Constructing an Audience Ratings Convention

Mark Balnaves and Tom O'Regan
Edith Cowan University and University of Queensland

Mark Balnaves is Professor of New Media at Edith Cowan University. He has co-authored and co-edited on the diffusion of media in the Penguin *Atlas of Media and Information*, on research methods in the Sage publication *Introduction to Quantitative Methods* and on audiences and media ratings the University of Queensland Press book *Mobilising the Audience*. His co-authored *Media Theories and Approaches: a global perspective*, in press, with Palgrave Macmillan, provides an overview of trends in media studies. Contact: m.balnaves@ecu.edu.au

Tom O'Regan is Professor of Media and Cultural Studies and Head of School, English, Media Studies and Art History, at the University of Queensland. He is author of *Australian Television Culture* and *Australian National Cinema*, now seminal in the field. His current research is in the areas of contemporary film studio, Hollywood film and television production ecologies with a particular emphasis on their Australian instantiation, new "service" models for audiovisual media, independents and cultural policy, audience development in the arts and media industries, and contemporary reframings of cultural policy in the light of creative industries and creative city policy paradigms. Contact: t.oregan@uq.edu.au

Abstract

Broadcast ratings and survey forms of media research undertaken for advertisers, producers and broadcasters have been an integral part of our broadcast system for so long that it is difficult to see how much work had to go into constructing them and then having them accepted and utilised by the different industry players. Early promoters of ratings and media survey research more generally had to ‘invent’ then sell the very concept of ratings and survey research to reluctant radio stations, advertisers and advertising agencies. This paper will examine how the new forms of knowledge and practices and techniques associated with the application of social survey research techniques became integral to the very management and orientation of the businesses associated with broadcasting.

1 *Acknowledgements:* The authors are grateful for the help of Des Foster, Ian Garland (ASTRA), Ian McNair, Don and Mark Neely, and Gwen Nelson in exploring the history of their respective audience ratings companies or the contemporary moment. The research is part of an Australian Research Council (ARC) project The Emergence, Development and Transformation of Media Ratings Conventions and Methodologies in Australia, 1930–2008.
**Introduction**

Audience ratings are a peculiar phenomenon in the world of commodities. Although broadcasters collect all their revenue from the audience market, their success in the audience market depends upon their ability to compete successfully in the content market - because they must compete for audience attention (Napoli, 2003). Bill McNair, an original founder of Australian audience ratings, recognised this phenomenon early last century together with the need for a system that measured audience attention, despite genuine resistance from the broadcasters themselves. As radio evolved as a medium, he was ever mindful of the need to sell, argue and manage on the basis of numbers and surveys using statistically representative samples of audiences; a new concept at the dawn of mass radio broadcasting. McNair kept calling this activity from the 1930s to the 1960s 'scientific research' – research as a means of better knowing audiences on a mass scale. Today we would simply call it ‘market research’.

Typically what was being promoted and needed to be accepted was the general concept of ‘scientific research’ – market research – and its two typical variants: customised and syndicated research. Market research companies like McNair and Anderson (both of whom ran competing ‘ratings’ right through to the early 1970s) and Morgan (who ran the Gallup public opinion polls) were all companies whose very business model was predicated on undertaking both kinds of ‘scientific’ research for different clients drawn from the suite of advertiser, advertising agencies, radio and later television stations and networks. These companies offered to develop and deliver in conjunction with various industry players (1) ‘customised research’ undertaken for a particular purpose, commissioned by usually a single client and ‘private and commercial in confidence’ to that client; and (2) the even more novel concept of ‘syndicated research’ – like the ratings – where a variety of industry players pay for a regular and repeatable surveys conducted upon the same ongoing premise which are made possible by the very scale of the buy-in by advertisers, agencies, and stations. The very idea of competitive industry players paying for survey research covering the whole of broadcasting in a particular geographical area was initially anathema to the broadcasters themselves – unsurprisingly the push for the adoption of these novel
arrangements came from advertisers and advertising agencies seeking ways of distinguishing among competing broadcaster claims.

The two forms of research activity inevitably fed into each other and sold each other. The undertaking of customised survey research for a particular firm as part of its management priorities generated confidence in survey research techniques and the firm delivering them; while the provision of regular syndicated research surveys paid for by the sector as a whole provided ongoing industry recognition of the ‘scientific methodology’ helping to popularise its methods among sceptical industry managers, leading to greater confidence in commissioning ‘syndicated research’. Through the development and eventual rolling out of media survey research nationwide, the Australian broadcasting audience became progressively constructed as a national audience (Balnaves & O’Regan, 2002).

For many decades of the 20th century Australia was not able to be grasped as a ‘national audience’ in any media or in any media organisation conceptualisation. There are various historical reasons for this arrangement. Firstly, there was an industry resistance to measurement of audiences; ironically from the media itself, like radio. Secondly, there were, over time, limitations in what ratings could do in its classifications or which areas, geographically, it could reasonably cover and with what frequency. In this paper the authors will track the emergence of the national audience, from housewives and their classification into ‘A,B,C,D,E’, through to what became a ‘two currency’ regime in audience ratings persisting until the 1970s.

**The historical account of ratings**

In the United States there has been an extensive historical study of the development of audience ratings (Beville, 1985; Webster, Phalen & Lichty, 2000). These histories focus on how different technologies of audience ratings emerged; how audiences were estimated, valued, and sold; how advertisers and media firms predicted the behaviour of audiences; and how audience economics has shaped media content.
There have been few if any qualitative and ethnographic type accounts of audience ratings and how people involved in creating the market for ratings affected the market for ratings. Peterson (2003, p. 8) gives a useful definition of ethnography as a “description based on intimate, long-term reflexive encounters between scholars and the peoples they are studying”. In what follows, the authors provide an overview of the first part of a detailed historical and qualitative account of the creation of audience ratings as a technology and its impact on conceptions of the national audience. While not a participant observation study per se, interviews with senior managers involved in developing audience ratings, together with archival materials from the National Archives and audience ratings companies, form the materials for the historical research. The choice of those to be interviewed was based on what Burgess (1986) calls theoretical sampling – selection of individuals based on theoretical purposes and key participants’ own knowledge of those involved in the phenomenon under study (in this case the founders of audience ratings). The accounts from participants given below are relatively long, but the authors are following Morley’s (1980) and Radway’s (1991) ethos that emphasises the need to hear participants’ voices within the context of the themes of the research, rather the analyst’s voice alone.

There is, as Donald MacKenzie (2006) points out, a very human side to the actions of institutions, and the trend in sociology and media research has to look at this human side. Audience ratings are a human engine, they shape rather than observe what is going on. They are "an active force in transforming its environment, not a camera passively recording it." (2006, p. 12.)

The interviews cited here are from participants involved in the building of the audience ratings system, managers from both the McNair and Anderson’s rating systems, the two major audience ratings methodologies in broadcasting up until the 1960s.

“Systematic research has hardly been tried”, Bill McNair and early research

Rather, as commodities themselves, the ratings were constructed in response to market pressures, including competition and monopolization
as well as continuities and discontinuities in demand. The ratings producer was no scientist motivated by curiosity, but rather a company seeking its self-interest through the profitable manipulation of demand. (Meehan, 2003, p. 201.)

Meehan gives the impression that audience ratings emerged in a relatively uncontested way as part of a capitalist engine that discovered research and measurement and simply adapted them as a way of increasing profit. Audience ratings history, however, is far more complex than this. Bill McNair, the founder of the McNair ratings system, gave his account of the need for audience ratings in his 1937 work *Radio Advertising in Australia* McNair lamented that in Australia “systematic research has hardly been tried. The agencies with competently staffed research departments are in the minority; and on matters affecting newspaper and magazine circulations and radio owners’ listening habits very little information has been collected.” (McNair, 1937, p. 44.) Of those who tried to do surveys of audiences, “in most cases the results have not been published; this has been of little use to the great body of advertisers. The few surveys on which information can be obtained have differed widely in method and scope.” (McNair, 1937, p. 248.) At the same time, McNair noted the problems that had emerged in the United States where market and audience research had just taken off.

In America when consumer research first obtained recognition, the new technique was quickly abused. Business executives with no statistical training would run off long questionnaires abounding in irrelevant and ambiguous questions, and have them taken by untrained interviewers from door-to-door through all sorts of unrepresentative localities. Completed questionnaires which showed unfavourable results were freely discarded in order to make the final percentages more impressive. Such methods could not but cast discredit on the whole practice of consumer research. (McNair, 1937, p. 248.)

McNair took it upon himself to promote independent and rigorous audience and market research. His book is a report on his own surveys of the radio audiences and listening habits of the time, pitched to convince the media industry of the need for sustained market research. We can also see from his book just how uninterested radio companies and the broader public were in research and how difficult it was for McNair to get data. “Although one would naturally expect the broadcasting stations if
their case was a good one, to take the lead in compiling audience statistics, in actual fact they have tended to discourage work in this field.” (McNair, 1937, p. 248.)

McNair’s concern with the independence of audience ratings and research has its own history, of which McNair was aware. In 1908 in Britain the Observer proceeded for damages against the Advertisers’ Protection Society, formed in 1900 to ensure that manufacturers who paid for advertising would get better value for their money. Its first action was to invite newspaper owners to divulge circulation figures and when they refused they provided its members with private estimates for leading newspapers and journals. The Society had published an estimate of 5,000 for the Observer’s circulation. The Observer could show that its net sales were over 80,000, but the Society was acquitted, “because, in the struggle for circulation which then consumed Fleet Street, there was no basis for more accurate estimates.” (Harris & Seldon, 1954 p. 14.) After this court case newspapers began publishing their circulation figures. The Society then embarked on a campaign for independently audited figures that led to the establishment of the Audit Bureau of Circulation in 1931. With the establishment of this Audit Bureau the principle of syndicated research had become established in a contested environment.

McNair set the scene for the first attempt at an academic and discipline-based study of audience ratings, to the extent where he submitted his 1937 book as a PhD to a university in Britain. His application was knocked back but his academic interest in audience ratings remained, like his competitive counterpart Anderson.

**Early Segmentations and Early Ratings: “the housewife”, McNair and Anderson**

McNair’s study was the first to touch on specifically ‘audience’ issues. But there were attempts at audience segmentation before McNair’s survey. For example, J. Walter Thompson (JWT) employed psychologists A.H. Martin and Rudolph Simmat to measure consumer attitudes towards advertising (Simmat, 1933). Simmat was brought to Australia to work as research manager for JWT in Sydney in 1929 and began some of the early research into the segmentation of audiences. Indeed, it is arguable that the first segmentation work began started in Australia. Simmat divided Australian society
into four market segments based on the combination of income and housewives. Classes A and B were high income housewives; C and D were average or below average income housewives. Class D says Simmat “barely sufficient or even insufficient income to provide itself with the necessities of life. Normally Class D is not greatly important except to the manufacturer of low price, necessary commodities.” (Simmat, 1933, p. 12.)

Simmat standardised interviewing techniques because experience had shown him that women were usually more effective as fieldworkers than men. “Experiments have indicated that persons with a very high grade of intelligence are unsatisfactory for interviewing housewives ... usually a higher grade of intelligence is required to interview the higher class of housewife than is required to interview the lower grade housewife.” (Simmat, 1933, p 13.) Simmat and his team had interviewed 32,000 Australian housewives by 1932 and advertising was from that research then targeted to specific audiences, with "sophistication" “definitely soft-pedalled” for Classes C and D. “We believe that farce will be more popular with our Rinso [detergent] market than too much subtlety.” (Simmat, 1933, p.14.) Soap and detergent were the major advertising markets during the 1920s and crossed all market segments.

McNair – studiously – does not mention Martin or Simmat’s work in his book, even though he was familiar with it. JWT also supported the McNair radio survey. McNair in his book uses income as the primary segmentation and argued that the household was more important than the housewife as the measure (McNair, 1937).

Gwen Nelson, arguably the first woman in audience ratings research, managed McNair’s sampling, fieldwork and interviews and was a business partner. She gives an insight into why the home was seen as ‘the audience’.

The household was a household, it was a family. They weren’t scattered here, there and everywhere. There was one radio set and the family sat around and listened to the radio. The advertiser thought it was ridiculous, nobody would sit around and listen to a voice. They’ll read it when they see it in print, in the newspaper, and they’ll believe it, but they won’t take any notice of an odd voice talking to them over the radio. And also, the housewife was always at home, or mostly at home, we had no problem
with what we would do with all the outs we found. We were able to get a very reasonable cross section of all the groups and all social structures, which was ideal for personal interviews. (Nelson, 2000.)

The history of audience research in Australia started, therefore, with Bill McNair, but the history of audience ratings as a business started with McNair and Anderson as competitors. McNair was a researcher with J. Walter Thompson (JWT). Anderson was a ‘radio man’ at 2GB Macquarie. According to Don Neely, one of Anderson’s senior managers, Anderson, to get legitimacy, “got the support of the Advertisers Association, and formed an honorary board of governors to give him some sort of status” to match the methodological strength of McNair (Neely, D., 2000).

Anderson chose the diary method as a way of measuring audiences, after talking to colleagues in the United States. McNair, by contrast, experimented with different methods because there were no measurements of radio at all, and settled for personal interviews and the recall method. McNair’s son, Ian, himself a manager at McNair’s, recounted the early history. “JWT was one of the main agencies handling radio work for Lux – Lux radio theatre may have been on then – and Colgate and so on and so forth. They were some of his main clients. In fact, the very first research he did, well, one of the first he did, was for Kellogg’s, I know, because I very nearly got the name Kellogg as my middle name. That must have been around 1933, in fact, I was born on the day that J. Walter Thompson got the Kellogg account. And I think they still have it. My dad said he was nearly going to call me Ian Kellogg McNair, but he changed his mind.” (McNair, 2000.)

Both methods – personal interview and recall and diary – ran side by side as audience measurement techniques for ratings, until both firms merged in the 1970s. For the whole period, from 1944 through to the merger, the debate about the two methodologies was personal and public.

It was a very public argument, because my dad liked sending out circulars and letters, and quotations and methodology arguments and George responded in his way, so the argument was going on all the time. The argument was centered around personal interviews versus diary. At that time, McNair was doing the personal interviews, in fact all through the
fifties and early sixties anyway, the method for both radio and television was face-to-face personal interviews, house-to-house, very controlled samples and all about yesterday’s listening and yesterday’s viewing. And the argument was, because it’s yesterday’s listening, it’s fresh in people’s mind, and they would be able to tell us, and most people do tell, exactly what happened yesterday. George’s argument was diary-based. People do fill out diaries, but the thing that was – the thing that perhaps won the day for diaries – was the fact that by getting seven days of recording from the same respondent they were able to find out from the same people what they watched on Monday, Tuesday, Wednesday and so on, and get a good cumulative audience over a week. Whereas, with the face-to-face one day recall method, you had to match samples each day so you got sample error each day. (McNair, 2000.)

The motives for collecting audience ratings were also personal. Anderson’s commitment to methodology and service, according to Don Neely (2000), meant that the company was never really financially flush.

It was personal. When we were measuring a country television station, we would over-provide, for the reason George said ‘it’s their livelihood. Forget the agencies, it’s the station that’s got to live or die by these numbers, and I want them to be right.’ Now the fact that we’ve quoted a price and are providing a service which was more expensive than we got was quite irrelevant to George. (Neely, D., 2000.)

George Anderson also stuck by his diary methodology, even though there was no direct evidence that it had advantages over recall. He had not, on Don Neely’s account, any evidence of the advantages of the methodology but he stuck with it. That is, until the All-Radio Methodology Study (ARMS) in the 1970s. “Out of this study came the finding that the most reliable form of radio audience measurement was the personally placed and personally collected, pre-listed diary. Surprise, surprise, this was exactly what George Anderson had said.” (Neely, D., 2000.)

When McNair’s and Anderson’s merged, both protagonists fell out of the picture, with their sons, or managers who had been with them for years taking over. Ian McNair’s recollection of the reasons for agreement for using diaries and dropping recall were, he said, based on methodological grounds. It was not so much personal interviews versus anything else, “it was the fact that you got data over a whole week
from a diary; that really persuaded us to continue on with that method for both radio and television.” (McNair, 2000.)

**From audience as home to audience as individual**

The early methodologies and the early assumptions about the audience are, as Mark Neely, Don Neely’s son and currently head of radio research at Nielsen, says, “simple models” (Neely, M., 2008). There was no conception of a ‘national audience’ read off each minute of each day, no conception of millions of individuals doing a myriad of things, and no conception, indeed, of a ‘national audience market’.

In contrast, at methodological level, the field researchers and managers of audience ratings from the 1940s through to the early days of television knew their participants intimately. The phone ownership was not very high, according to Nelson, and it was impersonal. “Your interviewer had to get the confidence of the person you were interviewing, and make them feel that they were really contributing to something. Which I think they did. And they’re the early days. Every month I went somewhere. We were doing three surveys a year in Melbourne, two in Adelaide, two in Brisbane and one in Perth. I went to Perth every year for 20 years.” (Nelson, 2000.)

Nelson’s account also gives an insight into the audience ‘market’, limited by geography and frequency of data collection, but with intimate knowledge of, and closeness to those interviewed.

In the pre-television period, as Des Foster, general manager of 2GB at the time notes, there was no guarantee that audience ratings would be used, especially if they gave a negative result. According to Foster (2000), “They weren’t cheap, and some stations, I recall, who didn’t do well in the ratings and were not doing well, revenue wise, would say ‘we don’t want them,’ they didn’t have the surveys, they didn’t want to know about them."
But once the audience ratings structure was there and once it was established by McNair and Anderson and accepted as a method of currency, changes in the machineries of knowing in audience ratings became intimately linked to changes in demographics or technology or the needs of stations and programmers. The first big change in the ‘simple model’ was in the introduction of the transistor radio and, of course, television.

Foster, who grew up with the radio as a medium for groups, remembered the shift from household to individual, via the transistor. “Suddenly the 12-18 year old group became a very identifiable segment. But I mean all these age demographics became very relevant and it was possible to tailor programmes for them.” The transistor was portable and you could carry it around with you. “Now junior’s got his own radio, and sis has got her radio and they could all be doing their homework in their different rooms, listening to different programmes. This mobility allowed a switch of focus from night time radio and radio was never ever to dominate at night again.” (Foster, 2000.)

While the recall and diary methodologies continued until the merger of McNair’s and Anderson’s, the underlying conceptualisations of the audience had changed during the 1950s. Programming formats and attention to demographics became interlinked. There was, by the time television became standard, no radio or television manager that was not looking carefully at the demographics of their audience and measuring their performance by the audience ratings.

The McNair A,B,C,D,E typology carried over to the television era. Nelson supervised in 1956 the first TV survey done in Australia. “And I remember very clearly, there were 3.2 percent of Melbourne residents, according to us, that had a TV set. No one in the upper classes around Richmond/Carlton area. The interviewer used to check if they found anyone who had a TV, what they viewed yesterday, and how many people viewed. There was one house in Richmond, a very poor house, very much an E, there was a huge group of people. And as the woman said, ‘we’re the lucky ones. We’re the only one in the street that has a TV set. And we charged adults sixpence and
children threepence and they all bring their own chairs to watch TV.” (Nelson, 2000.)

Australia differed from the United States in its methodology and approach. As you can see from the accounts of senior participants in the creation of the audience ratings system in Australia, there was a dedication to the art of audience ratings in and of itself, to the point where a company was willing to lose money to ensure accuracy. Why did Australia keep two competing ratings systems, both of which collected data from the same audiences? Why would hard-nosed station managers and media magnates buy from two methodologies that produced different figures and were expensive? The answer lies in the nature of the competition between McNair and Anderson. Both thought that their methods were the best. Both were respected in the media industry and both were trusted. When the audience became more fragmented and there were more stations to measure, the economics of keeping two ratings agencies changed. But embedded in that older system there was a perception of checks and balances, even if McNair’s and Anderson’s were expensive to run as two separate systems. Those checks and balances, interestingly, did not disappear when the two companies merged in the 1970s.

The merger happened at the end of 1973, which saved the problem for the industry, because there was still an amount of sympathy for McNair, and because he’d been in the industry for so long, having been dumped in the tendering situation. So they were delighted when we announced that we were getting together. But they insisted that the methodology be the Anderson methodology, the format was the Anderson format; in fact it was an Anderson show. Except that McNair was part of the deal. And that really did help, having the two get together. There was never really enough market for two companies in Australia. One company made a lot more sense. So the industry was happy. They had a research committee, they and we agreed to have an auditor, Dr Arthur Meadows was the first auditor, and he would come in and go through all the diaries and all the paperwork to see if we did what we said we’d do. And I believe today there still is an auditor. (Neely, D., 2000.)

The shift to an audited regime in audience ratings in Australia, the authors would argue, is a direct consequence of the way that McNair and Anderson developed over
time. The ‘simple model’ that Australia inherited had benefits over the monopolised model in the United States. For example, the 1963 Congressional committee into broadcast ratings established in the US came about because there did not appear to be checks and balances in the system.

The hearings suggested that the illusion of exact accuracy was necessary to the ratings industry in order to heighten the confidence of their clients in the validity of the data they sell. This myth was sustained by the practice of reporting audience ratings down to the decimal point, even when the sampling tolerances ranged over several percentage points. It was reinforced by keeping as a closely guarded secret the elaborate weighting procedures which were used to translate interviews into published projections of audience size. It was manifested in the monolithic self-assurance with which the statistical uncertainties of survey data were transformed into beautiful, solid, clean-looking bar charts. (Bogart, 1966, p. 50.)

Australia, as the historical accounts suggest, was in a very different situation. The owners and creators of the audience ratings system were at pains to educate their clients and at pains not to abuse the data or to exaggerate its significance. The gloss on data and abuse of consumer research, identified and opposed by McNair, was not a part of the Australia machineries of knowing audiences.

**Conclusion**

The authors have provided some extensive quotes from interviews taken with some of the founders of Australia’s audience ratings system. Audience ratings have never been simply a snapshot of Australians delivered to radio or television stations. Audience ratings, moreover, were not constructed simply in response to market pressures, as though there were never individuals like McNair and Anderson. McNair and Anderson had to convince the market that audience ratings were necessary to the broadcasting industry.

McNair and Anderson differed in their methodological approaches to audience ratings. However, both were dedicated to providing the best service possible, even if it involved at times a loss to the company bottom line. There was indeed in Australia a ‘social convention’ associated with the audience ratings. The efforts of McNair and
Anderson to educate their clients and the public are part of that convention. In retrospect, this social compact is extraordinary, especially when at the same time McNair and Anderson’s United States counterparts were trying to produce an ‘illusion of accuracy’. Radio and television stations in Australia might not have liked the outcome of McNair and Anderson’s ratings if they produced a negative result, but there was never a national inquiry into their methodologies, as there was in the United States for legitimate and illegitimate companies alike.

Change in who counts as part of the audience has been briefly touched on by the authors. Who counts as ‘audience’ is also a part of the audience ratings convention. There was a time when some broadcasters could ‘ignore the audience’ and refuse to purchase McNair and Anderson reports. But decisions on housewives versus homes, important young demographics versus old, were all a part of complex decisions between the users of ratings and their creators. By the time television arrived, no broadcaster or advertiser could ignore the ratings.

Audience markets and economics, it would seem, are not social actors in their own right who make decisions and act on them and all we have to do is to read off their consequences. The legacy of the McNairs and the Andersons, it could be argued, is an accountable and sophisticated Australian public audience ratings market that operated over decades, compared with sometimes controversial and contested international counterparts. Interestingly, the methodologies developed by those like Anderson also remain in practice, with Anderson-like diary formats still used for radio in Australia, and, interestingly, for television in some regional areas, including Darwin.

This paper has not discussed the ‘people-meters’ era. There are other issues that emerge in that historical account. However, this paper has attempted to show that there are complexities in the machineries of knowing in the audience ratings industry, especially in an industry where what is created is not what is sold.
References


**Copyright Statement:** Articles submitted for ANZCA08 remain the copyright of the author, but authors by virtue of submission agree to grant the Department of Communication, Journalism & Marketing at Massey University a copyright license to permanently display the article online for public viewing as part of this conference proceedings, and to grant the National Library of Australia a copyright licence to include the ANZCA08 Proceedings in the PANDORA Archive for permanent public access and online viewing. Articles first published in this ANZCA08 proceedings may subsequently be published elsewhere by authors, provided the next version acknowledges this original publication in the ANZCA08 refereed proceedings.