this Symposium by his colleague Dick Dodson, so I do not propose to make further comments on this subject at this stage in my paper. Except to say that I look upon it as an essential scheduling requirement.

Turning to the question of print creative wear-out, the published conclusion in the Millward Brown document(2) is admirably cautious. They write:

"On the available information it would not be prudent to make definitive statements about the nature of wear-out for all brands in all contexts. As a starting point for campaign planning, however, it would be wise to acknowledge the real possibility of an execution wearing out with increased exposures".

And they go on to explain that there was one particular model which seemed to provide a best fit to their data. This was a model which gave the first three OTS behind an execution an efficiency of 100%, but when the fourth (and subsequent) OTS were reached they appeared to make no contribution to advertising awareness whatsoever.

Millward Brown are rightly cautious, but I see no need for such caution. I have believed for many years that print ads experienced substantial wear-out. In the early days I had no strong evidence, but this belief was gradually built-up in step with my growing understanding of the key difference between print and

television, as far as the way in which they are consumed by readers and viewers. Namely time, time of exposure under the control of the reader but not the viewer.

The normal pattern of use of a publication is to turn the page, after an interval of time governed by the level of interest of the reader in the content of the page. If there is nothing of interest the reader passes on in a fraction of a second; but if the content is of interest the reader stays with it, possibly for many minutes. This is equally true of both editorial and advertising material.

In contrast, for television the viewer has no control over the length of the commercial; they can leave the room, or switch their attention elsewhere, but the normal pattern of behaviour is to remain passively viewing.

This key difference between the media is totally in line with the fact that Millward Brown have found no evidence that television advertisements wear-out, in terms of their ability to generate advertising awareness, whereas, as we have seen, print ads seem to experience a drastic fallaway in effectiveness after 3 OTS.

Make no mistake. This finding should have a major impact on the way people think about using print, with the publishers making the running. White space in the media market is their product, and the publishers should be leading the efforts to explore in depth the implications of this new, and revealing, insight into how their product works.

I know it might be thought similar to the conclusion reached by Michael Naples in his review 'Effective Frequency' published in 1979(6), but there is one crucial difference. Millward Brown links OTS levels to creative treatments, but in the case of the Naples study the possibility that effective frequency might be associated with the number of creative treatments was not an issue.

There is a second vital difference between 1979 and today. I think that it is fair to say that the Naples work had relatively little impact on print campaign planning in the years after 1979, but today, thanks to pressure from advertisers for more effective advertising expenditure, the vast flow of scanner data, and the work of researchers like Millward Brown the 1990's will be different. I do not think I exaggerate when I say that in the next 7 years to the year 2000 more will be learnt about effective advertising than in the previous 70 years.

In summary, and putting together these two considerations, time of exposure and print creative wearout, it is clear to me that print schedules need to be planned on the basis of reach and frequency in total, as is currently the case, per convenient time period (such as 4 weeks), and per creative treatment. There is, of course, nothing new in principle about such thinking. It so happens that all TV campaigns are planned in such a way, as a consequence of the nature of the medium.

For good measure, and while we are contemplating the need for a radical re-assessment of readership survey objectives, there is another problem area which needs to be thought about. This is related to the incidence of replicated and parallel reading events and the way in which they effect the readership estimates arising whenever a recent reading(RR) type survey design is employed. This is by far the most commonly used design.

Very briefly, may I remind you that for all RR designs, that is where the basic measurement of readership is the percentage of readers of any issue of a publication within the last publication period (the AIR measure), replicated readership arises for occasional readers whenever reading takes place on two or more days, while parallel reading arises whenever two or more issues are read on the same day.

Replicated reading is the most significant of these two effects, and there is a tendency by researchers to treat it wholly as an error factor, something to be eliminated. In truth, this is wrong. Replicated reading generates significant opportunities for repeat ad exposure and is a major source of strength for some types of publication. It should not be treated solely as an error factor, but, on the other hand, it does result, of course, over-estimation of coverage.

Parallel reading has the opposite effect. It results in an under-estimation of both coverage and frequency.

Hence what is required for good scheduling is to be able to draw a distinction between reading events which generate coverage and those which contribute towards frequency, so that replicated reading in particular is correctly handled, or at least handled in a better way.

These required changes constitute a daunting objective in research design terms, but I believe that they are comprehensive. I cannot think of any other requirements which are just round the corner. So this is not a development without end.

But let us be clear on one thing, right from the beginning. We cannot hope to achieve anything like perfect accuracy in the data we collect. In the very nature of the additional questions we will need to be asking, the accuracy of response might be less than the responses to the questions we are currently asking. But I have no doubt that the end product, the effectiveness of the print schedules planned on the comprehensive data base I have described, will be considerably better than the schedules currently produced.

SECTION 2 MAKING PROGRESS

International Cooperation

It will be clear from all I have said so far that the work I am talking about will have to be on a large-scale, and this means it will be expensive. It is also clear that the need for progress is worldwide, so it would make eminent sense to contemplate planning a development project on a worldwide basis.

I am well aware the cynics will be thinking that it is difficult enough to get agreement on any project within a country, so what are the chances of a multi-country cooperative effort being successful? To which I would respond that the chances are good, because this is such a challenging prospect and because the arguments which support it make so much sense. It is a project in tune with the requirements of the 1990's.

There is also the question of pressure for the harmonization of techniques between countries. Advertisers are beginning to plan across countries, but this is not easy at the moment, with a diversity of readership survey definitions and questions. It will be of benefit generally to print if the introduction of necessary improvements to national readership studies is coordinated, so that cross-country planning is facilitated.

Harmonization will be easier to achieve if everyone seeks to move to a higher level of performance, than if attempts are made to persuade A to move in line with B. Hence the project I visualise is an integrated, multi-national development study, divided into three phases:

Phase 1 Planning. Nothing more than a coming together of those who are prepared to express an interest in the concept, and who are willing to contribute the very small sum of £650 each to fund a planning stage. My aim is to build Phase 1 support at this Symposium, although prior to this meeting I have been engaged in a certain amount of lobbying. As a result of this lobbying I am happy to say that the NRS in the UK, and CESP in France have joined, together with the Survey Research Group, the largest supplier of readership research in the Far East. The SRG have also agreed to undertake whatever piloting might be required, at direct out-costs, so I think we have made a most encouraging beginning.

There is a first meeting planned for Phase 1 sponsors, and any others interested, after this session. The intention at this meeting is to discuss next steps, within a fairly tight time-table. Namely planning and piloting during the remainder of 1993, the conduct of a number of integrated national development studies during the first half of 1994, with interpretation and recommendations by the end of the year. That is, just in time for the next Symposium!

Phase 2 Development studies. I visualise that at least 6 participating countries will be required to conduct a study as outlined below, with each country funding its own work and contributing to the cost of central data processing and interpretation. It is possible, however, that a group of small countries could join together to fund a single study.

These development studies could be phased over time, to facilitate step by step refinement, but this would extend the life of the project, and could complicate interpretation.

Phase 3 Implementation. Hopefully, after completion of the development studies, we will be able to satisfy ourselves that we have devised a survey design which adequately meets the objectives which I have outlined. That is, we will have a design recommendation, developed on an international basis, for the various national readership survey operations to consider.

It is likely that at this stage the national readership survey operations in different countries will fall into two categories. There will be those who have contributed by funding one of the development studies, and those who have not participated.

We are now looking very much into the future, but it seems only fair to me that if the project is successful, those who have not participated in the development work, but who wish to take advantage of the lessons learnt, should pay an annual licence fee for the right to do so. This could be used to create a fund for further development studies.

This might be difficult to enforce, legally, because it is in the very nature of our work that all the design elements of a study are freely published, and I am told that it would be difficult to copyright an improved design, or register it as intellectual property. Equally, however, we are a small community, drawing closer and closer together. It seems to me that moral pressure against the pirating of a design, on which a great deal of co-operative development money has been spent, could be quite persuasive.

Development Studies

The initial planning for the type of development study I have in mind was carried out in 1989, while I was still at IPC Magazines, and in conjunction with Millward Brown. So the concept is not exactly new. Furthermore, I am pleased to be able to tell you that I have recently discussed this project again with Millward Brown, and they would be happy to serve Phase 1 as the lead design consultant.

The outline which follows sets down my initial thinking. It is put forward as the basis for discussion among the Phase 1 sponsors. It deals separately with the four areas in which improvements are required, namely:

- the quality of the reading contact
- the time of exposure
- reach and frequency per creative treatment
- distinguishing between coverage and frequency with respect to replicated and parallel reading

Outline of the Study Design

The plan is that we should work with groups of similar titles, and that each national study should cover 2 or 3 such groupings. The structure for each interview should be similar, with four types of questions being asked. Namely:

- questions intended to identify average issue readers for titles in the publication groups being covered by that particular development study
- questions intended to tell us more about the characteristics of the issue read, the reader, and the reading event, with a selection of such questions being selected for inclusion in the on-going readership study
- questions about a sample of advertisements in the publications which qualified the respondent as an AIR reader, asked only in the development study and purely for calibration purposes
- questions necessary to establish the relationship between the respondent and the target market of the advertisements being used for the calibration analyses.

The success or failure of this project will depend crucially on the extent to which the relationship between the way in which people respond to questions, and the patterns of reading behaviour to which those questions relate, are largely consistent between titles and between countries. In truth, of course, and to a great extent, this assumption has already been made, implicitly, since it is the justification for the widespread use around the world of the RR technique.

However, please note that I am not saying that all publications, in all countries, are read in the same way, but only that the relationship between the responses we obtain to our questions, and behaviour, will be similar. In this, and in all other areas in which we are working, we must be prepared to make trade-off decisions. There is no perfect survey; in all cases we will be faced with the need to choose between alternatives all of which have negatives. And at all decision stages we must keep firmly in mind the long term objective of enabling advertisers, agencies and publishers to maximise the effectiveness of print advertising.

Assessing the Quality of a Reading Contact

The quality of a reading contact would be evaluated by means of questions relating to the treatment given by respondents to a sample of advertisements in issues of publications for which they qualified as AIRs. That is, they happen to have read an issue of the publication within a publication issue period prior to the interview.

These calibration questions would be developed along the lines of the well-established Evaluative Assessment research facility provided by IPC Magazines for many years, and now conducted by Millward

Brown. I have described it elsewhere(7). The EA technique is a refinement of the Starch approach, and here again we have a link with the AG.MA work. But the crucial difference is that the proposed new approach is to analyse the treatment given to individual advertisements, reader by reader, and reading event by reading event, rather than working with overall averages. Only in this new way will it be possible to allow for the relationships which clearly exist between the interests of a reader, and the make-up of a publication.

The EA questions are designed to establish the nature of the communication achieved with the respondent, in a number of different ways, both qualitative and quantitative. The essence of the technique, which I can assure you works well, is to use the advertisement as a prompt and to focus most of the questions on the thoughts which were going through the mind of the respondent at the time they were looking at the advertisement during their natural reading of the publication.

One of the most informative analyses of the EA data is to distinguish between those respondents in and not in the target market. It will be found that on some measures the differences between these groups will be greater than 30%.

If this type of relationship is revealed it would enable planners to start thinking about such tricky, but very relevant questions, as how to compare the value of a prime reading contact with someone on the margin of the target market, with a less good reading contact with someone who is a prime market prospect' Current scheduling practice ignores such trade-off queries.

Time of Exposure and Reach and Frequency per Creative Treatment

These two objectives have been combined because it would be possible to build them into scheduling models using the same type of information. But as far as these objectives are concerned we are really thinking only of magazines, since the reading life of a newspaper, or even colour supplement, is normally short. Clearly it would be quite impractical to contemplate collecting information from a readership survey which related to exposure to specific creative treatments, but if we had information on the age of the issue seen, it would be possible to model the reach and frequency achievements of each creative treatment employed. Given the age of the issue, however, it would also be possible to model how long it took to build-up exposure. So we are lucky in that the same information could be used in two quite separate and important ways.

However, we cannot expect that respondents will be able to tell us the age of the issue they have seen with any high degree of accuracy for old copies, while there are various ways in which such a description as the 'current issue' will be found to be imprecise. This being the case, and on a judgement basis, I suggest that we aim for a three way split of issue age into the current issue, the immediately previous one, and all others.

Distinguishing between Coverage and Frequency

I have already made some comments about the significance of replicated reading, and have suggested that the way round this problem area is to draw a distinction between coverage and frequency, at least in those instances where replicated reading is likely to give rise to a significant over-estimation of coverage. Such as, when an occasional buyer keeps the copy and reads it over an extended time period. This type of reading is likely to be common for certain types of magazines, especially monthlies, in which case it would be a major contributor to the high level of readers per copy for some magazines shown by RR surveys, which some people find so difficult to accept. But I stress that such repeat reading should be counted for frequency.

This is a fairly complex area, but again it is a question of seeking the best balance between meeting the objectives of a readership survey on as broad a front as possible and achieving refinement in any particular area.

The approach requiring least respondent effort would be to reject from all coverage calculations all reading events which took place in the last issue period (that is, they would normally qualify as AIR events), with old copies bought by the respondent. Such events should not contribute towards coverage, since we can be very confident that they would have been read previously, but they should contribute towards frequency.

This approach would require least respondent effort, because it is likely that all information required would be already collected for other purposes. If a further refinement was required one approach, theoretically not 100% correct but fairly respondent friendly, would be to establish how often an issue has been picked-up, and also the time interval between the first and last pick-ups. If this time interval was greater than the issue period, the reading event should be rejected for coverage, but included for frequency.

The Sample

Ideally the sample for each publication group will be large enough to ensure an adequate number of the following different types of reader is available for analysis:

- primary readers; personal buyers and those living in the same household as a personal buyer.
- early pass-on readers; non-primary readers who had read the current or immediately previous copy.
- all others.

Further calculations are required to determine how big the samples should be, but it is likely that at least 600 reader/issue contacts will be required for each publication group. The number of people required to be interviewed in the first place in order to identify such reader/issue contacts will, of course, depend on the AIR readership levels achieved by the publications in the publication group being covered by the study.

Since it will also be necessary for the interviewers to carry a number of issues of each title being researched it can be well

appreciated that the cost will be high.

Pilot Test

As a development project there are various aspect of this outline design which need to be considered further in the light of actual interviewing experience. It would be essential, therefore, for one, or maybe more, pilot tests to be conducted. This work could be carried out in a single country.

References

1. APX - The New Deutschmark in Print Media Buying. Rolf Speetzen and Jurgen Wiegand. Readership Research Symposium. San Francisco 1993
2. People, Brands and Advertising. Seminar papers published by Millward Brown International. December 1992
3. Investigating Readership Lag. Gordon Pincott. Readership Research Symposium. Hong Kong 1991
4. The Ideal Survey. Alan Smith. Readership Research Symposium. Hong Kong 1991
5. Taking Account of the Time factor. Neil Shepherd-Smith. Readership Research Symposium. Barcelona 1988
6. Effective Frequency. Michael Naples. The Association of National Advertisers, New York 1979
7. Evaluating Magazine Advertisements. Alan Smith and Mark Hodson, Admap October 1987

