

AIM, ADVERTISING IMPACT MEASURE

Wim van der Noort, Interivew-NSS
Brian Shields, International Herald Tribune

"You mean, your statistics are facts, but my facts are just statistics."

Jonathan Lynn, Anthony Jay: script of *Yes Prime Minister* (1986)

Summary

Starting in 1999, The International Herald Tribune launched AIM, the Advertising Impact Measure. AIM stands for a series of studies measuring the *quality of reading* of international newspapers and magazines. The measures are linked to the major international surveys measuring print audiences for the pan-European market, i.e. the European Media & Marketing Survey (EMS) and the European Businessmen Readership Survey (EBRS) and also to the Asian Businessmen Readership Survey (ABRS) for the Far East & Pacific region.

To go beyond the currency of Average Issue Readership, enhancing readership research with additional measures reflecting (differences in) opportunities to see *advertisements*, is a much-debated topic in all Readership Symposia so far. Despite the considerable amount of effort and money that has been spent in developing and validating such measures the emergence of an accepted standard seems even less likely than it is for the AIR currency.

Nevertheless the demand from advertisers and agencies for more relevant print measures for media planning and buying will not diminish, since evidently print is lagging far behind in this respect compared to TV, Radio and Internet.

In this paper we present the approach taken in the AIM studies, providing media planners with a (in our view) practical and sound tool for advanced planning. The topics to be discussed are the general design and philosophy of the AIM studies, the technical measurement issues involved and last but not least the acceptance in the market.

Introduction

Measuring readership, it might be assumed, presupposes knowledge of what a reader is. After at least half a century of serious readership research designed for the commercial benefit of those planning, buying and selling print advertising, all evidence would suggest we are not yet near to coming to agreement on even this most basic notion. "Read or looked at, it doesn't matter where or when or for how long" or "Read or looked at for at least two minutes" or "Read" without further qualification, are still all oft-used mechanisms. Different criteria for readership, one supposes, bring forth different numbers and types of readers.

For advertisers, it might be taken that "readers" (however defined) is a term used to differentiate between those with a real potential of seeing their advertisement in a given title, compared with those with no chance of the same. Michael Brown, in his excellent "Effective Print Media Measurement" points out that readership scores should be taken as meaningful largely on a relative scale, and that absolute figures, though they are always expressed in these, are of dubious value. That is, readership of Publication A only makes sense, or has meaning, in relation to the score attributed to Publication B. Neither has any real status independent of the other¹.

So if we assume that advertisers and their agencies knowingly or unknowingly work on these assumptions, we are left with a set of publications with relatively - meaning both reasonably and in respect of each other - reliable readership scores. However, their concerns are not about "readers" per se, but with these readers in relation to their advertising messages. All work on the assumption that some advertisements are, independent of content and creative make-up, more likely to be seen, and hence have a chance to work, than others. All sides in the advertising business accept this: media owners have premiums and differential rates for different positions; agencies and their clients still clamour after first right-hand facing editorial matter or front of book (except in Chinese publications where left-hand "end" of book takes priority).

¹ As buying currency the 'relativity' argument holds true, however, the reduction of the importance of absolute measures is seriously compromised when total campaign spend is set against the schedule analysis measures of coverage and gross rating points if, as the case may be, absolute values are considerably overstated.

However, recent (meaning both that produced in the last twenty years and employing the recency model) readership research has, with a few exceptions² (and in the international arena, no exceptions), overlooked the differences in the way readers read and publications are read³. Rather the debate is often obviated by application, at the analysis stage, of the term 'Opportunities To See' (OTS). Here the difficulties and realities of differential exposure to advertising are both acknowledged and avoided. That is, all readers are deemed to be equal and all have an opportunity to see of 1. Most readership surveys go no further, certainly not in respect of producing data that, numerically, can sensibly qualify the measurement of "readers"⁴.

But for all **real** OTS to be equal to 1, however, the following (minimum) conditions have to pertain.

- 1) All pages of all titles are looked at or read from
- 2) All pages in all titles are looked at only once

However, these assumptions contain problematic elements. In respect of the first, you can't see an advertisement on an unseen page, and it is unrealistic to suppose that all pages are seen in all titles by all readers. For the second condition, pages viewed twice or more carry an opportunity to see equal to the number of separate page exposures. Thirdly, it might be added that there is an assumption also about the equality of intensity with which each page is read - an advertisement on a page that is read comprehensively would probably, in reality, be regarded as better placed than one that is turned-over with an immediate flick of the wrist, coupled with an uncurious eye. We will refer to this dimension later as quality of page exposure.

It is as if, through the later to be applied mechanism of OTS, readership research has given itself leave to treat these issues as "subjective" or "qualitative" matters, that, if at all, are to be dealt with by non-empirical techniques. At best these are taken to be of secondary importance; in reality, in methodology and questionnaire construction, they are often taken to be of no importance.

Returning to the position of the advertiser - without whom there would be no 10th Worldwide Readership Symposium: is the extent of their concern really confined to the ubiquitous "reader", with investigation thereafter halted by the meaningless, at best, misleading at worst, acronym OTS?

AIM, Advertising Impact Measure, takes as its starting point the premise that measures can reasonably be adopted which provide better answers to likely **real** opportunities to see than currently given by the expression (all) $OTS = 1$. We know that if only half of a publication is seen - that is, half its pages are read from or looked at - that the **real** average OTS of advertisements appearing in the title is less than 1. Also, if certain pages are seen on more than one separate occasion, then the OTS of these pages is greater than 1. And measures have been devised to investigate the quality of page exposures.

AIM therefore works on the basis that the "average" or general relationship between a reader and a publication can be penetrated by a set of brief, transparent and easily answerable questions. The use of the word average is not accidental. We are dealing with *average* behaviour in relation to the nature and quality of the reading experience. This is congruent with the readership measure itself, which is often expressed as "average issue". [In other studies where 'regular' readership is the norm, the AIM technique could be equally applied, but obviously only to those who comply with that criterion]. The AIM measures report on individuals' readership behaviour relating to the last issue of a given title that they read, and the score given for each title is a weighted average, by readership frequency, for those representative of average issue readership.

The objects of measurement are:

- 1) Proportion of publication read
- 2) Average number of times that each page that has been seen has been read or looked at
- 3) Quality of page exposure, that is the relationship between time spent reading and pages (in absolute terms) read

² IPA, ISBA and PPA "Quality of Reading Survey" conducted by RSL being a notable exception. There remain problems, however, with the application of the results. Questions following Guy Consterdine, Lynne Robinson and Hilary Birt's paper (The Quality of Reading Survey – An Instrument for change) at the Florence Symposium revealed that when the number of Page Exposures for some (monthly) magazines is related to the total time spent reading for these titles, the resultant time per page measure seriously calls into question the nature, meaning and value of the Page Exposure itself.

³ The 'Through the Book' method sought to go further than establish exposure to a publication in a given period. For a quick overview, stripped of its technical considerations, of the TTB versus Recency debate, Harry Henry's personal memoir "Readership Research" is a good starting place. This traces lines of argument from the early Hulton Readership Survey (1947) days, through to the late Timothy Joyce's paper at the Barcelona Symposium. Neil Shepard-Smith in his paper "The Ideal Readership Survey" produces an excellent treatment of the subject in considerably more depth and tackles the question of 'readership' and meaning of 'impacts' and 'OTS' in a fashion that will not disappoint.

⁴ Harry Henry (ibid.) correctly identifies the crucial question as being, "you may be able to tell me how many people on average will read or look at a copy of the publication in which my advertisement appears, but can you tell me what is the likelihood that they will read, or even glance at, the advertisement itself?"

Research design

Quality of reading data for large scale readership surveys in general can only be gathered by means of separate surveys, since the questionnaires for readership surveys are already lengthy enough collecting the basic readership data needed. The research model of AIM, as a separate follow-up survey, has now been applied to three major international media surveys:

1. AIM-EMS
linked to the European Media & Marketing Survey
2. AIM Business Europe
linked to European Businessmen Readership Survey⁵
3. AIM Asia
linked to Asian Businessmen Readership Survey⁶

In the European Media and Marketing Survey (EMS) only 'proportion read' is collected as additional qualitative measure for international print titles. In the European and Asian Businessman Readership Survey no qualitative measures are involved at all.

Two ways are open for collecting additional quality of reading data:

1. *Single source approach* A single source approach can be accomplished by re-contacting respondents from the original readership survey and ask them the additional quality of reading questions. The advantages of the single source approach are obvious. First it enables us to directly identify readers of individual publications from the existing database. This makes fieldwork very efficient, since also for low penetration publications, readers can easily be found, assuring sufficient sample for each different title. Second, the sample is drawn from a known universe, so potential biases from non-response can be easily re-weighted and third an extensive data record containing media and target group information is already available for each respondent, reducing the questionnaire length in the re-contact survey to a minimum. This approach was adopted for AIM-EMS.
2. *Separate samples* A second approach is to conduct a fully separate survey among a fresh sample, using the same universe definitions and selection criteria as in the parent-survey. This approach was taken with regard to AIM Business Europe and AIM Asia, since the survey procedures for the European and Asian Businessmen Readership Surveys don't allow re-contacts to be made with the respondents. This method is less efficient in terms of sampling procedure and fieldwork, but is in itself equally valid, provided that the readers selected for the AIM survey can be considered as equivalent representatives of the target group of the original survey, in this case executives in leading companies. The AIM survey in these cases is carried out using the same sample sources as in ABRS and EBRS and selecting the same company and job functions. A refined sample stratification and weighting scheme is used to ensure that the distribution of the AIM sample over countries, company types and job functions reflects the ABRS and EBRS universes.

Interviewing method

The interviews are conducted using Computer Assisted Telephone Interviewing (CATI). For AIM-EMS (the first AIM-study) this was a rather obvious choice, since EMS itself is done by means of CATI. The same research model then was applied to AIM Business Europe and AIM Asia, first of all for reasons of comparability. Moreover, a telephone survey is practically the only way to perform the extensive screening needed to find respondents with eligible job functions according to the universe definitions and being also recent readers of the publications to be measured.

The fieldwork in 16 European countries for AIM EMS and AIM Business Europe has been conducted by INTERVIEW•NSS from the European Data Collection Centre in Amsterdam by native speaking interviewers. The fieldwork in the Asian countries for AIM Asia was conducted by Taylor Nelson Sofres from their call centre in Hong Kong either in the local language or in English. Most of the interviewers are bi-lingual and the CATI software could switch between languages by one stroke on the keyboard if necessary.

Selection and definition of readers

The purpose of the study is to obtain indicators for the quality of reading that are representative for the average issue audience of a publication. The measurement involves a rather detailed specification of the reading behaviour for the last issue read. Therefore it is only meaningful to include readers in the survey for whom this last reading event was not too long ago. So only *recent readers* are selected for the AIM surveys, defined as:

- having read the publication in the last 7 days for dailies
- having read the publication in the last 30 days for weeklies and fortnightlies

For representing the average issue audience, the over-representation of infrequent readers is re-weighted by applying the theoretical reading probability derived from the reading frequency.

⁵ IPSOS-RSL

⁶ IPSOS-RSL

Sample

In the table below the publications measured and the sample sizes for each of the studies are given.

TABLE 1 Sample sizes

	AIM-EMS 1999-2000	AIM Business Europe 2000-2001	AIM Asia 2000-2001
Dailies	n=	n=	n=
Financial Times	235	150	110
International Herald Tribune	251	149	115
Wall Street Journal Europe	135	122	
The Asian Wall Street Journal	-	-	130
Weeklies			
Asiaweek	-	-	176
Business Week	212	149	128
The Economist	239	150	127
Newsweek	253	150	126
Time	250	150	136
Far Eastern Economic Review	-	-	125
Fortnightlies			
Fortune	170	118	128
Forbes Global	-	44	11
Total	1745	1182	1312

Weighting and ascription

Each survey can be considered as consisting of separate representative samples of recent readers of the publications involved. In case a respondent is questioned about two publications, he also belongs to two different samples.

All results are weighted and projected in accordance with the parent-survey. The universe estimates for the total audience of recent readers (i.e. readers last week or last month) are calculated, using the cumulated reach after 6 insertions for dailies, 4 insertions for weeklies and 2 insertions for fortnightlies.

Re-weighting of the sample is done separate for each publication, based on the distribution in the parent-survey for the total audience of recent readers of the following variables:

Country
Age
Sex
Business sector

To reflect the average issue audience, results were additionally weighted using the theoretical reading probabilities derived from the reading frequency question.

The results of the AIM-EMS are also ascribed to the full database of the European Media and Marketing Survey, making it possible to analyse the AIM impact measures for a host of different target groups.

Questionnaire

The questionnaire starts with the screening questions establishing for which publications the respondent qualifies as a recent reader. The quality-of-reading questions are asked of a maximum of two titles per respondent to avoid a too-lengthy questionnaire being administered, which might affect the quality of the response and the interview. In cases where a respondent qualified for three or more titles, a selection of 2 titles was made, giving priority to publications with a low penetration for which it is harder to find readers.

The questionnaire was on average less than 10 minutes, which enhanced a high response of nearly 60%.

The AIM study conceptualises quality of reading in two basic dimensions:

1. The actual *reading behaviour* of the last issue in terms of page exposures and reading intensity, summarised in an overall AIM factor per publication.
2. *Attitudes* to the publication, giving insight into the reader/publication relationship, tapping on differences in functionality and Umfeld between titles.

Measuring page exposures

The following behavioural impact measures can be derived from the AIM survey:

1. Proportion of pages read
2. Time spent reading
3. Average number of reading occasions ('pick ups')
4. Multiple page exposures
5. AIM factor

The AIM factor is a calculation bringing together page exposures and reading intensity per page read. It is a multiple of total page exposures, exposures per page read and reading intensity per page read, which will be explained further on in this paper.

First we will go a bit deeper into the question of how multiple page exposures are measured in this survey. In the Quality of Reading Survey (QRS, 1998) undertaken for the UK press market, the PEX score was introduced as a successor to the 1986 MPX study. The measurement of PEX is based on the percentage of pages read on the most recent reading day multiplied by the number of reading days at any issues in the last publishing interval. This approach is theoretically correct, assuming that the last reading day is a randomly assigned day and therefore an unbiased indicator for the reading behaviour on the other days of the publishing interval. In practice this approach might easily lead to overclaiming, for respondents tend to project their best reading behaviour in the issue on the only day they are questioned about. Beside this, as demonstrated by Ron Carpenter (WRRS, Florence, 1999) there is an apparent over-estimation of the total number of reading days in the last interval.

The measurement of page exposures in the AIM study differs in the following ways from PEX. Instead of the number of reading days, the number of separate reading occasions ('pick ups') is asked for, making the concept of multiple page exposures also applicable for daily newspapers. Second the number of reading occasions is related to a particular issue (i.e. the last issue read) instead of any issue in the last interval. And last, the proportion of pages read is measured for each of the claimed reading occasion. For sustaining the memory of the respondent, he had to specify for each reading occasion first when and where it was, before estimating the proportion of pages read.

This measurement procedure resulted in realistic estimates of the number of page exposures.

The multiple page exposure factor published in AIM indicates on how many different occasions a read page is seen by the reader and is calculated in the following way: summing up for each reader the proportion read on each reading occasion (brut proportion read) and then dividing the brut proportion read by the net overall proportion read.

Attitudes

Respondents were asked to indicate their level of agreement (on a scale ranging from strongly agree, agree, nor agree nor disagree, disagree strongly disagree for the titles for which they had given quality of reading information. These statements were

- (*title*) is a publication I would miss if it were no longer available
- I read (*title*) for selective specialist information that I find useful for my work
- I value the broad international perspective on a variety of topics that I get from (*title*).
- (*title*) provides me with a different perspective and viewpoint that I don't get from my national media
- (*title*) provides me with a general international perspective and overview that I find useful for my work
- It is important to me as a person to be broadly informed on international events and their repercussions and (*title*) gives me this
- (*title*) gives me insights into global affairs and events that are important for my personal and business life
- In conversations I sometimes find myself quoting from or referring to information or news I read in (*title*)

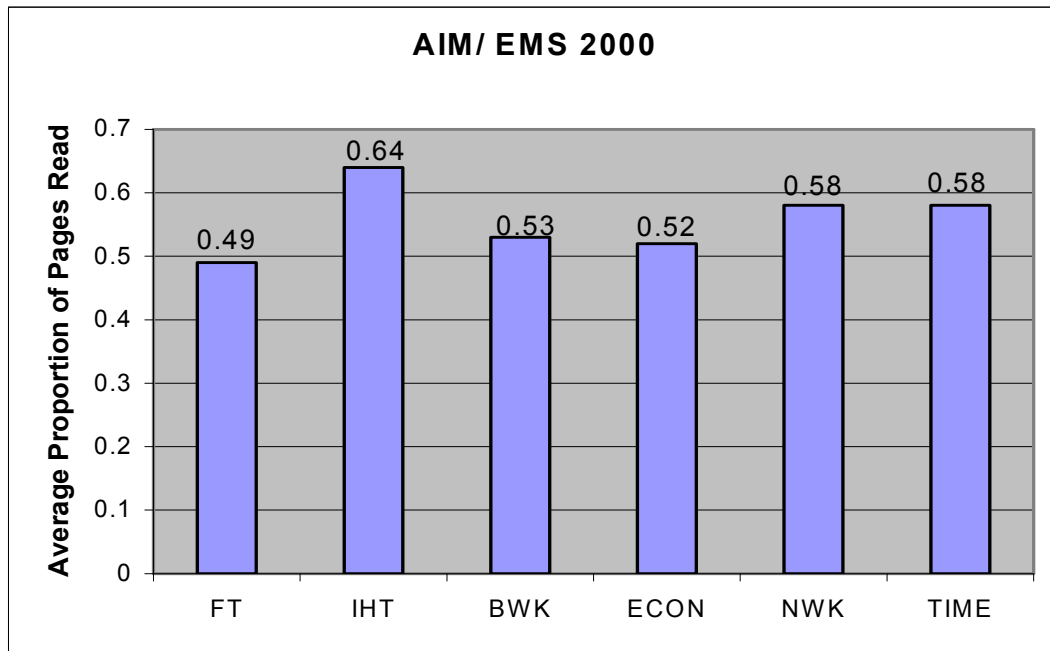
Findings

Having conducted the research and obtained the results, what inferences, in a general sense, can be drawn, and how can the findings be applied?

First, that there exists a sizeable difference between the highest and lowest scoring publications on each of the variables measured, proves how moribund the whole concept of an "opportunity to see" is. Below are figures for proportion read for AIM related to EMS. The same is equally true for average exposure per page read and for time spent reading each page that is read.

The advantage, where merit is attached to the highest scores, does not always lie with the same titles, but clearly there is a range of performance, and not a singularity, as OTS would suggest. See Table 2.

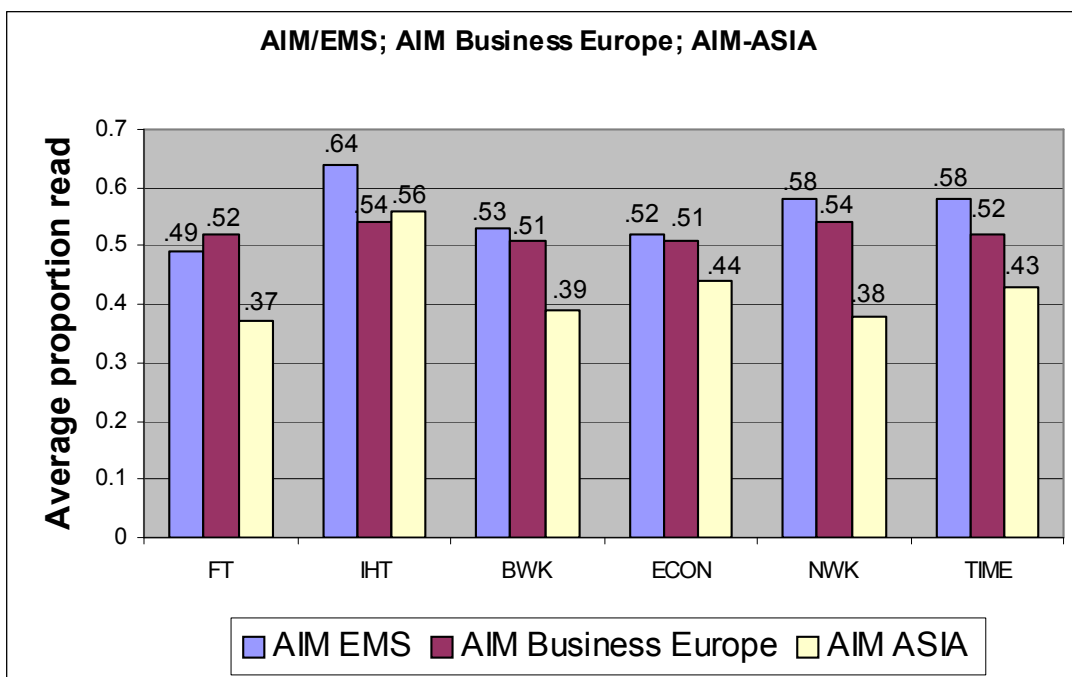
TABLE 2



Secondly, publications do seem to exhibit similar generalised traits, and hence strengths and weaknesses over a range of different universes, and indeed continents.

Figures are given here for five titles on the parameter of proportion of publication read for AIM linked to EMS, and AIM conducted on the same universe definitions as EBRS and ABRS. See Table3.

TABLE 3



What is perhaps of greatest interest is not the numerical value of the proportion read score for any title across the three separate surveys, but rather the relative position of each of the titles against the other four across the studies. The relative position of the publications remains consistent. One again is reminded about Michael Brown's rejoinder about scores being valuable in the relative sense only.

However, these data were not produced simply to disprove the paucity of the concept and numeration of OTS. They were intended to have a function in the planning process. Clearly each of the three main dimensions measured can be used independently of each other and can be set alongside the readership measures from each of the surveys they are meant to be reflective of. There was pressure though to go further. To produce a figure for each title that combined the effect of the three principal dimensions - proportion of publication read, average exposure per page read, quality of page exposure.

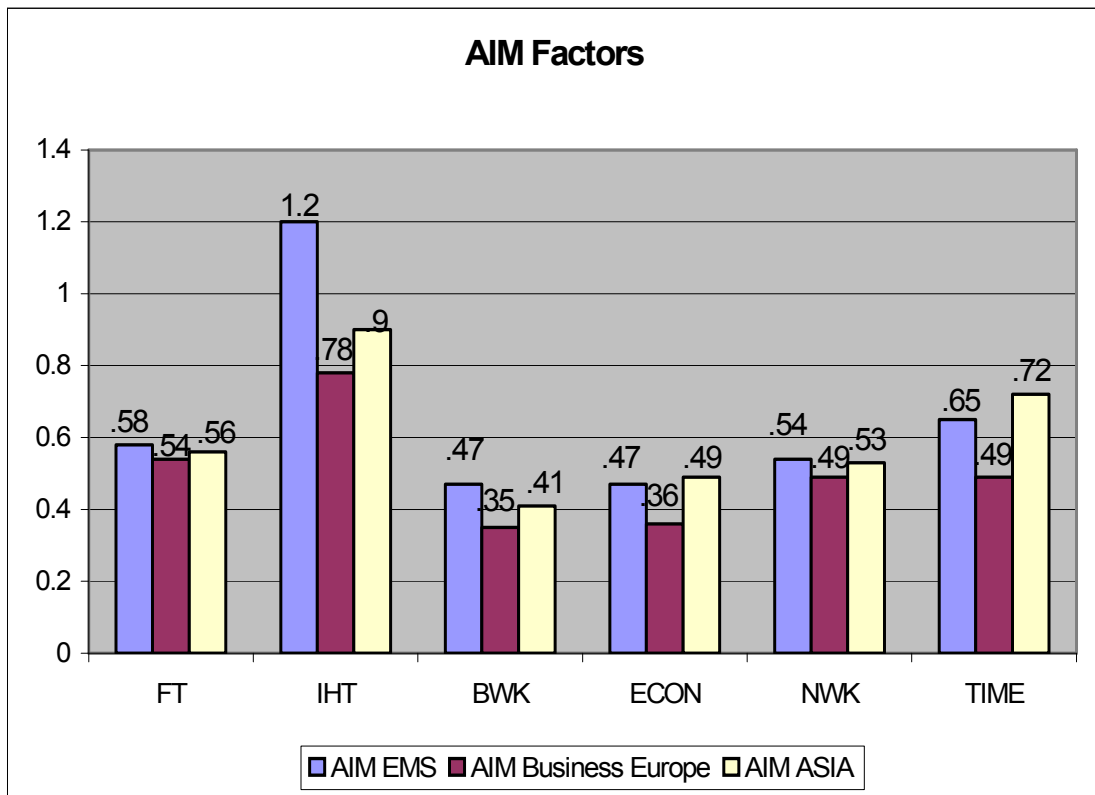
The production of a single relative factor was necessarily an arbitrary business. It is not something that is "out there" in reality, like the height of a building or the speed of light. This is something more akin to IQ, for example. It is an artificial construct, though perfectly transparent.

The following was arrived at:

$$\text{AIM Factor} = (\text{Proportion of pages read}) \times (\text{Average exposure per page read}) \times \sqrt{\frac{\text{Time spent reading}}{\text{Average pages read}}}$$

It should be noted that proportion of pages read cannot exceed 1, and average exposure per pages read cannot be less than 1. The range of scores for the quality of exposure variable has a great deal more variance than the other two. To leave it unmodified would render it the largest influence in the overall factor score. Hence the reason to smooth the variation by application of the square root. Clearly there are other ways to produce a single factor. A planner, for example, may not wish to accord relative equality between the three dimensions and the calculation can be adjusted accordingly. However, whatever the final estimation of an AIM Factor it will, like the three dimensions taken individually, have the merit of differentiating between titles and producing scores across the three surveys that, relatively, exhibit similar patterns. See Table 4.

TABLE 4



Support from the market

It is doubtful if the AIM series would have gone off the ground if it were solely conceived and supported by one publishing house. The survey has attracted sponsorship and financial backing from ten leading international advertising agencies - Carat International, CIA, Doremus, Initiative Media, MediaCom, Media Edge, Mindshare, MPG, Optimedia, TMP Europe and Universal McCann. The AIM surveys have not been a methodological construct to undermine one element within the media planning and assessment process (OTS) that is employed to generate the final gross rating points figure for any given schedule. The data are used to produce - in the eyes of the user - give better considerations of an individual medium and by extrapolation, a better understanding of the collective effects of chosen media in a given schedule.

For example, AIM, as linked to EMS, is fully enmeshed within the whole survey database. Scores on any of the AIM dimensions can be produced across a range of target markets for the measured international dailies, weeklies and fortnightlies, as well as being integrated into schedule analyses

So far AIM data have been collected at re-contact or reconstruction stages. For many readership surveys there is no reason that information cannot be obtained (if for a limited number of titles initially, and for the international English language media landscape this would suffice) within the survey proper.

Why has this not been done or is not really even now been given due consideration?

It is technically possible; and, international media agencies and their clients are supportive of the initiative. The resistance comes from some media owners, ones whose best interests, as they perceive them, are not best suited to the further investigation of an opportunity to see. Any debate generated by AIM - even if it focuses on how best the most pertinent variables that affect likely exposure and impact of advertising might be measured - may serve to move forward the substantive issue from a position of *should* we be doing such research to *how* best to do it.