Balnaves, Mark, Tom O'Regan, and Ben Goldsmith. "The Ratings Provider." *Rating the Audience: The Business of Media*. London: Bloomsbury Academic, 2011. 123–144. *Bloomsbury Collections*. Web. 26 Apr. 2023. <<http://dx.doi.org/10.5040/9781849664622.ch-006>>.

Downloaded from Bloomsbury Collections, www.bloomsburycollections.com, 26 April 2023, 14:07 UTC.

Copyright © Mark Balnaves, Tom O'Regan and Ben Goldsmith 2011. You may share this work for non-commercial purposes only, provided you give attribution to the copyright holder and the publisher, and provide a link to the Creative Commons licence.

6 The Ratings Provider

Most sampling frames in China are based on very outdated census figures which, even then, take account only of demographic and geographic factors, and not information on set ownership or viewing behaviour ... [The] most serious problem is one peculiar to China. It is standard practice to offer gifts and incentives to panellists taking part in the ratings surveys. Many of these gifts come complete with TV station logos. For whatever reasons, it is highly probable that many panellists feel they should return the favour by recording high viewing to the sponsoring channel. Many may feel, especially with Party control of the media, that it is their 'duty' to watch certain programmes such as the news. Andrew Green (Green 1996)

Andrew Green of Saatchi & Saatchi is here bemoaning the fact that there were in 1996 political and market difficulties in setting up ratings in China, not least the lack of a proper census from which to create an establishment survey, but of course also distortion. As we will see, his company was not passive on this issue but put proposals to China for change and these proposals have for the most part come to fruition. Green was concerned that because of the absence of a ratings provider as the point of coordination there had as a result been a massive duplication and multiplication of measurement services, all going their own way. As a stakeholder in China, Saatchi & Saatchi was not going to rely on many measurements that provided no basis for genuine currency. This might sound ironic, given the authors' last chapter where a ratings provider, like Nielsen, is incumbent, but everyone has still gone their own way. Green's argument, though, is sound from the point of view of ratings as a convention.

For all of the seeming apparent attention given to the ratings provider in news stories and industry debate over the ratings, surprisingly little attention has been given to the ratings provider, whether as an independent owner or in joint-industry collaboration, as the key coordinator of the ratings convention. The ratings provider fits into the convention as the nodal point around which knowledge accrues and the survey operates tactically. The provider is the knowledge broker – it is the intermediary which makes the ratings as a coordination rule a reality. The companies do the data collection and they provide information from this collection; they are the entity being evaluated by formal and informal auditing. This ensures that they are a

ready target for criticism. In this chapter we want to shift the focus on to the ratings provider and the kinds of politics, knowledge, behaviours, tactics and strategies appropriate to their carriage of the ratings.

The Official Truth

Saatchi & Saatchi, one of the biggest spenders in the Chinese advertising market in 1996, had a problem. It did not trust audience research in China or its evolving ratings system. Television in China is organized around the five complex political layers of the Chinese bureaucratic system: the national level (*zhongyang*); 30 provinces (*sheng*); 2,148 counties (*xian*); 622 cities (*shi*); and 'work units' or *danwei*. Each level has a Communist Party and a separate government structure. There were in 1996 estimated to be 330 million households of which 83.4 per cent had television reception. The Propaganda Department of the Chinese Communist Party (CCP) controls what is shown. The Ministry of Radio, Film and Television is represented at national, provincial and city levels. China, for Saatchi & Saatchi, was complex in political and market terms and its concerns about hypoing were matched by its concerns about a proper establishment survey to represent television and radio households.

As a result, Saatchi & Saatchi, on behalf of the Hong Kong Association of Accredited Advertising Agencies (4As), presented to all the major TV stations in China carrying out their own surveys five major recommendations on reforming China's ratings system:

- 1 The service should be national, covering all major provinces and cities. Each survey currently in existence involves different companies in different cities using different techniques;
- 2 The service should be continuous. All the current advertising agency surveys suffer from the weakness that they measure only one point in time that is not representative of other periods of the year;
- 3 The service should be led by research buyers rather than research sellers, with the television stations as the primary partners and founders, as in most western markets;
- 4 There should be only one ratings system in China. The alternative, as has been seen, is a mess;
- 5 The service should be reliable, independent and transparent in its operation.

What we can see in China mirrors, as the authors have tried to show, the early history of audience ratings in the West. In those countries which established

commercial broadcasting media early, such as Australia, the United States and Canada, audience measurement started with customized research undertakings of advertising agencies and stations. Fan mail is a good example of this type of customized research. It could be used to assess the popularity of programmes. The trajectory from customized to syndicated research occurred because a collective understanding of audiences in a growing broadcast market was needed. Without it no systematic trade was possible. A ratings service – the convention of standardized measurement – *cannot be imposed*. The need for it has to grow out of the experience, needs, understandings and interests of the various media players, and out of their common recognition that consensus on a convention provides stability and certainty in the market. As a corollary, a sufficiently developed, competitive and differentiated media market must be in place to require competitors to become collaborators for the purpose of supporting an independent media broking service.

The idea of an independent, industry-wide audience measurement service took some time to emerge in the United States and elsewhere, and it developed at different speeds. The United Kingdom could have adopted ratings for radio but decided against it and only did so when the market became competitive. Because audience measurement began as customized research by advertising agencies and radio stations this research was often limited by the interests of the commissioning service. Customized audience research was currency, as it still is in markets without a convention, and advertising agencies in these circumstances are well placed to function as ‘brokers’ and coordinators of a fragmented market place. In these circumstances customized research is ‘*the only way for advertisers to learn about the audiences for media outlets*’ (Miller 1994: 63). This is the case for African countries and, indeed, in advanced industrial states where the panel sizes for ratings may be too low. As a result, advertising agencies and the individual station face a credibility problem with advertisers and media outlets. Each has a vested interest in the research. Advertising agencies have a potential conflict of interest as media buyers and planners while stations have a vested interest in promoting research favourable to themselves. In the history of the United States, these custom studies inevitably ‘created a cacophony of incomparable audience claims’ (Miller 1994: 63).

As the authors have argued, and as China as a modern exemplar demonstrates, there needs to be a different approach if ratings are adopted, one that can arbitrate competing claims and be accepted by all as a fair and representative measure. For this to happen the several interests involved with competing media outlets, advertisers and advertising agencies have to come together to help construct and accept a standard measure and the convention that goes with it. Syndicated audience research was born out of this necessity to construct ‘a common template for assessing the audience’ so as to permit the orderly development of

an advertising and media market. Writing of the US experience Miller argues that advertisers sought a trade-off in survey design. More outlets could be measured with less detail, but to the detriment of rich data which is more costly. An underlying demand for rich demand, though, remained (Miller 1994: 63).

Ratings research is a large-scale research undertaking. Both costs and logistics ensure that ratings research is syndicated because its costs are borne by the various media players although, as we heard from Gale Metzger, in the United States the broadcasters themselves usually ended up footing the larger part of the bill. These outlets have an interest in coming together despite the often intense competition among them. In the United States they had 'something important in common – the same market to sell' (Barnathan 1964: 172). In the case of China, Green is pointing out that Chinese stations need to understand this as they already saw themselves in competition. To this end it makes sense for the different Chinese television stations to support the provision of regular, comparable, 'longitudinal information about audiences that can be used to sell advertising space and time' (Miller 1994: 63). China, in fact, recognized that television and radio have a common interest in ensuring that the ratings data is taken up by advertisers and media sales people. To ensure the credibility of the ratings system among the several groups who rely on it, the organization providing the syndicated research must be structurally separate from both media businesses and advertisers. Ratings providers must function and be seen to function as independent brokers. The Chinese government partnered with TNS in its initial incarnation. It has been and will be up to the players in the market to see whether the agreements to the convention are kept. What ratings providers seek is to establish an 'official truth': an agreement for the description of audiences for a number of media products which can be offered for sale to a number of clients. As Miller perceptively puts it:

A service's case for becoming a convention is heightened to the extent that it can claim to capture all of those viewing, reading, or listening – every audience member for every program or articles, at all times. This criterion is particularly important to media clients who will pay the most toward the service's existence. The service's bid to offer the official reality will also benefit from a demonstration that its method detects only real viewing, reading, or listening and not false or misleading impressions. This criterion is particularly important to advertisers on whom the media clients rely for income. (Miller 1994: 67)

Persuading the full range of media players from the competing media companies that a particular technology or service for collecting and analysing audience data is suitable is no small achievement. Most importantly, the need for these services

to prove themselves is continuous, especially when new systems or new players win ratings services. The early use of ratings was, for the most part, a check on performance of individual radio stations and their programmes, much like an internal audit on performance as well as a currency for trade. Indeed, there was no expectation that audience ratings *should* be made public on a routine basis, but that would come later. As a consequence discrepancies in data were a largely intra-mural matter for the broadcasting industry. Early ratings surveys were also not conducted on as regular basis as today. In Perth in Australia, for example, there was for many years only one sweep a year over a few weeks. Diary and personal interview data were necessarily limited. With no computational power the data analytics were narrow as was the 'window' to the audience. Agencies and stations could only ask so much of the data, and ratings research could only collect and analyse so much data from the audience.

The Silent Revolution

Australia is a good example of the impact of peplemeters and the expansion on the window to the audience. In 1980 McNair-Anderson in Australia was partially acquired by the UK-based Audits of Great Britain (AGB) ratings research group. AGB were investing heavily in meter and other related technologies. The cost involved in developing meter technologies was outside the scope of the Australian company. This had been the case even in the early days of ratings in the United States where Nielsen took a massive personal risk in his investment in the Audimeter. AGB McNair lost the metropolitan TV television ratings tender to Nielsen for start-up in 1990. Both tenders were for peplemeters – 'an electronic push-button way of recording the presence of individual viewers, with periodic prompts and the meter itself automatically recording which channel is on at the time' (Barwise and Ehrenberg 1988: 176).

The trial of peplemeters was in 1989 for rollout in 1990. The longstanding use of diaries in the television market stopped in December 1990, although the diary method is still used today for remote regional areas, including Darwin. After Robert Maxwell died in 1991 AGB worldwide went into a management buyout including the Australian and New Zealand operations, and ACNielsen (as the Nielsen company was known) then purchased the Australian interests from the management buyout team in 1994. ACNielsen as a result became the main provider of Australian radio and television ratings. It held this position until AGB Italia won the metropolitan TV ratings contract in 1999.

The peplemeter trial was run alongside the diary method over several months and showed as much as a 20 per cent increase in viewing time reported by the

peplemeters, especially at non-peak times. Peplemeters, it was argued, are better at picking up single-person viewing. Single viewers tend to under-report their viewing in diaries. This under-reporting was widely known within ratings agencies prior to the introduction of the peplemeter. Single viewers often under-reported for privacy reasons such as not wanting other members of the household to know and because they were less likely to remember to fill in the diary when unaccompanied.

By registering continuous viewing, peplemeters pick up more audience movement within programmes. The peplemeter permits the 'assessment of all televised offerings that can be received by the sets; it does not tire no matter how long the sets are on; and it counts reception time continuously' (Miller 1994: 68). Pushing buttons on a remote that sends messages to a set-top box on the computer seemed to remove the problem of fatigue in filling out diaries and truthfulness. For all of the importance of the turn to peplemeters, the most significant 'revolution' in ratings after 1970 was not the replacement of diaries by peplemeters but the quiet revolution associated with the coupling of ratings with computational power. In this context peplemeters and computers permitted a greater detail and range of operations, better interrogation of data and additional manipulations of ratings data in analysis routines.

Computer-related programmes started to become a feature of syndicated audience measurement from the early 1970s in Australia. In this period the computers were all mainframes requiring specialist programmers within ratings agencies to run them. Over the 1970s a range of different types of services were explored or set up, including combined media-monitor products. McNair in 1970 had its single source Prime Prospect Profiles. From 1971 AGB also provided marketing information for products and services from a single-source data collection called Consumer Market Profiles. AGB McNair installed the first online ratings data delivery service to all the main advertising agencies in 1981. AGB Australia introduced the world's first electronic diary consumer panel in 1986. The computer concept provided for use of in home, online portable computer terminals recording purchases of all barcoded products to map grocery buying. The peplemeters were established in 1990 with a Sydney, Melbourne, Brisbane, Adelaide and Perth panel of 2,400 people replacing the Sydney and Melbourne panel of 1,200. Peplemeters permitted the ratings results to move towards becoming an overnight service as results from the previous day were 'uploaded' to the ratings' agency computers via the telephone line for processing overnight. This 'overnight' phenomenon had initially emerged as a premium ratings service in the United States in the early 1970s for large markets such as New York and Los Angeles. It was becoming a feature of the basic service in Australia. Large-scale and immediate data collection also brought with it greater demand for publication of ratings results, unknown in the early history of ratings. People today can look in a magazine and see the programme and station ratings.

The advent of peplemeters coincided with the PC revolution which had not only distributed to the individual desktop computational power previously the preserve of mainframes but had also encouraged the development of software programs designed to carry out data analysis routines previously the preserve of programmers. This combination ensured that a greater range and variety of data manipulation of ratings results was able to become routinely available and delivered to the burgeoning workstations of advertisers, television and radio personnel and the like.

Ratings operations also moved from the back room of ratings agencies to the front room of media businesses, advertisers, agencies, media planners and media buyers. What were previously more incidental, ad hoc and expensive research undertakings on the part of ratings agencies for selected clients, requiring a programmer with advanced programming skills, now became more frequent, routine queries on the part of dispersed users of ratings. The databases were now available on their desktops and delivered through software programmes without the need for specialist programmers. The ratings book was being replaced by the ratings database.

For example, ACNielsen's 'Media Advisor' was integral to its provision of ratings services over the 1990s. This personal computer software package was designed to run under Microsoft Windows and allow the user to examine audience trends over time. The database program used statistics such as ratings, share, reach, frequency, hours and minutes viewed, audience profile and the number of people watching as its base. It contained a variety of software modules designed to cover all aspects of analysis and planning for programmers, media business sales departments, schedulers, advertising agencies and advertisers. Media Advisor included modules for dealing with smaller periods of time (quarter hour periods), with larger day-parts, and for tracking the performance of a programme over time. Over the decade additional features were added: tools to enable the user to analyse a television advertising campaign after the fact, to display the ratings according to a person's viewing patterns, to analyse audience 'loyalty' for programmes and stations, and to build media schedules 'based on TARP or budget objectives'. Software was also developed to better handle reports generated by Media Advisor and these included modules for 'viewing and formatting' and 'managing' reports. The official ratings provider from 2001 for metropolitan television markets, OzTAM, provides a similar facility as Media Advisor with its AGB WorkStation. It has a similar audience database that promises individual by individual, minute-by-minute data delivered overnight, 365 days of the year. The development of these ratings software/databases ensured that the ratings became available to a larger variety of ratings clients who now had available to them a greater variety of pathways into the ratings data previously the preserve of the agencies and their largest clients. Advertising

agencies and media companies alike were now conducting their own analysis of the ratings data. With this move to a ratings database came a corresponding increase in the complexity of data collection and analysis, and the complexity of intermediate decisions on the audience and its relative value to advertisers and stations. Media planners emerged as intermediaries in the buying and selling of advertising based on the ratings and consumer profiles research. Other intermediaries emerged capable of interpreting ratings data and linking ratings data either to other datasets or customized qualitative and quantitative survey instruments. The 'window' to the data was clearly changing to respond to new circumstances including advertiser needs. In the 1930s soap manufacturers were the largest advertisers and their judgements were based on information on a simple categorization of the audiences as 'housewife'. By 2000 there are thousands of advertisers with judgements based on hundreds of different classifications of audience.

There is no doubt that the need to know about what audiences and consumers do has increased in complexity and computers have provided one means of drilling down to provide this information. Slowly and imperceptibly a paradigm shift was occurring in ratings research with the move from the ratings book to the online, disk and CD-based ratings computer database. At one level not much had changed: the databases provided the same basic information as the ratings books did and peplemeters were still the basis for this information. But the presentation of the information had changed. It had moved from being pre-structured to cover the most common analyses of the data in a book based delivery system to a more open-ended resource capable of offering diverse users a variety of manipulations of constantly updated data based on their individual needs. With the ratings book the end user had very limited means of handling, responding and interpreting data – they were inevitably constrained by the costs of data manipulation. With the ratings database it was now the users who were doing the manipulations. The ratings agencies were now concerned to facilitate their users in their use of the software/database, to ensure that the analytical tools met client needs and to develop tools to allow additional queries to be made using the software. 'Knowing their audience' was now the business of ratings agency clients in a new way: the programme-makers, the schedulers, the space sellers and buyers, the media campaign managers/creators and so on were now 'constructing several pictures' themselves from these databases. If the ratings have always had an 'end user' focus, this embodiment in a ratings database placed special onus upon the users of the data to manipulate, extract, evaluate and generally make meaningful the ratings information.

The ratings database also placed a different onus on the ratings agencies. They now needed to provide users with a sophisticated understanding of the nature and character of the audience and of the database. Ratings agencies were having to

cooperate with their clients not only to facilitate the generation of information but also to illustrate, train or otherwise inform an increasingly larger range of clients in how to use the system and carefully explain its limits. Because the focus was now on the client generating meaningful data from the system, ratings providers had a significant and ongoing information management and software training problem. Products like Media Advisor required people who could use the software in sophisticated ways and who knew the limitations of the database and had developed appropriate research techniques.

Like any software and database there were significant risks for users and providers alike. Clients were being presented with a mountain of information and a variety of pathways through this information. And this carried a risk in that too much information can complicate rather than simplify the decision-making of users. Another risk is that the end users might have an imperfect understanding of the meaning of ratings and ratings methodologies – leading to poor decisions on media placements, media vehicles, presentation to clients and so on. Consequently training needs increasingly occupied personnel in ratings agencies that were looking for ways to disseminate information and develop capacity for handling media database product on the part of those already in or likely to go into advertising agency and media positions. For clients to use the ratings efficiently and effectively they need to understand both the potential and limits of the data. With the increasing range of manipulations able to be done the tolerances and uncertainties in the ratings instrument are becoming more visible and less a specialized understanding.

With the move to PC-based information, ratings services around the world discovered new clients were emerging for these services. As the principal costs for ratings measurement are ‘first copy’ costs – the costs associated with the measurement process and producing the database – once the database is set up the service can be sold to additional clients at minimal expense. These new clients placed further training pressures on ratings providers. The rapidly developing PC market also made it possible for additional services to be developed which sought to develop more comprehensive understandings of audiences in relation to different media and not just in relation to different outlets within a particular medium. From one side organizations such as Roy Morgan developed Roy Morgan Single Source based on a large interview sample (50,000) with questions related to lifestyle, attitudes, media consumption (TV, radio, pay-TV, cinema, newspapers and magazines, catalogues and the internet), brand and product usage, retail visitations, purchase intentions, service provider preferences, financial information and leisure activities. From the other side ratings providers like Nielsen started to bring together its range of previously separate software/database products.

Nielsen's 'Panorama' is one such syndicated multi-media marketing database that integrates consumer demographics, product usage and media consumption. Here the aim is to add value by fusing together its various products. In mid-2001 Panorama was drawing on Nielsen's TV and radio ratings data, its readership surveys, and specifically designed product and service usage questionnaires with plans to 'incorporate consumer goods purchasing information from the ACNielsen Homescan consumer panel'. This fusion product promised media planners and sellers the capacity to 'go beyond demographics when analysing media consumption patterns'. It was sold as enabling media strategists to 'conduct detailed analyses of viewers by demographics, print and television media, and product usage'; brand managers to 'define target markets and profile consumers by demographics, lifestyle habits, product usage and media consumption'; and media sales persons people to 'learn about your target's buyergraphics and doergraphics' and profile programme audiences against category and major brand users to demonstrate to advertisers the niche audience efficiencies of lower rating off peak programmes.

'Superior Technology': ATR-OzTAM and ACNielsen Controversy in Australia

The trend towards more tailored ratings, customized at the individual level, 'micro-audiences' and the shift of audiences to diverse media has had a similar effect on television as it had on radio. These different considerations no doubt weighed heavily on the minds of the technical experts and executives from the networks when they met to write the specifications for the tender for metropolitan television ratings in Australia in 1999 and to evaluate the rival tenders. The ratings debate of late 1999 and the first half of 2001 was both like and unlike previous debates. Like the first peoplemeter contract, it was over a change of provider as the incumbent: ACNielsen lost the lucrative metropolitan television ratings contract to newcomer AGB Italia's Advanced Television Research (ATR) – which provided OzTAM with its data. Unlike previous controversies, which were between different research technologies and methodologies, the OzTAM/ACNielsen controversy was over different results using the same peoplemeter technology. This similarity undoubtedly made the reported discrepancies both surprising and disturbing.

There were two parts to this debate. The first part, in late 1999, was in the aftermath of the announcement of a change in ratings provider. It centred on the networks' motives for dumping the incumbent in favour of a comparatively unknown newcomer. The second part of the controversy occurred in the first half of 2001 as discrepancies between the two systems became evident and set the stage for public dissension over just 'whose' figures were accurate. Any change in

the provider of ratings conventions is bound to excite controversy. The incumbent has all the advantages of incumbency: it is a known quantity, with experience and a reputation as a known agency for providing reliable and comprehensive measurement for a variety of media outlets and related syndicated services. In Australia ACNielsen had been providing the 'convention' for several years and in its provision of several syndicated services was well placed to provide ratings services for converging media environments. Furthermore this was an incumbent that not only traced its history back to the beginning of ratings research in Australia but also carried the Nielsen name – a global name in ratings technology and the major supplier of ratings services within comparable English language markets.

By contrast the new provider had to establish its credibility with both prospective users, particularly those not involved in making the decision to select the new provider, and with the broader public. In this case ATR was a 'new player' unknown inside English language television markets. Its reputation for ratings services came from Latin America and continental Europe. Rather than setting its credibility against that of the incumbent, ATR/OzTAM promoted its service as an innovation – a technically superior service. Additionally, it attempted to secure the services of credible figures drawn from within the Australian ratings industry in the period leading up to and immediately after being awarded the contract. Ian Muir, a prominent figure in audience measurement in Australia for several decades, became OzTAM's chief executive officer and was joined later by other former ACNielsen personnel. (Interestingly, Muir had led the losing AGB McNair bid for the peplemeter contract a decade earlier; his 'crossing to the other side' was critical to the local standing of the AGB Italia bid.) ATR also relocated one of its European experts to Australia for the critical start-up period (Wilmoth 2001).

Further complicating the credibility problem facing the new provider was the move on the part of the television networks to 'own' the data previously owned by the provider. ATR is supervised by OzTAM, a company established and financed by the three commercial networks. While ATR does the work, OzTAM owns the results and sells them to subscribers. This structure raised important and legitimate concerns as to the potential for conflicts of interest in the provision of ratings data. While it is not uncommon for media outlets to 'own the data' internationally, this was not the Australian practice. This move raised several concerns about just how independent, rigorous and believable the OzTAM data were likely to be. Just how much control over information was being demanded by those who paid for it and relinquished by the ratings provider? These integrity concerns were a real issue for those users not involved in the decision to change ratings providers, such as advertisers and media buyers. They were also a major concern for network television's competitors such as pay-TV television providers. In a local climate where free-to-air television

networks are widely perceived both locally and internationally as having inhibited pay-TV television and other new kinds of television development, the ATR decision was going to be inevitably read as another move by the networks to protect their patch at the expense of others. Certainly, pay-TV television providers had less reason to cooperate with OzTAM – the provider of only free-to-air ratings – than with ACNielsen, an independent company, who provided ratings for several media services.

As Miller notes, persuading the full range of media players that a particular technology for collecting and analysing data for their audiences provides ‘the basis for their costly transactions’ involves the research firms in ‘an elaborate ritual of testing and “validating” the new measurement system’ (Miller 1994: 66). ATR/OzTAM had to ‘prove’ itself particularly in the crossover period between the two systems. With the television networks initially claiming that their choice of ATR was based on its ‘superior technology’, they undoubtedly expected the new system to provide very little in the way of surprises. It would be more a tweaking of the system to reveal greater levels of detail on exposure to programmes and advertising, and to prepare for the circumstances of television environments as much as seven years hence. What they got in the initial set-up period of the ATR panels in early 2001 was something quite different.

In the first weeks of the new rating period of 2001 there were considerable differences between the OzTAM and ACNielsen results. In the week beginning 11 February, OzTAM figures for Melbourne as a 30.5 per cent audience share for channel Nine; 29.5 per cent for channel Seven; and 19.3 per cent for channel Ten. ACNielsen figures had Nine on 33.4 per cent, Seven on 26.7 per cent and Ten at 21.4 per cent. The OzTAM figures stunned the industry, giving rise to concerns that it might be seriously flawed. While the discrepancies between the old and new results narrowed over the year as the OzTAM panel was ‘fine-tuned’, by May it was clear that the OzTAM ratings system had resulted in a consistently improved outlook for Seven (Balnaves, O’Regan and Sternberg 2002). Nine may have won seven of the twelve ratings weeks up to that point but it was only narrowly beating Seven with a 30.5 per cent share compared to Seven’s 30 per cent. Seven had narrowed Nine’s lead considerably – a circumstance which encouraged projections by analysts such as Credit Suisse First Boston of increased earnings for Seven and diminishing earnings for Nine over coming. SBS’s ratings also went up (SBS thought this was because of better coverage of non-English-speaking households on the part of OzTAM); ABC’s went down; and the results for the Ten Network were mixed though improving over the year (Balnaves *et al.* 2002).

At the same time the ratings results were also indicating trends towards less overall television viewing. This was predicted to translate into pressure on advertising

revenues as media buyers reassessed the value of free-to-air television. It also showed that audiences were moving into and out of programmes in ways which considerably complicated understandings of programme popularity. David Keig of Keig and Co had used OzTAM's facility to 'take a snapshot of who is watching every minute' to 'track the way a viewer flicks from network to network' (Dale 2001). Based on Keig's analysis of the average prime-time programme in March 2001, only '36 per cent of those who dipped ended up watching the whole show, while 51 per cent of the supposed audience watched less than a third' (Dale 2001). For Keig, 'Most programs now are like revolving doors, with people coming in and out all the time'. These audience behaviours would need to be factored into media buying and advertising placement.

A damaging controversy ensued. The controversy was not initially centred on the difficult methodological issues surrounding the construction and implementation of ratings systems, particularly its panels. Quite the contrary, it began as a controversy with a far simpler narrative. This was a 'war' among the commercial interests of the various parties: Nine supporting the ACNielsen results because it conferred upon them an unambiguous no.1 spot; Seven and later Ten supporting OzTAM as its figures gave them their best chance in years to whittle away Nine's lead. There were allegations of 'greed' on the part of those reluctant to accept the 'umpire's verdict' – with Nine being seen as reneging on its own earlier ringing endorsement for the new official convention holder, OzTAM. Indeed, the Nine Network's CEO, David Leckie, was widely reported as having pushed the hardest for the AGB Italia bid. (In a curious ways this helped the 'standing' of the OzTAM system among non-television station users of the system as it indicated the 'independence' of the system in that it could provide bad news for those who paid for it.)

The consensus underpinning the ratings system seemed to be falling apart. As none of the vested interests involved could settle the matter in their favour, queries about both systems intensified. Was there some circumstance polluting what ought to have been an objective, empirical project of determining the ratings? There was a crisis in confidence in the system. Journalists reported warnings from media buyers that their clients were holding back from buying television advertising as the controversy continued. With advertiser media-buying decisions starting to be affected a growing consensus from advertising agencies, media planners and media buyers was that the damaging controversy needed speedy resolution. Intermediaries drawn from advertising such as the Media Federation (representing 'most media specialist agencies') emerged seeking to 'broker a truce between warring factions'. From their point of view the issue would not be resolved by maintaining two parallel rating systems – that would only 'continue to show that they were different'. The central issue for the Media Federation was in its President Peter Cornelius's words,

‘why they are different?’ Because of the *prima facie* similarity between the two systems, the methodology – how ratings panels and data collection techniques are constructed – became the issue around which the conflict could be resolved. Was one of the panels flawed? Was it ACNielsen’s panel or OzTAM’s? Or both? Or neither?

With such a damaging stand-off undermining the credibility of the convention and with two panels causing consternation in circumstances where advertisers had become accustomed to one agreed standard, an independent audit became necessary. Professor Peter Danaher, a New Zealand expert on sampling, was brought in to ‘adjudicate’ the dispute – that is to explain the differences between the systems. Central to his brief was to assess the representativeness of the two panels. In making his assessment, census data on the makeup and character of the Australian population would be central. While the report is confidential, the absence of commentary would seem to indicate that Danaher’s report contained few surprises. It is almost as if his appointment as auditor itself brought the controversy to an end.

When Danaher reported his results the differences between OzTAM and ACNielsen results were narrowing, there was a growing consensus that the differences between them had been exaggerated, and that there were not serious enough flaws in the OzTAM panels to warrant revisiting the ratings contract.

As Julius Barnathan noted in the context of public discussion of ratings in an earlier period in the United States, discrepancies between ratings systems are not issues to do with the rigging or fixing of ratings by broadcasters, but turn on the ‘operational flaws in the techniques used by each rating service’ (Barnathan 1964: 172). Typically, these flaws are ‘known to the sophisticated researcher but unknown to the layman’. The problem the layperson encounters in these public controversies is the uncertainties of the ratings – the fact that they are ‘estimates’ not ‘hardfact yardsticks’ (Barnathan 1964: 173). While the discrepancies between the two systems may be unsurprising to an insider and be may well be within acceptable bounds, to an outsider these same discrepancies point to flaws which undermine confidence in the ratings and therefore the decisions flowing from these.

All concerned had been educated in the ‘the problem’ facing contemporary ratings technologies. Panels by their nature are ‘samples’. They are liable to error, as the media players know. OzTAM’s website foregrounds the fact that ‘like any survey, a sample will never be as accurate as surveying the entire population’, so ‘results obtained can and do fluctuate’. Anyone running an analysis using OzTAM’s data will therefore need to ‘always consider the size of the sample from which you are drawing your results from’ (OzTAM 2001b). Samples are part of the trade-off for ratings research. No two samples are going to be exactly the same. It is difficult

even with larger panels to be representative in a fragmenting media environment increasingly characterized by niche audiences and a culturally diverse society. It is therefore likely that both OzTAM and ACNielsen figures were both right.

The concerns that the OzTAM and ACNielsen debate has generated about the representativeness of panels, their size and therefore the accuracy of modern audience ratings techniques will not disappear. The increasing individualization and sheer quantity of data create greater uncertainty and will put pressure on panel size. The problem existing panels pose for broadcasters such as SBS, according to Ken Sievers is that the current sample sizes have been 'relatively efficient in measuring the mass audiences of the commercial stations who achieve quite high levels of ratings' but present difficulties when dealing with smaller, niche audiences (Sievers 2000). OzTAM and ACNielsen were competing in the tender to become the 'official' reality for audience measurement, but when their results conflicted the debate was not resolved among the players themselves but in the public domain, with the protagonists being the very owners and users of the ratings. The issue of 'truth' was not simply methodology but the process of coordination and agreement among the different players who define and agree on the social convention. Public trust is also a part of the agreement. The OzTAM and ACNielsen debate brought in all the media players because they all have a stake in how ratings, as currency, are defined and accepted.

'Superior Technology': Nielsen versus Hooper, Nielsen versus Arbitron

Some months ago, the rating people were vigorously assailed by a distinguished committee of the National Association of Broadcasters for producing 'what appears to be conflicting testimony' on the standing of shows. Last December, they were attacked by Sponsor, a trade publication for radio and TV advertisers, in a report headed What's Wrong with the Rating Services? Among other criticisms, the Sponsor article voiced suspicion regarding relations between some rating services and certain of their customers. Said the report: 'Stations, agencies, all bring pressure to bear to keep ratings high. Sponsor has seen letters from stations to rating services promising to buy the service "when you can show us on top".' Currently, a second important group of broadcasters is investigating the whole ratings industry.

B. Davidson (Davidson 1954)

Sautet has an interesting thesis about modern firms in competitive environments. Firms have visions of what the future could be like (Sautet 2002: 83). The successful

ones are 'alert' to opportunities and the discovery of not yet perceived or used knowledge.

The US audience ratings history is a story of entrepreneurial action and success by Nielsen in the face of powerful competitors and defence against continuous attacks on credibility from a range of players. At some points in its history Nielsen was an innovator and at other points it was an incumbent wielding its power and taking advantage of innovations created elsewhere. In the Australian system, after a period of entrepreneurial competition, the ratings contract became the mechanism that framed the actions and bids by particular entrepreneurial firms. In the contemporary moment, there are in US audience measurement many competitive rivals in different fields but the last great rival to Nielsen in the United States was the American Research Bureau (ARB). The ARB changed its name to Arbitron in 1973 because of concerns that a name like the American Research Bureau sounded governmental.

Nielsen has been able to obtain remarkable longevity as a firm. It has operated in an 'open market' and has not had the 'contracts' to hide behind. As Penrose wrote:

A firm specializing in given products can maintain its position with respect to these products only if it is able to develop an *expertise in technology* and *marketing* sufficient to enable it to keep up with and to participate in the introduction of innovation affecting its products. (Penrose 1995; Sautet 2002: 135)

This would suggest that a firm has to maintain its advantages through continuous developments in technology and in their expertise in marketing. Both fronts seem important. Marketing here should be interpreted broadly as Nielsen aggressively used and uses the courts, congress and various other venues to support its strategy or defence. This becomes particularly important when we consider the very idea of 'superior technology'. C.E. Hooper mobilized Hans Zeisel to support him in the critique of CAB's techniques, claiming superiority of technology in audience ratings. The broadcasters also never forgave CAB for denying them access to and formal participation in its work until 1936.

But Hooper was soon to find himself under an even more aggressive attack from Arthur Nielsen. He launched a public and academic campaign against Hooper, his diary method and, not least, sampling. Moreover, Nielsen was determined to have the Audimeter and automatic recording accepted as the 'superior technology'. In this, Nielsen was brilliant. His established and close links with strategic partners paid off. The Advertising Research Foundation (ARF) conducted the first major critique and recommendation for audience ratings standards, completed in 1954. It examined the conditions under which a study was made, questionnaire design, interviewing procedures, sampling, executive execution of survey, editing, coding

tabulation and interpretation of results. This process involved over 100 meetings and conferences, not including consultation and written feedback from the seven ratings services. The study covered all the major methods of the time:

- 1 diary
- 2 recorder
- 3 personal coincidental
- 4 personal roster recall
- 5 personal unaided recall
- 6 telephone coincidental
- 7 telephone recall
- 8 combination telephone coincidental and telephone recall
- 9 combination telephone coincidental and diary
- 10 combination telephone coincidental and personal roster recall

While the report did not name specific ratings services, it favoured Nielsen's recorder method. By 1950, Nielsen's Audimeter was in 1,500 homes providing data on:

- 1 homes using radio
- 2 programme ratings
- 3 station coverage
- 4 homes per dollar
- 5 total audience
- 6 audience flow
- 7 commercial effectiveness
- 8 talent popularity
- 9 programme preferences
- 10 product sales related to programme ratings
- 11 network audience
- 12 unaffiliated stations audience

- 13 commercial audience
- 14 minute-by-minute audience
- 15 cumulative audience
- 16 audience turnover
- 17 frequency of listening
- 18 audience for spots
- 19 non-network programme audience
- 20 network sustaining programme audience
- 21 station area data
- 22 Pacific network ratings

In 1950 Hooper and Nielsen brokered a deal where Hooper would have local radio measurement and Nielsen national radio measurement. Hooper introduced a TV measurement service in 1952, but he was killed in a boating accident in 1954. Television was in 83 per cent of American homes by 1958; however, Hooper's city TV ratings services were sold to ARB in 1959 with Hooper's company retaining local radio, which had expanded to 200 markets. By 1955 Nielsen had spent US\$12 million on the Audimeter and his anxiousness for its success was not surprising. He was also head of the largest market research organization in the world, with 3,500 full-time employees in 17 offices in 11 countries and an annual gross of US\$24 million, of which 22 per cent was from his broadcast division.

What is very interesting is that Nielsen, like Hooper, was quite happy to aggressively argue, with statistics appended, on the weakness of competitive audience ratings technologies, but then to adopt them himself if needs be, just as Hooper had done with recall. Nielsen dropped the Audimeter for the TV markets and substituted it with diaries until the 1970s when the Storage Instantaneous Audimeter (SIA) was deployed. This meter collected data from television sets in homes and sent the data back to Nielsen. Twice a day Nielsen would retrieve data from each of the 1,200 SIA homes. The SIA cost network subscribers US\$300,000 per year. The Nielsen Television Index (NTI) provided continuing estimates of TV viewing and national sponsored network programme audiences, including national ratings 52 weeks a year. It measured TV station audiences in over 200 local markets. An 800,000 diary sweep verified the accuracy of Nielsen's 1,200-home panel. By 1977 Nielsen's US TV services cost its 3,000 clients over US\$26 million. Of Nielsen's overall revenues of US\$270 million, the TV division represented 10 per cent.

Arbitron at the same time ran into difficulties with its meter, having contracted a company called E-Systems to provide a device that was supposedly 96 per cent accurate. However, on delivery 91 per cent of the meters failed within four weeks of operation. The court case against E-Systems was won, but took six years between 1973 and 1979 to come to a US\$3million judgment. Arbitron never caught up with Nielsen in the television market as it went through management and takeover difficulties that delayed or affected the speed by which plans went ahead. With the evolution of computers, however, Arbitron provided innovative services like its Audience on Demand (AoD) service that offered clients the ability to use data in their own computer calculations, including reach and frequency. By 1974, Nielsen and Arbitron each produced 170 reports providing more data than the industry could possibly handle. This proliferation of data also created another form of distortion, where companies would do calculations on the data, even if the sampling tolerances did not allow it. Companies would, with no sense of shame, charge *extra* to distort results, in some cases charging up to US\$8,000 (Banks 1981).

It is not necessarily superior technology, therefore, that wins the day. The ratings provider has to be proficient at dealing with all the politics of the ratings field and be able to negotiate with all the players successfully. This includes, not least, ensuring that within its own methodology, execution of methods and analysis of results that distortion is rigorously controlled, or if coming through hypoing, aggressively pursued. The best example of this is in 1966 when Nielsen discovered that 6 per cent of its Multinetwork Audience sample had received questionnaires asking them to watch particular commercials on a certain day. Nielsen found out that 5 per cent of the National Audimeter sample had also received the same letters. On further investigation Nielsen discovered that Rex Sparger, one of the former investigators of the Congressional committee on ratings, had stolen the names of the Nielsen sample from Nielsen's rubbish. Sparger did not deny this and also said that he planned to write a book revealing hypoing, but Nielsen found that Sparger himself was involved in the practice. Sparger had a contract with Charles Lowe, the producer of the Carol Channing special that went to air at the same time as his hypoing spree. Sparger was alleged to have been paid US\$4,000 to rig the show's ratings. Nielsen sued for US\$1.5 million for illegally obtaining trade secrets, but the matter was settled out of court. Nielsen, the company, knew that hypoing was not just a technical issue for the methodologist but a hot political issue that can embroil government committees and the public. In this case, given the quick actions of Nielsen, the federal government did not order a probe.

There can be little doubt that Arthur Nielsen's strategic nous was extraordinary and his personal experience became embedded in his company's strategic actions generally. What Nielsen learnt, like his competitor Hooper, was that the debates over

innovation in audience ratings methodology and technology were primarily about the disposition at any one time of the broadcasters, the advertisers, governments, technical expert committees, research institutes and not least the public. This did not mean that innovation did not happen, but it did mean that Nielsen as the central entrepreneurial firm in audience ratings was able to deploy knowledge – not yet used or not yet perceived knowledge – in dramatically effective ways that often wrong-footed his competitive opponents. His harnessing of professionalism, discussed in the next chapter, is another example of this.

Summary

Syndicated audience research such as ratings is necessarily contentious. Some of this contention is structural because the reporting of syndicated research in ratings comparisons can never drill down to the levels of specificity required by the various players. Standardization produces tension. In this sense the ratings will always ‘fail’ to satisfy the need for a more detailed picture of the audience. By definition ratings systems only know the audience in certain capacities. The various interested industry parties – broadcasters, programme producers, advertisers, regulators, media planners, ad agencies – need to supplement ratings with customized research. The ‘problem with the ratings’ is not so much a problem with methodology and sampling, as is commonly claimed, but with the ways various players perceive the ratings system and its limits. In circumstances where the ratings are seen as delivering ‘the perfect’ picture of the audience the structural imperfections of syndicated research – arrived at as they are by negotiations among several parties and interests – are ignored. Customized research is necessarily able to give a more in-depth picture of audience behaviours. It is not constrained by syndicated research’s need to establish aggregate measures and agreement among dispersed stakeholders.

Research firms who supply customized research routinely comment on the imperfections and blunt character of ratings research, pointing to the need for alternative (additional) knowledge and research on audiences. In much the same way, public and academic commentary on the ratings system routinely draws attention to the conventional, constructed character of the ratings, their closure to social differentiation, their hegemonic construction as arbiters of programme value and decision-making within the industry in comparison, for example, to the fickleness and changeability of audiences themselves (Ang 1991: 171–94). These several constructions of the ratings persist, alternatively overly inflating and diminishing the value and purpose of the ratings. This public ‘sound and fury’ over the ratings can sometimes obscure the often precise understanding of the limits and therefore value

of the ratings on the part of the various actors and the complementary character of customized and syndicated research.

It is important to acknowledge the public character of ratings research as syndicated research. This aspect of ratings research is sometimes ignored when the ratings are considered as industry-imposed orderings of diverse audience behaviours and pleasures. From the perspective of the ratings provider, the ratings system is at base an auditing and therefore accountability measure. Those involved in delivering ratings systems understand themselves as providing a public service. They necessarily operate 'above' the interested stakeholders who nonetheless support this auditing service and make it possible. With the move towards ratings databases and the interlinking of several databases the function and character of syndicated and customized research alike is changing. It is becoming increasingly common for syndicated research instruments such as ratings databases to form one basis for customized research activities which draw on other databases and qualitative research methodologies. In this way 'ratings research' databases are becoming an integral part of customized research. The 'service' character of the ratings instrument is being extended in ways analogous to the 'data mining' undertaken by private corporations on Australian Bureau of Statistics census data. Does this distract from the core task of providing independent, reliable and benchmarked data? Probably not, as any compromising of independence and of the ratings standard setting function would diminish the market value of the research base for diverse users. But it certainly does mean that the ratings instrument is becoming less embodied in a separate system with quasi-firewalls between it and customized research activities, and becoming more a function capable of being rigorously separated out from other uses of the datasets.

There has been an historical trend evident in broadcast and media markets around the world for broadcast markets with multiple ratings providers to be replaced by markets with single ratings providers. Ever since Nielsen took over the local television market ratings and had these alongside its longstanding national market ratings in the United States, it has been not only the main game but the only game. There have been rumblings over its monopoly provider status and criticisms that it acts as a law unto itself; and significant controversy has surrounded its operations. In different ways similar concerns have been expressed about single ratings providers in Australia, the United Kingdom and continental Europe. Indeed, in the contemporary period ratings provision is often characterized by quasi-monopolies and multinational companies. While there have been considerable controversies surrounding those firms and agencies that provide these single ratings measures, much of this attention focuses on the ratings monopoly rather than on why a monopoly became necessary. Trends towards 'single measures'

within ratings provision connect up with the increasing demands made upon ratings providers to develop a single coordination rule, to construct and operate increasingly larger panels, and to provide more frequent and timely ratings results. They are also connected to the scale of the technological apparatus that needs to be increasingly brought into play. It is clear that in the computer era deep pockets are necessary to conduct ratings operations and to undertake the innovations essential to the continuing relevance of the ratings convention.

At the same time ratings providers, no matter the conditions under which they operate, need to negotiate common expectations of performance and behaviour that are largely taken for granted. This includes expectations of integrity, structural separation and transparency; trade-offs required of ratings providers in managing internal expectations of research performance and innovation on the one hand, and broadcaster and advertiser willingness-to-pay on the other; and not least managing their dual identity as a provider of a quasi-public service in undertaking syndicated research and as a market research provider critical to the operation of commercial markets. The US experience confirms the existence of these expectations, even if the way they play out may be different, country by country.