CONTENTS

AJMR Forum ............................................ Page 4

Research services offered by Australian advertising agencies
Alan T. Shao and David S. Waller ....................... Page 11

Return on marketing investment:
A practical end to end approach for linking marketing investments to shareholder value
Liane Ringham ........................................ Page 21

Consensus versus polarisation:
Measuring ordinal variability
T Paul Hutchinson ....................................... Page 27

Understanding mail survey non-contact:
an examination of misaddressed survey invitation returns
Benjamin Healey and Philip Gendall .................. Page 37

Book Reviews .......................................... Page 46
The Australasian Journal of Market & Social Research (AJMSR) is the official journal of the Australian Market & Social Research Society (AMSRS). It is published twice a year in May and November.

All full members of the AMSRS receive a complimentary subscription to AJMSR as part of their membership.

It is also possible to subscribe to AJMSR:

<table>
<thead>
<tr>
<th>Category</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australian subscribers</td>
<td>$66 AUD (including GST)</td>
</tr>
<tr>
<td>New Zealand subscribers</td>
<td>$66 AUD</td>
</tr>
<tr>
<td>International subscribers</td>
<td>$70 AUD</td>
</tr>
</tbody>
</table>

Original papers are invited. There is an editorial policy to ensure that a mix of theoretical and practical papers is presented. Papers should be of some relevance to at least a segment of the population of practising market researchers. Purely academic marketing papers are probably more suitable elsewhere. All papers are refereed by independent assessors.

Contributors of papers will receive five copies of the issue in which their article appears.

Manuscripts should be typed and double spaced. Diagrams and figures should be provided electronically, and of a professional quality. In writing papers, authors should consider the style of articles in previous issues.

On acceptance of an article, authors will be asked to supply their article electronically, preferably in Microsoft Word format. However, conversions can be made from other packages.

AJMSR is an open publication. All expressions of opinion are published on the basis that they are not to be regarded as expressing the official opinion of the Australian Market & Social Research Society. The AMSRS accepts no responsibility for the accuracy of any opinions or information contained in this publication and readers should rely on their own enquiries in making decisions.

All papers for submission should be forwarded to:

Professor Lester W. Johnson
Editor, AJMSR
Melbourne Business School
The University of Melbourne
200 Leicester Street
Victoria 3053
Australia

All enquiries regarding subscriptions, advertising and other related matters should be directed to:

Australasian Journal of Market & Social Research
Australian Market & Social Research Society
Level 1, 3 Queen Street
Glebe NSW 2037
Australia

Phone: +61 (0)2 9566 3100
Fax: +61 (0)2 9571 5944
Email: publications@amsrs.com.au

Copies of the articles in AJMR may be made for personal or classroom use, without charge and with the publisher’s consent. Copying for any other purpose must first be approved by the publisher.

AMSRS recognises the contribution of Professor Lester Johnson (Editor) and the Melbourne Business School.
June 2005

Dear Colleague

Welcome to the first edition of the Australasian Journal of Market and Social Research, the new name for the official journal of the Australian Market and Social Research Society. The journal's re-launch follows the Society's name change on 1 July 2004 and better reflects the significant role research companies play in public utilities, government and not-for-profit sectors. This edition continues to meet our aim of bringing quality refereed articles on topical issues of interest to our members and readers.

Benjamin Healey and Philip Gendall have answered our call for contributions from our New Zealand colleagues. Their paper investigates the issue of mail survey non-contact, and examines whether certain sub-populations are more likely to be prone to it than others. It also outlines practical implications for cost management, response rate calculations, identification of potential bias, and the development of sampling methodologies to counter survey error introduced by non-contacts.

Liane Ringham presents a model for linking marketing activity to financial performance, and identifies three challenges to be met if an organisation is to successfully implement marketing return on investment (ROI).

Paul Hutchinson considers consensus versus polarisation and how an appropriate statistic be calculated. Presenting two sets of ordinal observations, he posits a different concept of ordinal variability to that recently presented by Blair and Lacy, and Yeger.

Alan Shao and David Waller examine responses from 68 managing directors of Australian advertising agencies. They analyse the services offered by advertising agencies and explore the advertiser's influence on agency functions — with a particular emphasis on the research services offered.

The Forum section of this edition is an extract from the Australian Market and Social Research Society and the Association of Market Research Organisations (AMRO) submission to the Australian Communications Authority, in response to the Discussion Paper, "Who's Got Your Number - Regulating the Use of Telecommunications Customer Information".

Duncan Stuart takes a look at a new book written in response to the current move in market research away from the traditional reporting role. He considers evidence in this book that research has moved toward a more interpretive role and how it is integrated with the decision-making needs of client organisations.

As promised in the last edition, Peter Link reviews two books on developing brands and advertising. Part of the seven-volume set titled Qualitative Market Research: Principle and Practice, these books provide a useful addition to this expansive set.

We again encourage our readers to contribute to the AJMSR. The journal offers an opportunity for you to share your expertise with your colleagues. All the necessary details for contributors can be found on the facing page.

Greg Wayman
National President
Submission to the Australian Communications Authority

Re: Who’s Got Your Number – Regulating the Use of Telecommunications – Customer Information

In May 2004, the Association of Market Research Organisations (AMRO) and the Australian Market and Social Research Society (AMRS), the two peak bodies of the Australian market and social research industry, lodged the following submission with the Australian Communications Authority in response to its request for feedback on its 18 March 2004 Discussion Paper, “Who’s Got Your Number – Regulating the Use of Telecommunications Customer Information”.

Market and social research is an important service that benefits all Australians. From political polls to surveys of automotive customer satisfaction to studies such as the *Household, Income and Labour Dynamics in Australia*, market and social research provides valuable information about the world in which we live. This information helps government, commercial and not-for-profit organisations make informed decisions based upon the interests and needs of their constituents and clients.

In the following pages we provide an introduction to market and social research and the ways in which it serves the public good. We also explain how access to telecommunications customer information improves the quality of this service and why restrictions on the research industry’s access to information such as those contemplated in the ACA Discussion Paper pose serious risks to our ability to provide quality, affordable research.

This submission also highlights several important characteristics of market and social research generally, and the Australian research industry in particular. First, market and social research is not “direct marketing”. Researchers do not contact people to sell them anything, but to gather information on a particular topic in order to help a client organisation make strategic or planning decisions. It is for this reason that market and social research organisations around the globe are often exempted from “do not call” registries and similar restrictions. Second, the market and social research industry is bound by rigorous ethical guidelines, which place limits on the ways in which researchers are allowed to use personal information and ensure that individuals’ privacy is preserved.

Market and social research serves the public interest

Market and social research aims to provide accurate and timely information on topics relevant to government, commercial and not-for-profit organisations, which contributes to their decision-making processes. To achieve this objective, researchers systematically collect and objectively record information from individuals or organisations. Information from any given individual is collated with information collected from numerous other individuals and is then analysed to develop a view of the behaviour, needs, attitudes, opinions and motivations of a population. Armed with this knowledge, businesses are
able to develop programs and products to meet the desires of their customers and governments can tailor initiatives to meet the needs of their constituents. It is, therefore, in both individuals' and the public's interest for people to participate in research, both in order to have input into such decision-making and so the samples obtained truly represent the spectrum of Australian society that will be affected by such decisions. In fact, the Australian public recognises the importance of having input into the decisions of business and government and acknowledges that market and social research is an effective way of influencing those decisions. (Appendix A discusses a research survey that supports these findings.)

The way in which market and social research uses telecommunications - customer information

As discussed above, quality market and social research depends upon a researcher's ability to contact a representative sample of people. In general, access to telephone listings enables market and social researchers to obtain representative samples of the populations of interest that are as close to random as possible. This ensures that the findings of the research are valid, and enable as researchers to make inferences about the population at large based on these findings.

The list that forms the basis of market and social research contacts is referred to as a 'sampling frame'. Market and social research draws upon telecommunications customer information found in both publicly available directories (e.g. electronic White Pages) and marketed products and services (e.g. Sensis-style services) for a range of purposes in order to construct sampling frames that are as accurate, up-to-date and as relevant as possible to the research task at hand.

There are two main purposes for which market and social researchers use telecommunications customer information.

1. To obtain samples in given geographic areas

Researchers may require samples of individuals in given geographic areas for two reasons. The first and most obvious reason is that people in a circumscribed location (e.g. electorate, water catchment area, Centrelink area office, utility franchise, public transport route) are the target group of interest for the specified research project. Sampling for a project of this type is generally referred to as 'geographic sampling'. In addition, where the target group of interest is the whole of Australia (e.g. Federal election polls), it is often necessary that the sample be stratified by location (e.g. states, metropolitan/regional split), so that it represents the proportions in the actual population. In both cases, a properly drawn sample limits sampling error and means that the research can be conducted with a minimum sample size.  

The second scenario is one in which information about an individual's residential (or business) location is used to infer something about the person, in order to increase the probability of reaching scarce targets (or increase the 'incidence'). This is sometimes referred to as 'geo-demographic sampling'.

1 In contrast, direct marketing typically uses telephone listings to contact as many people or households as possible.
For example, if one were interested in examining the impact of a public health campaign aimed at reducing children’s exposure to passive smoking in the home, our target group of interest would be parents who smoke. The incidence of this group in the overall Australian population is quite low (approximately 6%). Therefore, the probability of finding individuals that belong to this group by calling random Australian phone numbers is low and the associated costs are very high. However, the incidence of people who both smoke and have children is much higher in lower socioeconomic areas (approximately 12%). It is, therefore, possible to use publicly available statistics, such as data from the Australian Bureau of Statistics, to identify areas of low socioeconomic status, and then to obtain telephone listings of people who live in that area.

Telecommunications customer information is essential to both of these research scenarios. Because telephone exchanges do not correspond precisely to all geographic boundaries, and in light of the increasing transportability of telephone numbers, geodemographic sampling (e.g. by census collection districts or map grids) cannot be done by inference from telephone number prefixes. Instead, it must be done with reference to postcodes or, in some cases, specific street addresses. For this reason, it is imperative that market and social researchers have access to lists in which phone numbers and address information are linked - such as those held by Public Number Directory Providers (PNDPs).

Overall, market and social research uses such targeted sampling methods in order to increase the 'hit rate', that is, to find eligible respondents as efficiently as possible. With a higher hit rate, fewer households need to be contacted to obtain the desired sample size, which minimises the cost of such research exercises. Without the information needed to generate a high hit rate efficiently, the costs of such research would substantially increase and organisations might be forced to forgo important research because it is unaffordable.

2. To obtain valid telephone numbers for a list of households or businesses

At times, it is necessary in market and social research to use services offered by PNDPs to supplement existing data. This usually involves the sourcing of telephone numbers for households or businesses for which telephone numbers (or up-to-date telephone numbers) are not available. In some instances, client organisations will have lists of customers who have agreed to be contacted for research purposes, but for whom they do not have accurate, up-to-date contact details. Thus, research agencies must use this service in order to improve the list’s utility as a sampling frame.

Market and social research may also use services offered by PNDPs to validate data sourced elsewhere. This may involve the confirmation of the currency of telephone numbers in an existing list (e.g. of a client organisation’s customers) or, alternatively, numbers that have been generated by research agencies themselves for random digit dialing.

Harms from restricting or prohibiting the use of telecommunications customer information by market and social researchers

In all of the research scenarios described above, telecommunications customer information is used to source, supplement or confirm the currency of telephone numbers in order to produce good quality sampling frames, which reduces sampling error and results in the minimum number of people being contacted for any given research project.
Any restrictions or prohibitions on the use of telecommunications customer information have the potential to disrupt, and in some cases prevent, certain types of sampling and research.

If the use of electronic telecommunications customer information were not allowed in these cases, sampling for these kinds of projects would be less efficient, that is, a greater number of households would need to be called to achieve the desired number of respondents fitting the project specifications. Thus, a restriction would not benefit the public, but would impose a greater burden on the community, as more people will be contacted, many of whom will not be appropriate respondents in the specific research. Restrictions will also lead to greatly increased costs for both research organisations and the client organisations for which market and social research firms conduct research.

If unable to access information from PNDPs, one alternative the market and social research industry might be forced to contemplate is a return to the days before electronic versions of the White Pages were available, when researchers manually paged through hard copy directories with pen and ruler. This technique (if allowable under the industry standard, privacy and copyright law) is labour intensive and involves a decrease in efficiency and accuracy, which would result in increased costs for the industry and its clients.

Indeed, in some cases, the increased costs resulting from restrictions on the use of telecommunications customer information would preclude clients’ commissioning desired research. This would have two main consequences. First, this would undermine the ability of businesses and government to make appropriately informed decisions, the social and economic ramifications of which should not be underestimated. Second, it would threaten the viability of the market and social research industry itself and accordingly, threaten the livelihoods of the large number of people the market and social research industry employs and the industry’s contribution to the national economy.

Alternatively, were research to proceed using poorer sampling methods, the resulting data would be of a poorer quality, leading to ill-informed decision-making on the part of government and businesses. The societal costs of a reduction in the accuracy and quality of market and social research is difficult to quantify, but potentially enormous.

In contrast, there are significant social and economic benefits in ensuring that market and social researchers retain the right to use electronic telecommunications customer information for legitimate research purposes involving the construction of sampling frames. Continued access will ensure the welfare of the industry, as well as facilitating the flow of information about the population to those who make decisions that affect the population. More informed decision-making benefits everyone.

It must be noted that researchers outside of the market and social research industry, such as academic researchers and social scientists, also rely on access to telecommunications customer records. Indeed, we have been contacted by academic researchers in the public health field, keen to ensure that the AMRO and AMSRS mission mentions that access to this information for sampling purposes is essential to them as well.

The discussion above focused upon the public service that market and social research provides, and how restrictions would greatly impede that service. Below, we discuss some specific reasons why, if the AGA chooses to limit use of telecommunications customer information only to specific purposes, market and social research should be designated a permissible purpose.
Industry initiatives protecting individuals’ privacy

The market and social research industry has already implemented rigorous ethical and privacy codes to ensure that when researchers use telecommunications customer information, individuals are informed and their privacy is protected.

AMRO members are bound by the Australian Market and Social Research Society's Code of Professional Behaviour (see Appendix B). Academic researchers and social scientists follow a similar code of conduct. Industry members also subscribe to either the Market and Social Research Privacy Code, which has been approved by the Federal Privacy Commissioner (see Appendix C), or the National Privacy Principles. A CD containing a variety of resources to assist industry members in complying with privacy requirements has also been produced and distributed jointly by the industry's peak bodies.

Under these codes, where individuals are contacted based on information sourced from telecommunications customer information records, researchers are required to inform them:

- the source of the research sample;
- the organisation on whose behalf they are calling;
- the purpose of the contact (which is restricted to research), and
- that any participation in the research project is completely voluntary.

The market and social research industry recognises that the long-term success of market and social research depends upon the willing cooperation of the public and business community. Therefore, the industry has developed industry-wide practices that aim both to minimise intrusion and assure the public and business community that research is carried out honestly and objectively using processes that protect the identities and rights of individuals. This is necessary to encourage informed and willing participation in market and social research activities.

In short, restricting or removing market and social researchers’ access to electronic versions of telecommunications customer contact details would harshly impact the industry itself, government and the business community, and the public, without any corresponding benefit to telecommunications customers’ privacy.

The distinction between market and social research and direct marketing and implications for ‘opt-out’ or consent-based schemes

The distinction between market and social research on the one hand, and direct marketing (especially telemarketing) on the other, is sometimes misunderstood by the general public. However, it is an important one.

Market and social research involves the collection of information from individuals or organisations, which is collated and aggregated to give a picture of a market or part of a market. This process guarantees anonymity to respondents unless they expressly waive this right.

Direct marketing includes telemarketing, or the marketing of goods or services using the telephone as the principle medium of contact with the customer or 'prospect'. Telemarketing interviews collect information from individuals and organisations that is used to inform sales-based approaches either at the time or at a later date. Anonymity and confidentiality are therefore impossible.
There is evidence to suggest that many individuals who do not wish to be contacted for direct marketing purposes may be willing to be contacted for genuinely confidential market and social research. In recognition of this - and of the many benefits good research provides - a number of international jurisdictions have exempted market and social research from existing 'opt-out' or 'do not call' schemes. This reinforces the fact that market and social research is different from other activities, such as direct marketing, in which telecommunications customers' data can be used without protecting privacy, anonymity and confidentiality.

At the same time, many people still do not fully understand what "market and social research" entails (and, especially, its distinction from "direct marketing"), and would, therefore, be ill-equipped to make an informed decision about opting out of it. Furthermore, it is difficult to imagine a system in which people would choose to opt out of all market and social research. People often wish to have the opportunity to comment upon subjects in which they are interested or that are important to them (e.g. at the moment, people may welcome the opportunity to contribute their opinions about Australian troops in Iraq, but perhaps not about financial services). Moreover, even if contacted, people can choose to opt out of participating in research at any time; any participation is completely voluntary. Hence, the market and social research industry believes that an opt-out or consent-based approach directed at the use of telecommunications customer data for market and social research is in no one's best interest.

For these reasons, the industry believes that market and social research should be classified as a permissible purpose for using telecommunications customer information. Further, should the ACA decide to implement a mandatory 'opt-out' or consent-based system for the use of telecommunications customer information for all purposes other than primary purposes, the use for market and social research purposes must, at least, be clearly distinguished from other uses such as direct marketing.

Finally, it should also be noted that, if restrictions were only applied to the use of Integrated Public Number Database (IPND) data, our industry would benefit from the existence of more than a sole provider of non-IPND customer data (i.e. more than solely Sensis) to ensure competition in the directory and directory-related products and services industry, and the resulting implications for pricing in this industry.

Credits: Sam Cuming and John Sergeant.

To review the full paper along with accompanying appendices, visit www.amro.com.au
Research services offered by Australian advertising agencies

Alan T. Shao  
North Carolina Ports Professor of Marketing and International Business  
The University of North Carolina at Charlotte  
Department of Marketing  
Belk College of Business Administration  
Charlotte, North Carolina 28223 USA  
Email: atshao@email.uncc.edu

and

David S. Waller  
University of Technology, Sydney  
School of Marketing  
PO Box 123  
Broadway, NSW 2007  
Australia  
Email: david.waller@uts.edu.au

Abstract  
New media are changing the landscape of market research, and so clients are demanding new services to meet the dynamic nature of the industry. It is paramount that managing directors of advertising agencies can ensure that their firms provide suitable services to fulfill clients' needs, including research services. This study examines responses from 69 managing directors of Australian advertising agencies regarding services offered to clients, the availability of those services, and the extent their clients influence agency functions—with a particular emphasis on the research services offered. The findings indicate that Australian advertising agencies are more likely to offer general market research services, having them readily available, and that clients have an influence on agencies' research functions.

Introduction  
In the advertising industry, advertising agencies place a tremendous amount of effort on obtaining and maintaining a positive agency-client relationship by providing the services required by clients. A solid agency-client relationship is necessary for a client to effectively promote its product, and for an advertising agency to keep its clients. An understanding of this relationship is vital as an unproductive and negative relationship will likely lead to the termination of the agency's services, which can result in major outlays of time, money and effort for the client. This can involve delays in implementing new campaigns, time spent on the process of selecting a new agency, and the development of rapport, trust and confidence in the new agency (Quinn 1978; Newsome 1980; Cook 1988; Weibacher 1991). It has been suggested that less than half of client-agency relationships last longer than five years (Beard 1999), and the failure to build a long standing relationship can result in "significant costs and anguish", with the process of switching agencies and developing a new partnership taking up to two years (Michell and Sanders
Thus, the changing of advertising agencies can have a negative impact on the firm (Michell 1984; Cagley and Roberts 1984; Buchanan and Michell 1991; Mathur and Mathur 1996; Waller 2004), so it is vital to maintain a positive agency-client relationship.

An important issue in agency-client relationships is the "people factor", particularly in the area of agency selection and evaluation (Cagley and Roberts 1984; Cagley 1986; Wackman, Salmon and Salmon 1986; Harvey and Rupert 1988; Verbeke 1988; Marshall and Na 1994; Mathur and Mathur 1996; and Farn and Waller 1999a). This is understandable as advertising agencies provide services in "the creation, production and/or placement of the communication message" (Belch and Belch 2004), and services are highly dependent on the people who provide the service (Parasuraman and Zeithaml 1983). Further, when it comes to measuring agency performance, Laco (1988) identified five key performance measures: contribution to the achievement of client marketing objectives; contribution to the standing of client product(s), service(s) or brand(s); creative output; value for money; and service quality. He found that a study of United Kingdom clients that there was need for improved satisfaction in the areas of value for money and service delivery.

This paper aims to analyse the services offered by advertising agencies and the advertiser's influence on agency functions. It will focus on the research-related services offered by agencies, comparing both large and small agencies, to determine where these services fit in relation to other agency services, such as creative or account services.

**Background**

Most aspects of the relationship between advertising agencies and their clients have been well documented, and are reflected in the "the development of an interesting, diverse and relevant body of literature" (West and Paliwoda 1995). Waller (2004) presented the main theoretical issues regarding agency-client relationships in three main stages: (1) agency evaluation/selection; (2) relationship development and maintenance; and (3) agency review/termination. Bush and Bush (2000) provided an overview of research on agency-advertiser relationships and reported that much of the research has examined a broad range of organisational, managerial, and interpersonal factors that have the ability to disrupt the agency-advertiser relationship. Organisational issues such as agency selection, account switching, agency loyalty, and organisational buying have been closely looked at by Buchanan and Michell (1991), Cagley and Roberts (1984), Michell (1987), Michell (1988), Michell, Cetaquet and Hague (1992), Michell and Sanders (1996), Farn and Waller (1999a), Farn and Waller (1999b), and Prendergast, Shi and West (2001). Managerial issues such as relationship development, campaign development, agency performance, media planning, and advertising talent availability have been studied by Hotz, Ryans, and Shanklin (1982), Wackman, Salmon, and Salmon (1986), Verbeke (1986), and Abratt and Cowan (1999). Finally, interpersonal factors such as client role ambiguity and satisfaction in client-ad agency relationships have been closely examined by Beard (1996, 1999).

Wackman, Salmon and Salmon (1986) identify four sets of factors that can influence the success of agency-client relationships. The factors are: (1) Work product: the advertising and advertising plans; (2) Work patterns: the daily aspects of how the agency and client work together; (3) Organisational factors: including company policy, structure, and politics, and the qualifications/experience of personnel involved; and (4) Relationship factors: the "chemistry" of the relationship which includes the level of trust, respect, rapport, and comfort between agency and client personnel.
They found that relationship factors were the most highly significant predictor of a client’s satisfaction with its agency, followed by organisational factors, work pattern and work product factors. Yet, even though services are dependent on the people who provide them, the service needs of clients must still be fulfilled (Parasuraman and Zeithaml 1983), and there is growing interest in looking at advertising agency services and assessing agency service quality (Farn and Waller 1999b; Na, Marshall and Son 1999; Na, Marshall and Son 2003).

The main services provided by an advertising agency are: (1) Account services, which is the link between agency and client and managed by the account executive; (2) Marketing services, which can include the media department that analyses, selects and contracts media resources, and the research department that designs and executes research programs; and (3) Creative services, which focuses on the creation and execution of advertisements, and employs copywriters, artists and other specialists (Belch and Belch 2004). In one of the few studies using Australian data, Shao and Hill (1992) examined the advertising skills agencies could offer to their clients. In this survey of managing directors throughout Australia, the authors found there to be major shortages of copy translators, cultural advisors and language advisors. Less severe shortages were detected with copywriters, broadcast producers, broadcast directors, account executives, account supervisors and creative directors. The shortage of copywriters was particularly alarming since the creative process is at the heart of advertising operations.

The aim of this paper is to determine the services that Australian advertising agencies offer to their clients and the amount of influence the client has on these services, with a focus on the research activities.

Methodology

To determine the type of services that are provided by advertising agencies, a survey of managing directors of advertising agencies was undertaken. The sampling frame of potential respondents was obtained from industry sources and the Yellow Pages throughout Australia. A profile of agency respondents is listed in Table 1. Over one-quarter of the respondents (26.5%) had 10 or fewer clients, while 70.6% had no more than 30 clients. Most of those responding were from private companies, and the majority had been operating 20 years or less – with over one-third (38.2%) operating between 10 and 20 years. Finally, most respondents (61.8%) had net billings of less than $10 million. To determine whether there are any differences in the services offered by small and large agencies, respondents were divided into two groups based on net billings: those with billings less than $10 million (42 agencies) and those with billings greater than $10 million (22 agencies). It was assumed that the larger agencies would offer more services than the smaller agencies. (Note: four agencies did not state their billings and so they were deleted from further analysis.)

Results

Service offered

The respondents were presented with a list of services that are typically offered by advertising agencies, and were asked to indicate which of those services their agencies regularly offered to their clients. The list of 18 services included a number of creative, management and research activities that agencies can offer to their clients (see Table 2, which ranks the services based on the number of total responses). The top three services offered by over 90% of the agencies were creative activities: copywriting, creativity, and graphic design services. Other highly popular services included assistance with marketing strategy, commercial production, internet development, photog-
raphy, direct-response marketing, and print coordination. As for research services, marketing research was offered by 70% of the agency respondents. Other services offered were attitudinal analysis of consumers (61%) and pre-testing promotions (59%), followed by demographic analysis of target market (53%), psychographic analysis of target market (52%) and post-testing promotions (50%).

Comparing the agencies based on size of billings (less than or greater than $10 million) to determine whether there are any differences in the services offered by small and large agencies, a z-test was used. Table 2 shows that eight services had significant differences between the two groups: production of commercials; direct-response marketing; media planning; attitudinal analysis of consumers; pre-testing promotions; demographic analysis of target market; psychographic analysis of target market; and post-testing promotions. With each of these, as predicted, the larger agencies were more likely to offer the services than the smaller agencies, and five of the eight services were research services.

**Availability of advertising personnel**

Respondents were then presented with a list of 24 positions in an advertising agency, and to gain insight into the skill base

---

**Table 1: Agency Profile**

<table>
<thead>
<tr>
<th>Number of active clients this agency currently serves</th>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 to 10</td>
<td>18</td>
<td>26.5%</td>
</tr>
<tr>
<td>11 to 20</td>
<td>11</td>
<td>16.2%</td>
</tr>
<tr>
<td>21 to 30</td>
<td>19</td>
<td>27.9%</td>
</tr>
<tr>
<td>31 to 40</td>
<td>6</td>
<td>8.6%</td>
</tr>
<tr>
<td>41 to 50</td>
<td>3</td>
<td>4.4%</td>
</tr>
<tr>
<td>more than 50</td>
<td>10</td>
<td>14.7%</td>
</tr>
<tr>
<td>missing</td>
<td>1</td>
<td>1.5%</td>
</tr>
<tr>
<td>Total:</td>
<td>68</td>
<td>100%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ownership status of this agency</th>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Private</td>
<td>55</td>
<td>80.5%</td>
</tr>
<tr>
<td>Public</td>
<td>13</td>
<td>19.1%</td>
</tr>
<tr>
<td>Total:</td>
<td>68</td>
<td>100%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of years the agency been in operation</th>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than 5 years</td>
<td>11</td>
<td>16.2%</td>
</tr>
<tr>
<td>5 to 9 years</td>
<td>12</td>
<td>17.9%</td>
</tr>
<tr>
<td>10 to 20 years</td>
<td>26</td>
<td>38.2%</td>
</tr>
<tr>
<td>21 to 30 years</td>
<td>10</td>
<td>14.7%</td>
</tr>
<tr>
<td>31 to 40 years</td>
<td>2</td>
<td>2.9%</td>
</tr>
<tr>
<td>more than 40 years</td>
<td>7</td>
<td>10.3%</td>
</tr>
<tr>
<td>Total:</td>
<td>68</td>
<td>100%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>The agency's total NET billing in the last year*</th>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than $10 million</td>
<td>42</td>
<td>61.8%</td>
</tr>
<tr>
<td>$10.1 million to $20 million</td>
<td>5</td>
<td>7.4%</td>
</tr>
<tr>
<td>$20.1 million to $30 million</td>
<td>2</td>
<td>2.9%</td>
</tr>
<tr>
<td>$30.1 million to $40 million</td>
<td>2</td>
<td>2.9%</td>
</tr>
<tr>
<td>$40.1 million to $50 million</td>
<td>1</td>
<td>1.5%</td>
</tr>
<tr>
<td>more than $50 million</td>
<td>12</td>
<td>17.6%</td>
</tr>
<tr>
<td>missing</td>
<td>4</td>
<td>5.9%</td>
</tr>
<tr>
<td>Total:</td>
<td>68</td>
<td>100%</td>
</tr>
</tbody>
</table>
Table 2: Services offered to clients

<table>
<thead>
<tr>
<th>Service</th>
<th>Total (64 responses)</th>
<th>Less than $10m</th>
<th>More than $10m</th>
<th>Z-test (sig)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Copy writing</td>
<td>62</td>
<td>40</td>
<td>22</td>
<td>-1.040 (.298)</td>
</tr>
<tr>
<td>Creativity</td>
<td>60</td>
<td>39</td>
<td>21</td>
<td>-.408 (.684)</td>
</tr>
<tr>
<td>Design services (graphic)</td>
<td>58</td>
<td>38</td>
<td>20</td>
<td>-.056 (.955)</td>
</tr>
<tr>
<td>Marketing strategy</td>
<td>56</td>
<td>37</td>
<td>19</td>
<td>-.284 (.777)</td>
</tr>
<tr>
<td>Production of commercials</td>
<td>52</td>
<td>31</td>
<td>21</td>
<td>-2.107 (.035)*</td>
</tr>
<tr>
<td>Internet development</td>
<td>50</td>
<td>34</td>
<td>16</td>
<td>.756 (.450)</td>
</tr>
<tr>
<td>Photography</td>
<td>48</td>
<td>32</td>
<td>16</td>
<td>.304 (.761)</td>
</tr>
<tr>
<td>Direct-response marketing</td>
<td>47</td>
<td>26</td>
<td>21</td>
<td>-2.886 (.004)*</td>
</tr>
<tr>
<td>Print coordination</td>
<td>45</td>
<td>28</td>
<td>17</td>
<td>-.892 (.378)</td>
</tr>
<tr>
<td>Marketing research</td>
<td>45</td>
<td>27</td>
<td>18</td>
<td>-1.458 (.145)</td>
</tr>
<tr>
<td>Media planning</td>
<td>44</td>
<td>26</td>
<td>19</td>
<td>-2.200 (.028)*</td>
</tr>
<tr>
<td>Attitudinal analysis of consumers</td>
<td>39</td>
<td>19</td>
<td>20</td>
<td>-3.557 (.000)*</td>
</tr>
<tr>
<td>Pre-testing promotions</td>
<td>38</td>
<td>21</td>
<td>17</td>
<td>-2.110 (.035)*</td>
</tr>
<tr>
<td>Database development</td>
<td>36</td>
<td>22</td>
<td>14</td>
<td>-.862 (.389)</td>
</tr>
<tr>
<td>Demographic analysis of target market</td>
<td>34</td>
<td>17</td>
<td>17</td>
<td>-2.802 (.005)*</td>
</tr>
<tr>
<td>Psychographic analysis of target market</td>
<td>33</td>
<td>17</td>
<td>16</td>
<td>-2.452 (.014)*</td>
</tr>
<tr>
<td>Post-testing promotions</td>
<td>32</td>
<td>17</td>
<td>15</td>
<td>-2.105 (.035)*</td>
</tr>
<tr>
<td>Database management</td>
<td>29</td>
<td>18</td>
<td>11</td>
<td>-.545 (.586)</td>
</tr>
</tbody>
</table>

* p < 0.05

offered by agencies they were asked to indicate the availability of the talents offered by their agencies. They could indicate whether the position was readily, usually, sometimes, rarely or never available. Table 3 shows the rankings of each advertising personnel position according to their degree of availability based on the number of "readily available" positions. The positions that were most readily available were creative services: layout artists, copywriters, and art directors. The research-based positions (marketing analyst and marketing researchers) were not as available as many of the other positions listed and were ranked 15th and 16th respectively. This indicates that agencies prefer offering the creative and account services to their clients, and that some research skills are left for market research specialists.

Client influence on functions

Respondents were then asked to indicate the extent of influence clients have on their agency's functions. A list of 13 services was presented and respondents indicated the level of influence by using the following scale: no influence, minor influence, major influence, total influence, and non-applicable. The results ranked by client influence, based on the number indicating a "major influence", is summarised in Table 4.

In all cases, respondents indicated that there was some level of influence from their clients. This finding was not surprising since agencies provide the services for their clients. Interestingly, clients most often preferred to be involved in marketing strategy and marketing research activities – followed by copywriting, design services, and media planning. Marketing strategy and marketing research are functions that require intense interaction between agencies and their clients since they often work hand in hand. After all, the sole purpose of marketing research
## Table 3: Availability of Advertising Personnel

<table>
<thead>
<tr>
<th>Position</th>
<th>Total</th>
<th>&lt; $10 million</th>
<th>&gt; $10 million</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Readily Available</td>
<td>Usually Available</td>
</tr>
<tr>
<td>Layout artists</td>
<td>41</td>
<td>41</td>
<td>21</td>
</tr>
<tr>
<td>Copywriters</td>
<td>41</td>
<td>41</td>
<td>20</td>
</tr>
<tr>
<td>Art directors</td>
<td>38</td>
<td>38</td>
<td>31</td>
</tr>
<tr>
<td>Account coordinators</td>
<td>34</td>
<td>34</td>
<td>27</td>
</tr>
<tr>
<td>Media planners</td>
<td>37</td>
<td>37</td>
<td>27</td>
</tr>
<tr>
<td>Traffic personnel</td>
<td>37</td>
<td>37</td>
<td>17</td>
</tr>
<tr>
<td>Art directors</td>
<td>37</td>
<td>37</td>
<td>18</td>
</tr>
<tr>
<td>Creative directors</td>
<td>37</td>
<td>37</td>
<td>17</td>
</tr>
<tr>
<td>Mechanical artists</td>
<td>37</td>
<td>37</td>
<td>25</td>
</tr>
<tr>
<td>Account executives</td>
<td>37</td>
<td>37</td>
<td>25</td>
</tr>
<tr>
<td>Media buyers</td>
<td>38</td>
<td>38</td>
<td>16</td>
</tr>
<tr>
<td>Direct marketing specialist</td>
<td>38</td>
<td>38</td>
<td>15</td>
</tr>
<tr>
<td>Account supervisors</td>
<td>38</td>
<td>38</td>
<td>19</td>
</tr>
<tr>
<td>Marketing planners</td>
<td>38</td>
<td>38</td>
<td>11</td>
</tr>
<tr>
<td>Marketing analysts</td>
<td>36</td>
<td>36</td>
<td>11</td>
</tr>
<tr>
<td>Marketing researchers</td>
<td>36</td>
<td>36</td>
<td>10</td>
</tr>
<tr>
<td>Advertising managers</td>
<td>36</td>
<td>36</td>
<td>10</td>
</tr>
<tr>
<td>Database developers</td>
<td>36</td>
<td>36</td>
<td>10</td>
</tr>
<tr>
<td>Broadcast producers</td>
<td>35</td>
<td>35</td>
<td>13</td>
</tr>
<tr>
<td>Broadcast directors</td>
<td>35</td>
<td>35</td>
<td>18</td>
</tr>
<tr>
<td>Legal advisors</td>
<td>35</td>
<td>35</td>
<td>17</td>
</tr>
<tr>
<td>Brand managers</td>
<td>35</td>
<td>35</td>
<td>17</td>
</tr>
<tr>
<td>Cultural advisor</td>
<td>35</td>
<td>35</td>
<td>10</td>
</tr>
<tr>
<td>Copy translators</td>
<td>35</td>
<td>35</td>
<td>10</td>
</tr>
</tbody>
</table>

16 Australasian Journal of Market & Social Research | Volume 13, Number 1 | June 2005
Table 4: Extent clients influence the agency's functions

<table>
<thead>
<tr>
<th>Service</th>
<th>No Influence</th>
<th>Minor Influence</th>
<th>Major Influence</th>
<th>Total Influence</th>
<th>Not Applicable</th>
<th>No of Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marketing strategy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;$10 million</td>
<td>1</td>
<td>7</td>
<td>28</td>
<td>1</td>
<td>1</td>
<td>38</td>
</tr>
<tr>
<td>&gt;$10 million</td>
<td>3</td>
<td>3</td>
<td>15</td>
<td>1</td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td>Marketing research</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;$10 million</td>
<td>3</td>
<td>5</td>
<td>21</td>
<td>2</td>
<td>4</td>
<td>35</td>
</tr>
<tr>
<td>&gt;$10 million</td>
<td>1</td>
<td>3</td>
<td>13</td>
<td>1</td>
<td>2</td>
<td>20</td>
</tr>
<tr>
<td>Design services</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;$10 million</td>
<td>3</td>
<td>15</td>
<td>19</td>
<td>1</td>
<td>1</td>
<td>39</td>
</tr>
<tr>
<td>&gt;$10 million</td>
<td>2</td>
<td>8</td>
<td>9</td>
<td>2</td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td>Copy writing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;$10 million</td>
<td>2</td>
<td>18</td>
<td>17</td>
<td>1</td>
<td>1</td>
<td>39</td>
</tr>
<tr>
<td>&gt;$10 million</td>
<td>2</td>
<td>8</td>
<td>10</td>
<td>2</td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td>Media planning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;$10 million</td>
<td>4</td>
<td>10</td>
<td>15</td>
<td>1</td>
<td>5</td>
<td>35</td>
</tr>
<tr>
<td>&gt;$10 million</td>
<td>1</td>
<td>12</td>
<td>8</td>
<td>1</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Psychographic analysis of target market</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;$10 million</td>
<td>6</td>
<td>4</td>
<td>14</td>
<td>7</td>
<td>1</td>
<td>31</td>
</tr>
<tr>
<td>&gt;$10 million</td>
<td>1</td>
<td>8</td>
<td>8</td>
<td>3</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Demographic analysis of target market</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;$10 million</td>
<td>2</td>
<td>8</td>
<td>13</td>
<td>7</td>
<td>2</td>
<td>30</td>
</tr>
<tr>
<td>&gt;$10 million</td>
<td>1</td>
<td>6</td>
<td>10</td>
<td>19</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Direct-response marketing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;$10 million</td>
<td>5</td>
<td>13</td>
<td>12</td>
<td>3</td>
<td>3</td>
<td>34</td>
</tr>
<tr>
<td>&gt;$10 million</td>
<td>2</td>
<td>11</td>
<td>6</td>
<td>1</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Attitudinal analysis of consumers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;$10 million</td>
<td>4</td>
<td>11</td>
<td>11</td>
<td>5</td>
<td>3</td>
<td>31</td>
</tr>
<tr>
<td>&gt;$10 million</td>
<td>4</td>
<td>9</td>
<td>8</td>
<td>3</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Internet advertising strategy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;$10 million</td>
<td>2</td>
<td>13</td>
<td>11</td>
<td>1</td>
<td>1</td>
<td>28</td>
</tr>
<tr>
<td>&gt;$10 million</td>
<td>2</td>
<td>10</td>
<td>6</td>
<td>1</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>Production of commercials</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;$10 million</td>
<td>7</td>
<td>15</td>
<td>8</td>
<td>1</td>
<td>3</td>
<td>34</td>
</tr>
<tr>
<td>&gt;$10 million</td>
<td>2</td>
<td>11</td>
<td>6</td>
<td>1</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Pre-testing promotions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;$10 million</td>
<td>5</td>
<td>8</td>
<td>7</td>
<td>8</td>
<td>4</td>
<td>29</td>
</tr>
<tr>
<td>&gt;$10 million</td>
<td>1</td>
<td>6</td>
<td>9</td>
<td>1</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Post-testing promotions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;$10 million</td>
<td>7</td>
<td>8</td>
<td>7</td>
<td>7</td>
<td>4</td>
<td>30</td>
</tr>
<tr>
<td>&gt;$10 million</td>
<td>1</td>
<td>7</td>
<td>8</td>
<td>1</td>
<td>20</td>
<td></td>
</tr>
</tbody>
</table>

is to help decision-makers make informed decisions, and it is marketing research that often guides advertisers' marketing strategies. Since marketing strategies impact advertisers' performance for several years, there is little room for error. So advertisers prefer to be closely involved in the process. Clients indicated that they prefer to be less involved in the actual implementation of advertising activities, which include pre- and post-testing of promotions, as well as production of commercials. This may be attributed to the fact that more specialised skills are needed for these activities and the clients, therefore, prefer to leave them in the control of advertising experts.
Discussion
The objective of this paper was to enhance our understanding of the services offered by advertising agencies and the advertisers' influence on agency functions, with a particular focus on the research-related services offered by the agencies. The survey was completed by 68 managing directors from advertising agencies across Australia. As for the services offered to clients, the main services offered by over 90% of the agencies were creative ones: copywriting, creativity, and graphic design services; with marketing research services being offered by 70% of the respondents. Creative positions (layout artists, copywriters, and art directors) - which are typically the heart of advertising agency activities - were the positions that the agencies had readily available while the research-based positions, marketing analyst and marketing researchers ranked much lower on the list. This indicates that agencies prefer offering the creative and account services to clients, and that while some research services are offered and some have them readily available, most research skills are left to the abilities of market research specialists, such as market research companies.

The extent clients influence agency operations indicate that they indeed play an important role in the process to ensure that they obtain the results they are seeking from their promotion efforts. From the results, marketing strategy and marketing research are areas where clients have a major influence. There is less influence in research activities that are implemented like pre-testing, post-testing promotions, and production of commercials, where clients leave the more specialised activities to the experts.

Conclusion
In general, it appears that while creative and account management services are extremely important for advertising agencies in Australia, a full line of marketing research services are still offered by a number of agencies. In particular the more general marketing research services are offered and are readily available to their clients, yet they are also the services that have a major influence from clients. Further research should be undertaken to determine the use of market research by advertising agencies and how this is used in campaign creation and management. Analysing the link between market research services and agency-client relationship would also be worthwhile in the future to aid our understanding of its importance in the advertising process. As clients seem to have an influence on agency functions, it is important to understand what clients want so that agencies can better serve them in the future.
REFERENCES


Return on marketing investment:
A practical end to end approach for
linking marketing investments
to shareholder value

Liane Ringham

INSIDE STORY Knowledge Management Pty Ltd
Level 5/ 2 Barrack St
Sydney, NSW 2000 Australia
lianer@insidestory.com.au

Abstract
Within today's organisation marketing is under immense pressure to demonstrate its value to the firm. This requires methods used to link marketing activity to financial performance. We use the recently developed customer equity approach to marketing activity evaluation and link this to a shareholder value model, the result of which we name the Marketing Investment to Shareholder Value (MISV) model. The approach has the potential to demonstrate to CEOs, CFOs, and shareholders, the financial value of marketing activities such as above the line advertising and other activities that are currently considered hard to value. From our practical experience in working with tracking brands and ads and working with Australia's largest organisations, this paper identifies three challenges to be met if an organisation is to successfully implement Marketing ROI.

Introduction
The current challenge for marketing is for its outputs to be fully valued by the organisation. Without a language in common with analysts, CFOs, CEOs and the board, the risk is that marketing's inputs will not be valued. Many argue that this has resulted in a shift to short term tactical activities and the erosion of many marketing budgets. At the same time, it is clear that there is increased recognition of the value of brands and reputation, though relatively little published about ways of linking marketing activities and investments with these outcomes. A recent special section in an issue of the Journal of Marketing (Lehmann 2004) on Linking Marketing to Financial Performance and Firm Value provides a collection of papers on the current state of knowledge in this area. Also, marketers are becoming ever more interested in measuring the impact of marketing activity and investment on financial outcomes through appropriate marketing metrics (e.g. Australian Marketing Institute 2004; Swinburn 2004).

We believe that the concept of customer equity is a key means to achieve this outcome and is particularly relevant when revenues cannot be easily linked to marketing investments as with brand advertising and other investments that have short term and longer term impacts. Further, linking the financial estimates generated from customer equity valuations (which are in accounting terms essentially incremental gross margins) can be usefully input into a standard shareholder valuation model which, with some adjustments, can produce a set of managerial accounts that demonstrate the return on investment in marketing activities.

In the current paper, we provide a basic overview of the major approaches to value creation available in the literature and show how the customer equity and
shareholder value can be linked. We also identify how we have been able to link invests (weekly target audience rating points and weekly spend) to customer equity. In doing so, we demonstrate how marketing can potentially talk in the language of senior management and CFOs and demonstrate the value of its activities. We also share some of the key challenges in implementing marketing return on investment with a view to furthering this emergent field.

Customer equity – the missing link
In our review, the term customer equity first appeared in a published paper in 1996. At that time, it was used in a Harvard Business Review article written by Blattberg and Delghton (1996). Since that time, the concept of customer equity (or the net present value of current and potential future customers) has become a relatively hot topic (Dorsch and Carlson 1996). Indeed, two major books have appeared in the business literature, both of which use customer equity as the linchpin for the business strategy development and evaluation. These are Rust, Zeithaml & Lemon (2000) and Blattberg, Getz & Thomas (2001).

The focus on customer equity has as one aim, the redirection of primary management attention away from the product and towards the consumer. This means less attention to product-based ideas such as market share and more attention paid to consumer-based outcomes such as customer equity share, which are more relevant to understanding share of wallet. Another aim is the provision of a mechanism for examination of the return on investment (ROI) in various marketing activities including making allocation decisions for investments in different marketing activities. This has been an area of great interest among both marketers as well as academics for many years (e.g. see the work on Economic Value Added, Ehrbar 1998).

Linking to shareholder value
A whole stream of academic and practitioner thinking in the financial and economic arena has identified that the key goal of a business is to generate shareholder value (McKinsey et al 2000) and the positive and negative cash flow streams discounted over time should be the focus of any investment of the organisation (Srivastava, et al 1998). This has been discussed very recently as a priority area for research by Rust, et al (2004) in their survey of what we already know and what needs to be examined in the field of measuring marketing productivity. We use the customer equity idea as the middle part of a model that ends with shareholder value.

The Rust et al approach to customer equity
Rust et al (2000) and Rust, Lemon & Zeithaml (2004) link total customer equity econometrically to a variety of brand equity, customer satisfaction and other drivers. The approach involves the estimation of a Markov switching model using multinomial logit modelling and provides scope for linking customer choice drivers (e.g. brand awareness, brand image perceptions, customer preferences) to customer equity. The advantage of the switching model is that it enables marketers to value the impact of investments beyond the immediate time frame, as many strategic marketing efforts have a short term and a longer term impact and often the longer term impact is potentially more valuable.

This is quite different from the Blattberg et al (2001) approach, which is basically spreadsheet-based where, for a sample of customers, the firm needs to calculate the initial rate of acquisition, rate of retention, and the rate of add-on sales. They also need the margins for these three activities along with the expenditure for the three activities, all at the per customer level. These are then combined to calculate measures for what they call acquisition, retention and add-on equity. Finally, the per-customer amounts are
aggregated up to the firm level as a basis for making strategic decisions about whether to invest in acquisition, retention or cross selling strategies and the appropriate use of above the line advertising and below the line investments.

There is no real scope in their approach to analyse the impact of attitudes and perceptions on the various customer equity measures, something that the Rust et al approach allows. It is essentially a database method, but does not take into consideration marketing strategies that seek to position the brands in the minds of consumers and to understand the value of these investments.

Getting from customer value to ROI
Customer equity is a financial estimate of the value created by the firm by an activity or series of activities that change the driver or drivers of customer equity. Once costs are overlaid on the marketing activities that drive the changes in the customer equity, Rust et al have been able to calculate return on investment which is basically the increase in the dollar value created by the activity divided by the cost of the activity which is discounted over the life of the activity's expected impact as is typical of any capital investment in accounting terms.

This is truly a breakthrough as it enables business people and researchers for the first time to say for example, that the increase of the brand on a particular image dimension (e.g. caring) created by a particular investment in, say, a brand campaign has increased customer value by, say, $2.3 million dollars. This is surely more meaningful than the current state of play that enables the researcher and the marketing director to, say, for example, that a campaign has increase image for caring from 3.2 to 3.5 on a 5-point scale and that this is statistically significant.

Challenges in implementation of the customer equity approach
For three years, we have been working with tracking data and Rust's algorithms to work out practical ways we can value the efforts of marketing. We have, for example, worked with data in over the counter (OTC) pharmaceuticals, snack foods, banking, telecommunications, and gaming. We will confine our experiences in this paper to the advertising and brand spheres though many of the lessons transfer across to the valuation of activities such as PR, service improvement, etc.

We have identified at least three challenges. The first challenge lies in identifying the linkage between specific marketing activity and the drivers (e.g. brand awareness, brand image perceptions, brand preference). We have found first hand what Broadbent (1990) also has found, that basic regression techniques are unable to model the impact of retention factors, decay and the valuation of the creative impact of any campaign. We are indebted to the insights generated by the work of Brown (1986) and Broadbent (1990) who gave us clues as to how to proceed. Establishing these linkages is the key to any valuation of marketing activities beyond the short term and establishes the timing of cash flows that are critical to the final phase of generating incremental gross margins by period (say, for the next three years) for the shareholder value model.

The second challenge is a managerial one. To implement the Rust model, access is required to several pieces of information relating to the size and timing of the cash flow streams from the investment. The specific inputs depend on the nature of the business though typically include: customer contribution margin, frequency of purchase/revenue delivery, volume of purchase (if applicable) and size of market including growth estimates. Ideally this is at a customer level and also, where relevant, the information is needed for competitors. Our experience is that this information is still virtually impossible to
get in Australia for many of Australia's largest organisations, particularly if the information is at a cross product or total brand/customer perspective. We have worked to overcome this is by making estimates based on segment membership and expert opinion where data is missing or only partial. This is not as bad as it seems as these estimates can be put into our final sensitivity analysis.

Finally the third challenge is that any estimate contains a degree of error. The final challenge lies in the estimation of the error of the estimate. Rust's article gives some clues to this and the most rigorous way is to take the choice model classifications of individual customers and follow through the customers in real life to see what they actually do. We wonder if many marketers have the willingness or time in practical terms to track through a series of actual respondent behaviours in real life for several years. In view of realities of today's market place, we believe that the most practical outcome may be to conduct a series of sensitivity analyses on the estimates provided to the shareholder value model (i.e. the incremental gross margins). This will at least enable marketers and researchers to get up and running so that over time they can build and improve their models based on real life validations with follow up studies.

**Challenges in linking marketing investments to shareholder value**

The logic of linking marketing investments to shareholder value is illustrated in Figure

![Figure 1](#)
1. This can be termed the Marketing Investment to Shareholder Value (MISV) model.

There are several adjustments required to the usual capital investment thinking when it comes to valuing a marketing investment. The first of these is that companies need a set of managerial accounts as currently Australian and International accounting standards do not allow for the valuing of marketing assets. Hence, the marketing investment must be amortised in these sets of accounts over the projected life of the investment. Then the incremental gross margins (IGMs) after amortisation can be estimated. Finally, the net present value (NPV) of the IGMs is calculated using the adjusted risk index (discussed below). This NPV is essentially a measure of Market Value Added (MVA), a concept discussed by McKinsey et al. (2000, p 59) as growing in popularity as a measure of financial performance.

Second, the risk index used or discount rate applied to the estimates of gross margin should not be the overall discount rate applied to the business (typically the weighted average cost of capital - WACC; see McKinsey, et al. 2000, chapter 10). The risk index needs to be adjusted based on:

- Barriers to entry
- Legislation
- Customer dependency
- New product acceptance
- Customer mix
- Competition

In other words, the risk index needs to take into consideration the factors that could potentially affect the probability of gross margins not being generated in full and on time. So timing is everything. We can examine the timing issue by sensitivity analysis, systematically varying the risk and/or timing of gross margins in order to better understand the financial risk involved.

What advantages are achieved in linking marketing investments to shareholder value rather than customer equity alone?

By linking to a company’s accounts structured to represent a shareholder value model, several advantages are possible:

- The Rust et al. (2000, 2004) model stops at the NPV of the customer lifetime value estimated from the model. The full MISV model takes this one more step into the realm of shareholder value.
- A clear understanding of the timing of the positive cash flows - so the business can see when value is delivered and how long they have to wait.
- Sensitivity analysis gives an understanding of the conditions under which the investment will be vulnerable.

Hence, the estimated incremental gross margins that are the output from the Rust model are the inputs to the shareholder value model at the back end of the MISV framework. This provides a direct link from marketing investments to ultimate shareholder value so that CEOs, CFOs, board members, investors/owners can understand the true value of marketing in today’s organisation.
References


Consensus versus polarisation: Measuring ordinal variability

T Paul Hutchinson
Centre for Automotive Safety Research, University of Adelaide, South Australia 5005, Australia
Phone: +61 8 8303 5997
Fax: +61 8 8232 4995
Email: paul@casr.adelaide.edu.au

Abstract
Proposals have been made recently by Blair and Lacy and by Yager concerning how variability on an ordinal scale should be defined. The present article criticises these proposals for having some characteristics appropriate when the level of measurement is nominal rather than ordinal. It also shows that the use of numerical scores can easily lead to incorrect conclusions. An alternative idea is to compare distributions by notionally taking two random observations from one and two from the other, and saying there is evidence that the first is more variable if it is the source of both the smallest and largest of the four observations. Another alternative, a graphical method called the P-P plot, shows one cumulative distribution in terms of the other. Two datasets are used as examples. One concerns consensus in attitudes to homosexuality, and the other the perceived location (left vs. right) of political parties.

Introduction
The degree of consensus, as contrasted with polarisation, is sometimes the focus of research. There may be a desire to compare different groups of people, or compare different issues, in respect of variability on an ordinal scale – perceptions of institutions or personal attitudes, for example. It is more common for the typical or average attitude to be the focus, in which case the median is a familiar statistic that is usable even though the data are expressed only on an ordinal scale. (For example, the alternative responses might be strongly agree, agree, neutral, disagree, strongly disagree.) But when consensus versus polarisation is the issue, how can an appropriate statistic be calculated? The problem is that measurement of variability typically involves subtraction of one position from another, in order to get the distance between them – this is the case with the standard deviation, for example – and we cannot know if, for example, the distance from strongly agree to agree is really the same as that from neutral to disagree.

If numbers are assigned to the ordinal categories (e.g., 1, 2, 3, 4, 5), the variance and standard deviation can be calculated. When there are many categories, this is likely to be acceptable: in studies of political polarisation, Inglehart and Klingemann (1976) and Knutsen (1999) had 10-point scales, and did not question the appropriateness of the standard deviation. When there are few categories, there is likely to be scepticism about using the standard deviation, and an example below will show this to be justified.

However, any choice of numbers is rather arbitrary, and a measure of ordinal variability may be desired. Blair and Lacy (2000) and Yager (2001) have recently written on this, but the measures considered in those papers have an unexpected, and probably undesirable, feature (discussed below). A different concept...
of ordinal variability will be suggested below, and a graphical method for presenting two sets of ordinal observations that possibly have different variabilities will be discussed. A contrast of the more extreme categories of an ordinal scale with the more central may be familiar in the context of response style, but this is outside the scope of the present paper. The chief concern throughout is with statistical description, not testing. However, some readers may wish to know that statistical tests are available that are sensitive to changes in variability as well as central value – Hutchinson (2002) gives some relevant references.

**Smallest and largest**
The purpose of this Section is to propose a simple method of comparing the spread of two distributions. Symbols F and G will be used for the cumulative distributions.

**Method of calculation**
If the scale is ordinal, we can only know that the distance from A to D is greater than the distance from B to C if both B and C are between A and D, e.g. if the order is A B C D. Suppose we randomly take two observations from F and two observations from G, and we arrange these four observations from smallest to largest. If both the largest and the smallest came from F, we have evidence that F is the more variable; if both the largest and the smallest came from G, we have evidence that G is the more variable; otherwise, we cannot say anything. We should calculate the probabilities of these two possibilities, and compare them. (Admittedly, this method only measures variability in the region where the two distributions overlap: if the centres of the two distributions are appreciably different, it is unlikely that one of them will provide both the largest and the smallest of the four observations, and a comparison will not be possible. However, this should not be considered a criticism, as if two distributions really do not overlap very much, and the scale really is only ordinal, then it is inherently impossible to compare their variabilities.

In Example 1 below, there are four ordered categories, and the two distributions are from 1991 and 1993. Consider the probability that both the largest and the smallest observations came from 1991, or, more precisely, the probability of four observations being generated that suggest the distribution in 1991 was more variable than in 1993. This is obtained by considering the following list of possibilities:

* Two 1991 observations were respectively in categories 1 and 3, and two 1993 observations were both in category 2;
* Two 1991 observations were respectively in categories 1 and 4, and two 1993 observations were both in either category 2 or category 3;
* Two 1991 observations were respectively in categories 2 and 4, and two 1993 observations were both in category 3.

With p's being the proportions for 1991 and q's being the proportions for 1993, this probability is 2p1p3q22 + 2p1p4q2+q32 + 2p2p4q32. The probability of four observations being generated that suggest the distribution in 1993 was more variable than in 1991 can be obtained similarly.

Ordered categories are usually viewed as representing ranges of some underlying continuous variable, and thus there is unobservable variation within a category. Occasionally, the categories might be viewed as real, all observations of a particular category being necessarily equal to each other. If so, then such orderings as A < B = C = D or A < B < C = D would be seen as evidence that the distribution generating A and D had a greater spread than that generating B and C. Obviously this would lead to changes in the above formulae.

**Example 1: Attitudes to homosexuality**
One of the examples used by Blair and Lacy (2000) concerned attitudes to homosexuality (on a four-point scale, data in their Table 7). They argue that 1993
showed a decline in consensus as compared with 1991. This was so whether the criterion was the variance, calculated assuming equal-interval scoring, or the $\chi^2$ statistic that they introduced. I will say that there was greater consensus in 1993 than in 1991. (In passing, it may be noted that it is uncontroversial that there was less disapproval of homosexuality in 1993 than in 1991.) At this point, it may be noted that a calculation of the variance is irrelevant - it is undesirable to need to assume certain scores for the categories, which is why Blair and Lacy proposed their $\chi^2$ statistic. There is further comment on this below, and also description of how $\chi^2$ is defined and criticism of its properties.

The proportions in the four categories in 1991 were .755, .041, .044, and .160; in 1993, these were .663, .043, .073, and .220. By way of preliminary, it may be noted that the total proportion in the two extreme categories ("always wrong" and "not wrong at all") was less in 1993 than in 1991: it was 88.3 per cent, as compared with 91.5 per cent. This cannot be said to be strong evidence of greater consensus in 1993 than in 1991 - the proportion in the two extreme categories is sensitive to where the centre of the distribution is, as well as to its variability - but it does demonstrate that this claim is not completely unreasonable.

Using the method of calculation described earlier, if we randomly take two observations from the 1993 distribution and two observations from the 1991 distribution, the probability of four observations being generated that suggest the distribution in 1993 was more variable than in 1991 is 0.0012, and the probability of four observations being generated that suggest the distribution in 1991 was more variable than in 1993 is 0.0017. Thus there is evidence for greater consensus in 1993.

Example 2: Left vs. right location of political parties

Table 12.2 of Inglehart and Sidjanski (1976) gives data on the perceived location of eleven Swiss political parties on the left-right dimension. Five categories are shown, condensed from a 0-100 scale. For example, the proportions of respondents judging the Socialists to be in categories 1 to 5 (extreme left to extreme right) were .19, .39, .34, .06, and .02, and the proportions judging the Republican Movement to be in these categories were .09, .13, .24, .19, and .35. Inglehart and Sidjanski considered that these data imply more variability in respondents' opinions of the Republican Movement. This is plainly a reasonable opinion (the proportion in category 3 is smaller and the sum of the proportions in categories 1 and 5 is larger for the Republican Movement), but it may nevertheless be noted that Inglehart and Sidjanski do not specify exactly why they think this; after all, the proportions given seem consistent with respondents perceiving the Republican Movement as simply to the right of the Socialists - successive cumulated proportions are .19, .38, .92, and .98 in the case of the Socialists, and .09, .22, .46, and .65 in the case of the Republican Movement, and the latter are consistently less than the former.

Nevertheless, the "smallest and largest" method leads to the same conclusion that Inglehart and Sidjanski reached. The probability of four observations being generated that suggest the distribution for the Socialists was the more variable is .014, and the probability of four observations being generated that suggest the distribution for the Republican Movement was the more variable is .085. Thus the conclusion is that the Republican Movement distribution has the greater spread.

Comments

If there are many categories, the expressions for the required probabilities will be quite complicated. Depending upon the software available, it might be easier to obtain the probabilities by simulation - that is, given two vectors of probabili-
ties (the p’s and the q’s), draw two random observations using one vector and another two using the other vector; test whether both the smallest and the largest observations were generated by the same vector; repeat this a million times, and count the number of instances in which the required conditions occurred.

In the form above, there is not a comparison of a single population of interest with a standard population, leading to a numerical index that describes the population of interest. Rather, there is a comparison of two populations of interest. If an index is desired, the uniform distribution (i.e., with equality probabilities of each category) could be used as the standard of comparison.

**P-P plots**

**Method**

The “smallest and largest” approach makes no assumptions beyond those of an ordinal scale. In contrast, the approach of the present Section will be to assume a model for how the proportions in the several ordered categories arose.

One approach to the issue of comparing ordinal responses in two conditions is to assume that if only the appropriate transformation could be found, then the two distributions would be seen to both be normal. It is typically difficult to rigorously justify such an assumption, and if it is incorrect, there is a theoretical possibility that conclusions are not valid. Nevertheless, in my opinion this is usually the sensible way of attempting to understand this type of dataset. (Calculation of a single statistic, even one that appears free of assumptions, does risk missing some important feature of the data. Moreover, the “smallest and largest” statistic is based upon probabilities, and may be more difficult to comprehend than distance-based measures such as standard deviation that are practicable with metric data.)

The plausibility of the assumption of normal distributions may be examined by means of a P-P plot, which is the line traced out by plotting G(x) against F(x), as x varies. (The P-P plot has been invented in a number of different fields of application, and has acquired various names: see Hutchinson (1985) for more details.) Both F(x) and G(x) are often transformed into the corresponding normal deviates (z-scores). The reason for this is as follows. Since the observations are only on an ordinal scale, we could simply define an underlying metric variable to have a standard normal distribution (mean of zero and standard deviation of unity) for one of the sets of observations. It may then be the case that the distribution of this variable for the second set of observations is approximately normal, though with some different mean and standard deviation. If this is indeed so, then plotting the normal deviates will result in a straight line, the parameters of which (intercept and slope) will bear a simple relationship to the parameters of the distribution of the second set of observations (mean and standard deviation).

The concise explanation of this runs as follows. (For readers who are not accustomed to mathematical symbols and algebraic manipulations, this will, I hope, become clear in Examples 1 and 2 below.) Let F(x) represent the cumulative normal distribution, and let z(x) be the inverse of this. The meaning of there being a transformation such that there is a normal distribution in both conditions is that for each condition j, the probability of being in category i or lower is F([x_i - m_j]/s_j), for some choice of means m_j, standard deviations s_j, and category boundaries x_i. Therefore, if we transform the observed cumulative probabilities into their z-scores, we get (x_i - m_1)/s_1 for condition 1 and (x_i - m_2)/s_2 for condition 2. These are linearly related: (x_i - m_2)/s_2 = A.([x_i - m_1]/s_1) + B, where A is s_1/s_2 and B is (m_1 - m_2)/s_2.

Expressing this in another way, if there is a change in the location of the cen-
tre of the distribution, the proportion of observations in one of the extreme categories increases, and the proportion in the other extreme decreases. That is straightforward, but is the magnitude of increase at one extreme, relative to the decrease at the other extreme, consistent with the spread of the distribution remaining unchanged? Or does it suggest an increase or a decrease of the spread? Making an assumption about the functional form of the distribution - e.g. that it is normal - permits this question to be answered.

Example 1
The results in the case of the attitudes-to-homosexuality dataset are as follows. The proportions in the four categories in 1991 and 1993 were given earlier, and are repeated in Table 1. The cumulative proportions are also in Table 1, and thus the three points on the P-P plot are (.755, .663), (.796, .706), and (.840, .779), as well as the two end-points, (.00, .00) and (1.00, 1.00). From the cumulative proportions, the normal deviates shown in the final line of Table 1 may be obtained. (The method of doing so is by employing a table of the normal distribution, being careful to start with the known F and read off z, or by using appropriate software.) Thus the points on the transformed P-P plot are (.69, .42), (.83, .54), and (1.00, .77) (and the two end points disappear to ∞ and +∞, and cannot be plotted). These are shown in Figure 1. The range between the first and third cut-offs was 0.31 (= 1.00 - 0.69) of what the standard deviation was in 1991, and was also 0.35 (= 0.77 - 0.42) of what the standard deviation was in 1993. Hence the standard deviation in 1993 must have been a little smaller than in 1991 - there was less variability of attitudes (greater consensus) in 1993 than in 1991. Of course, the shift to less disapproval of homosexuality in 1993 is also reflected in these figures: the corresponding cumulative proportions were less in 1993 than in 1991, and so the normal deviates were less.

It is practicable to make some check of whether the conclusions are robust to changing the assumption of underlying normal distributions. For example, logistic distributions might be assumed instead. The appropriate transformation of the cumulative proportions, instead of the normal deviate, is then the logit. (Unlike the normal deviate z, there is an elementary expression for the logit of a probability p: it is ln(p/(1-p)), where ln denotes the natural logarithm.) For this dataset, the same conclusion about greater consensus in 1993 did follow.

Table 1. The steps in the construction of the P-P plot for Example 1, for which there were four ordered categories.

<table>
<thead>
<tr>
<th>Proportions</th>
<th>1991</th>
<th>1993</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>.755</td>
<td>.663</td>
</tr>
<tr>
<td></td>
<td>.041</td>
<td>.043</td>
</tr>
<tr>
<td></td>
<td>.044</td>
<td>.073</td>
</tr>
<tr>
<td></td>
<td>.160</td>
<td>.220</td>
</tr>
<tr>
<td>Cumulative proportions</td>
<td>.755</td>
<td>.663</td>
</tr>
<tr>
<td></td>
<td>.796</td>
<td>.706</td>
</tr>
<tr>
<td></td>
<td>.840</td>
<td>.779</td>
</tr>
<tr>
<td>Standard normal deviates (z's)</td>
<td>.69</td>
<td>.42</td>
</tr>
<tr>
<td></td>
<td>.83</td>
<td>.54</td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td>.77</td>
</tr>
</tbody>
</table>

Figure 1. P-P plot for Example 1, the dataset from Blair and Lacy (2000).
Example 2
For the Inglehart and Sidjanski (1976) dataset, a transformed P-P plot leads to the same conclusion reached by those authors and using the “smallest and largest” method above. The proportions in the five categories for the two parties are given in Table 2. The cumulative proportions are also there, and the four points on the P-P plot are (.19, .09), (.58, .22), (.92, .46), (.98, .65), as well as the two end-points. From the cumulative proportions, the normal deviates shown in the final line of Table 2 may be obtained, as for Example 1. Thus the points on the transformed P-P plot are: (-.88, -1.34), (.20, -.77), (.41, -.10), and (.05, .39), and these are shown in Figure 2. Thus 2.93 (= 2.05 - .88) times the standard deviation for the Socialists is the same as 1.73 (= 0.39 - -1.34) times the standard deviation for the Republican Movement. Therefore, the standard deviation for the Republican Movement must be appreciably greater than that for the Socialists.

Figures 1 and 2 appear to show very close relationships between the two (transformed) probabilities, but a warning should be given. The first part of this is that one cumulative probability is being plotted against another: the very method of constructing the plot ensures that as one increases, the other increases also. The second part is that concerning any curvature of the relationship, the eye probably sees this more strongly than is statistically justified. With four points, for example, there are three line segments that join them; if these get steadily steeper or steadily less steep, there will be quite a strong visual impression of curvature; and yet the probability of three quantities being ordered 1, 2, 3 or 3, 2, 1 is as high as one third.

Criticisms of other concepts
Mathematical approaches
We might consider developing a “tail heaviness” concept as follows. We say that F has heavier tails than G if there exists some y such that F(x) exceeds G(x) for x < y, and 1-F(x) exceeds 1-G(x) for x > y. That is, to the left of y, the left hand tail of F is heavier than the left hand tail of G, and to the right of y, the right hand tail of F is heavier than the right hand tail of G. However, this criterion is too strict for most purposes, as it is easy to think of pairs of distributions such that one seems clearly to have greater ordinal variability yet does not satisfy this criterion. For example, suppose there are four ordered categories, having probabilities .3, .2, .3, .2 for one distribution and .0, .2, .5, .3 for the other: the first seems intuitively to have greater variability, and does have a heavier left-hand tail, but the second has the heavier right-hand tail. And the tail-heaviness concept is not helpful with the attitudes to
homosexuality dataset because of the high proportion in the lowest category, in both years.

Several approaches that attempt to formalise the idea of one distribution being more dispersed than another, and which have a rather mathematical or axiomatic flavour, as tail heavierness does, are discussed by Sheaked (1985), Oja (1985), and Marzec and Marzec (1999). However, they are not very helpful in the present context, for one or more of the following reasons. (a) As in the previous paragraph, some are too strict, and in too many cases do not positively identify either of the distributions as the more variable. (b) Some are designed for a variable that in principle might take any value, rather than being limited to a few categories. (c) Some assume the variate is metric rather than ordinal.

Use of scores to calculate standard deviation
A straightforward approach is to assign scores to the categories, and then calculate the standard deviation. That the scores are arbitrary and thus this approach is undesirable has already been mentioned. The idea, introduced above, of normal distributions underlying the observed proportions, may be used to reinforce this by demonstrating that basing standard deviations on assumed scores can give rise to misleading conclusions.

To keep this demonstration simple, let an underlying continuous variate have only three categories, with the boundaries between them being at 1 and 2 on the continuous scale. Suppose that in one situation there is a normal distribution of attitudes with mean 0 and standard deviation 1. The boundaries are thus at z's of 1 and 2, and the proportions in the three categories are .34, .14, .52. Suppose also that in another situation there is a normal distribution with mean 1 and standard deviation 0.5 (i.e. less variability accompanying a higher mean). The boundaries are now at z's of 0 and 2, and the proportions in the three categories are .50, .48, .02. Thus this example resembles the attitudes-to-homosexuality dataset, in that the observations are skewed down into the lowest category. Now let us see the results of calculations that assume equal-interval scores (0, 0, 1, 2). Proportions .34, .14, .52 imply that the standard deviation on this scale is 0.43, whereas proportions .50, .48, .02 imply that it is 0.54. The apparently greater variability in the second situation contradicts the model that generated the proportions.

Blair and Lacy's $d^2$
The idea behind $d^2$, the measure of concentration proposed by Blair and Lacy (2000), and related measures, is as follows. The maximum degree of dispersion occurs when half the observations fall in the lowest category and half fall in the highest category; calculate an index describing how much more concentrated the observed dataset is than this. Blair and Lacy's equation (1) is $d^2 = \sum_{i=1}^{k} (F_i - 0.5)^2$, where $F_i$ is the cumulative proportion for the $i$th category, and there are $k$ categories. Independently, Yager (2001, Section 3) proposed $\sum_{i=1}^{k} F_i (1 - F_i)$; this is clearly equivalent, as $F_i (1 - F_i) = 0.25 - (F_i - 0.5)^2$.

Consider the simple case of just three ordered categories, and let $p_i$ be the proportion of cases in category $i$. The values of $d^2$ are shown below for several examples each having $p_3 = 0.2$:

<table>
<thead>
<tr>
<th>$p_1$</th>
<th>$p_2$</th>
<th>$p_3$</th>
<th>$d^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2</td>
<td>0.6</td>
<td>0.2</td>
<td>0.18</td>
</tr>
<tr>
<td>0.3</td>
<td>0.5</td>
<td>0.2</td>
<td>0.13</td>
</tr>
<tr>
<td>0.4</td>
<td>0.4</td>
<td>0.2</td>
<td>0.10</td>
</tr>
<tr>
<td>0.5</td>
<td>0.3</td>
<td>0.2</td>
<td>0.09</td>
</tr>
<tr>
<td>0.6</td>
<td>0.2</td>
<td>0.2</td>
<td>0.10</td>
</tr>
<tr>
<td>0.7</td>
<td>0.1</td>
<td>0.2</td>
<td>0.13</td>
</tr>
</tbody>
</table>

As the proportion of observations in category 1 increases at the expense of the proportion in category 2, so $d^2$ declines, as would be expected of a measure of concentration (absence of dispersion). But this only continues until $p_1$ reaches 0.5, after which $d^2$ starts to increase with
increasing p1. What seems to be happening is that $c^2$ is sensitive to a nominal-level concept of concentration, as well as to an ordinal-level concept: when a lot of observations crowd together in one category, even if this is an extreme category, $c^2$ increases.

This feature of $c^2$ may not always be appropriate; but users do need to be aware of it.

Discussion: First, catch your concept

This heading is intended to mean that the researcher should pay a great deal of attention to deciding upon what is really the best way of refining the vague phrase "ordinal variability" in his or her particular research context. For example, arguing about whether attitudes to homosexuality really exhibited greater consensus in 1991 or in 1993 is less productive than is refining what is meant by "consensus". The researcher cannot afford to be lazy and assume that some off-the-shelf index captures the concept that is required. In one context, Blair and Lacy's $c^2$ might be appropriate; in a second, the standard deviation based on equal-interval scoring; in a third, the "smallest and largest" idea.

Because Blair and Lacy approached the dataset in this way, I have accepted that it is appropriate to discuss the consensus or variability in attitudes to homosexuality. But there may be objections of a general statistical nature. The proportions in the very lowest category were about 70 per cent in both 1981 and 1993. It could be said that consensus and variability involve the notion of a distance from some point well below the middle of the distribution to some point well above; and with this dataset, we do not know where the middle of the distribution is, let alone a suitable point below that. If we took this view, we would presumably reject notions of consensus and variability that look at both the upper tail and the lower tail of a distribution, and instead look for one that is sensitive to the upper part of the distribution only. Further, there may be objections that are specific to the particular topic. Concerning the attitudes to homosexuality example, we might say this is a politicised issue, and whether or not people care about it is as important as the degree of their disapproval. Thus there may be three groups, which might be labelled disapproving, liberal, and uninterested. In this case, it may be that attitude is not an ordinal variable, and it is unfruitful to discuss the average and the variability. Instead, there would be issues concerning the proportion of people interested, and the proportion of interested people who are liberal as contrasted with disapproving. Turning to the perceptions of political parties example, we might propose that there is an objectively correct location of each party, and that the accuracy of respondents' perceptions is a useful concept.

This amounts to a plea to think about what you want to do with your data. This might seem to be a mere platitude. However, there has been an important change from a generation ago: the computing power available is now much greater. Previously, the driving force behind choice of a particular statistic might have been that it was easy to calculate, or that its distribution was normal, even if conceptually it was not quite right. But nowadays, the technicalities need be of no concern, and a statistic can be selected because it really does capture the appropriate idea; if it so happens that none has yet derived mathematically its null distribution, we simply resort to a randomisation strategy for statistical testing. That is, given the total list of observations, they are randomly assigned by the computer program to two groups of the appropriate sizes, and the statistic of interest is then calculated for this simulated dataset. This is repeated a thousand times (say), and thus the distribution of the statistic is obtained under the hypothesis that group membership had no effect on the observation. The observed value can now be compared with this null distribution, and a statement made that it does, or does not, lie in the most extreme 5% of the distribution.
References


Hutchinson, T.P. (1956). "Presenting one probability distribution as a function of another – some applications". American Journal of Mathematical and Management Sciences, 5, 103-123.


Understanding mail survey non-contact: an examination of misaddressed survey invitation returns

Benjamin Healey, Massey University
Philip Gandell, Massey University

Correspondence should be addressed to:
Benjamin Healey
Department of Marketing
Massey University
Private Bag 11-222
Palmerston North
New Zealand
Email: b.j.healey@massey.ac.nz

Abstract
Mail survey sampling frames inevitably contain out-of-date address information for some potential respondents. In addition to the cost of undelivered survey packages, the misaddressing that results from out-of-date data can lead to inaccuracies in reported response rates and bias in the population estimates inferred from survey results. To date, little research has been undertaken to investigate the extent of mail survey non-contact, or whether certain sub-populations are more likely to be prone to it than others. Furthermore, it is difficult for researchers to accurately assess non-contact in surveys because 'Return to Sender' receipts do not fully represent the number of invitations that did not reach their intended recipients.

This study examined non-contact return rates by sending deliberately misaddressed survey invitations to a sample of 1400 New Zealand households. The effectiveness of an envelope message encouraging recipients to return misaddressed items to the sender, aimed at improving overall 'Return to Sender' receipts, was tested as part of the study. Additionally, household composition was examined to identify demographic differentiators of misaddressed survey returning behaviour.

Only 53% of the misaddressed envelopes without a 'please return' message were returned, compared to 67% of those with the envelope message. In other words, the proportion of misaddressed survey invitations not returned can be reduced via a simple instruction on the envelope outer. Analysis of demographic characteristics suggests that households composed of individuals more likely to be 'independent itinerants' (e.g. younger people, renters, and those living in cities) are least likely to return a misaddressed survey envelope. These findings suggest practical implications for survey researchers related to mail survey cost management, response rate calculation, potential bias identification, and the development of sampling methodologies to counter survey error introduced by non-contacts.

Introduction
Mail surveys of the general population often use sampling frames such as an electoral roll or telephone directory. These frames inevitably contain various inaccuracies, but one type in particular, out-of-date address information, causes recurring issues for market researchers. The misaddressing that results from out-of-date data generates more than just a financial cost. Specifically, it can lead to inaccuracies in reported response rates and bias in the population estimates inferred from a study.

Response rate inaccuracies occur when non-contact is not adequately taken into account, often because only a subset of those sent misaddressed mail return it. For example, Hutt (1982, cited in Esslemont and Lambourne, 1992) posted 300 deliberately misaddressed
envelopes to households in the UK, of which 68% were returned. Similarly, Esslemont and Lambourne (1992) sent 200 misaddressed questionnaires within New Zealand of which 70% were returned. Moreover, a recent study in the United States by Braunsberger, Gates, and Ortinau, (2005) found that only 41% of the 1000 deliberately misaddressed questionnaires they mailed were returned unopened.

These results suggest that non-contact rates (and, therefore, response rates) are frequently underestimated. Indeed, Gendall, Finn, and Hoek (2004) found in a study of mail survey non-responders that an average of 13% of those never heard from and often assumed to have refused to complete the questionnaire should actually be classified as non-contacts because they had not actually received the survey package sent to them. As such, Gendall et al. proposed an adjusted response rate formula to compensate for the underestimation of response rates caused by undetected failure to receive.

Errors in response rate calculation as a result of unidentified non-returners may initially be considered trivial because they have no subsequent impact on the calculation of sampling error, and relate to an aspect of mail survey nonresponse (i.e. non-contact) that could be seen as inconsequential or unavoidable.

This view is supported by the fact that most efforts to reduce survey nonresponse have focused on reducing refusal rates. To this end, survey design elements such as repeated contacts, incentives and stamped return envelopes have been identified as improving response rates (e.g. see Dillman, 2000; Yammarino, Skinner, & Childers, 1991) and are commonly employed in practice.

However, the error associated with non-contact is potentially as great as the error due to refusals and therefore as worthy of consideration. Bias can be introduced by non-contact in a survey situation where there is a relationship between respondent characteristics and address inaccuracies; for example, if those who are not contacted differ demographically or behaviourally, or both, from those who either respond or refuse to participate.

Indeed, one might expect that individuals with a higher propensity for moving residence within a given period also have a higher likelihood of being associated with incorrect address information. Preliminary (unpublished) research examining a small number of general population mail surveys undertaken by the authors suggests that such a demographic relationship does exist; non-contacts, as identified by returned misaddressed mail, appear more likely to be younger, male, or undertaking study.

Given the potential for bias to be introduced by non-contact, further investigation is warranted to determine its practical implications and examine whether survey design techniques can reduce it. It may be possible, for example, to oversample individuals with a higher propensity to be non-contactable, such that the final ‘contacted’ sample is at least representative of the population and reliance on post-survey analytical adjustment techniques such as weighting can be reduced. However, before investigations of that nature can progress, a much better understanding of the extent and nature of mail survey non-contact nonresponse is required.

Research to date examining the component of mail survey nonresponse attributable to misaddressed mail is sparse. In fact, only the three studies identified earlier (Braunsberger et al., 2005; Esslemont & Lambourne, 1992; Hutt, 1992) have looked directly at the effects of misaddressed mail. Of these, only the Braunsberger et al. study went further than determining an overall rate of mail return by identifying a gender bias in return behaviour.
The present study was therefore designed to identify and examine in more detail the nature of mail survey non-contact by:

- assessing non-contact return rates from a survey pre-notification mailing to New Zealand households;
- testing a mechanism for increasing non-contact mail return rates; and
- analysing a range of sample household characteristics by non-contact mail return status.

These objectives were achieved by mailing intentionally misaddressed survey pre-notification letters to a sample of New Zealand households.

Method and procedure

A sample of household addresses was obtained by selecting 1400 individuals at random from the New Zealand electoral roll. Each household had a non-equal chance of inclusion as households containing a higher number of enrolled electors had a greater probability of selection. However, such a method replicates the selection process that occurs when individuals are sampled for a survey of the general public and was therefore appropriate.

Each sampled address was randomly allocated to one of two treatments, either an envelope with a 'please return to sender' message, or an envelope with no message. The message was centred on the bottom front of the envelope, and consisted of the following statement:

IMPORTANT: If this mail has not reached the intended person and cannot be forwarded, please mark the envelope "Return to Sender" and place it in a NZPost box.

Every 'banker' size envelope was addressed to a fictitious recipient, Miss P J Nettlefold, a name used in two of the prior studies mentioned earlier. A search of the electoral roll found only one person with the surname Nettlefold and they were excluded from the sample. A photographic search found no other surnames that were similar to Nettlefold.

The envelopes contained a pre-notification letter advising the recipient that they had been selected to complete a survey which would be sent through the mail at a future date. As the envelope was addressed to a fictitious person, the letter should not have been read.

To enable analysis at the household level, a number of electoral roll variables relating to individuals registered against the sampled households were retained, including age (within 5 year band), surname and title.

The sample size of 1400 and use of electoral roll information to profile household return behaviour represent substantial improvements over earlier studies in terms of robustness and potential for return correlate analysis.

Findings and discussion

Response to the mailing and message

Table 1 presents returns by message treatment. A response was recorded against those households that returned an envelope within 35 days (5 weeks) of it being sent.

The envelope message increased returns, with the 14% differential in response (26% of the non-message result) being statistically significant at the 99% level. Researchers who incorporate such a message in their mailings can therefore expect to identify more non-contacts, thereby benefiting from associated cost savings and opportunities for alternate contact or substitution.

Of note is that the non-message treatment achieved a return rate of 53%, lower than the 70% achieved for a non-message mailing in the Esslemont & Lambourne (1992) New Zealand-based study. The disparity may be due to methodological differences (e.g., relating to the
frame employed, sampling procedure, or
evelope size used). Alternatively, it may
relate to changes in response behaviour,
as evidenced by the decline in mail sur-
evay response rates over the intervening
period (e.g. see de Leeuw & de Heer,
2002). More research is required to
identify how each of these possible fac-
tors affect rate of return.

As demonstrated in Figure 1, while there
were differences in the level of response
for each treatment, the distribution of
returns over time was essentially the
same. Both treatments had slightly
fewer than 90% of returns made within
14 days of mailing.

These response distributions suggest
that a period of at least 10 to 12 days
should be given after an initial survey
mailing before a reminder is sent to allow
the majority of misaddressed mail returns
to be identified from an initial wave.
Furthermore, the overall response results
support the assertions of the prior stud-
ies such as Gendall et. al. (2004) and
Braunsberger et. al., (2005) that many
survey response rates are likely to be
underestimated due to a non-trivial level
of unreported survey delivery failure.

Are households that return the envelope
different to those who don’t?
To ascertain whether the household com-
position of envelope ‘returners’ differed
from that of ‘non-returners’, indicators
of average household age, total number
of electors, number of unique surnames,
and household gender split were devel-
oped using information contained in the
electronic version of the New Zealand
electoral roll. In addition, address-related
variables such as geographic location and
address type were examined. Significant
differences between ‘returners’ and ‘non-
returners’ were found on a number of the
derived variables.
Specifically, tables 2 and 3 suggest
that return rates are strongly related to
household composition. It appears that
younger households (average age 18-29)
are more than 2.5 times less likely
to return than elderly households (average
age 70+). Furthermore, households in
which the inhabitants share the same

Table 1: Messaged envelopes achieved a significantly higher return rate

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Sent</th>
<th>Returned</th>
<th>Avg. Return Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Envelope without message</td>
<td>700</td>
<td>368</td>
<td>(53%) 10.0 days</td>
</tr>
<tr>
<td>Envelope with message</td>
<td>700</td>
<td>467</td>
<td>(67%) 9.9 days</td>
</tr>
<tr>
<td>Total</td>
<td>1400</td>
<td>835</td>
<td>(60%) 10.0 days</td>
</tr>
</tbody>
</table>

Figure 1: The Time Distribution of Returns was Similar for Both Treatments
surname were significantly more likely to return the envelope than 'typical' mixed households (i.e. those with two, three, or four surnames).

Table 2: Households with a higher average age returned at a higher rate

<table>
<thead>
<tr>
<th>Average age</th>
<th>n</th>
<th>Returned</th>
</tr>
</thead>
<tbody>
<tr>
<td>18-29</td>
<td>138</td>
<td>30%</td>
</tr>
<tr>
<td>30-39</td>
<td>419</td>
<td>49%</td>
</tr>
<tr>
<td>40-49</td>
<td>376</td>
<td>63%</td>
</tr>
<tr>
<td>50-59</td>
<td>200</td>
<td>60%</td>
</tr>
<tr>
<td>60-69</td>
<td>128</td>
<td>77%</td>
</tr>
<tr>
<td>70+</td>
<td>139</td>
<td>84%</td>
</tr>
<tr>
<td>Overall</td>
<td>1400</td>
<td>60%</td>
</tr>
</tbody>
</table>

Note: Returns from all of the age group pairings except 40-49 and 50-59, and 60-69 and 70+ were significantly different at the 90% level.

Table 3: Single and many surname households returned at a higher rate

<table>
<thead>
<tr>
<th>Surname</th>
<th>n</th>
<th>Returned</th>
</tr>
</thead>
<tbody>
<tr>
<td>One</td>
<td>855</td>
<td>66%</td>
</tr>
<tr>
<td>Two</td>
<td>320</td>
<td>50%</td>
</tr>
<tr>
<td>Three</td>
<td>121</td>
<td>38%</td>
</tr>
<tr>
<td>Four</td>
<td>35</td>
<td>49%</td>
</tr>
<tr>
<td>Five or more</td>
<td>69</td>
<td>72%</td>
</tr>
<tr>
<td>Overall</td>
<td>1400</td>
<td>60%</td>
</tr>
</tbody>
</table>

Note: Single surname households returned at a significantly higher rate than Two, Three, and Four surname households at the 90% level.

Of note is that households in which five or more surnames were present recorded the highest rate of return. These tend to be rest homes or shared residences such as student hostels, which have different processes for dealing with mail to those used by 'typical' households.

Also of note is that the substantial gender differences in return behaviour identified by Braunsberger et al., (2005) were not replicated in this study. Braunsberger et al., found that 53% of the envelopes sent to female recipients were returned, compared to 28% of those sent to male recipients.

In determining whether a recipient was male or female, Braunsberger et al. used the original address details from their mailing list. That is, if the mailing name was originally for a male, they assumed that the recipient of the misaddressed envelope would be male. This was an understandable assumption given the limited household demographic information available in that study. However, unless the envelopes were sent to single-person residences, it is likely that at least some misaddressed envelopes were attended to by a household member other than the one on the original mailing list. As such, the gender differences identified are likely to have been subject to confounding factors.

With the above in mind, a different approach to gender response comparison was employed in this study. The gender composition of each household mailed was compiled using the name information for electors registered at each address. No significant difference in return behaviour was found between households that were predominantly male and those that were predominantly female. Similarly, no significant difference was identified between households comprised solely of females and those comprised solely of males.

To complement the household demographic measures in Tables 2 and 3, sampled addresses were categorised according to dwelling type and location. For dwelling type, an address for a rest home, hall of residence, or other group accommodation was classed as multi Residence. Addresses containing a Rural Delivery code or Post Office reference (e.g. PO Box), were classified as Rural Delivery and Post Centre, respectively.

Simple residential addresses (e.g. 10 Smith Street) were classed as Single
Addresses, while more complex addresses (e.g. 10-A Jones Street) were classed as Split Addresses, to differentiate those more likely to be family homes from those more likely to be flats. The location classification was based on the town or city of the address. Although imperfect, these classifications enabled a basic examination of differences in return rates, as presented in Tables 4 and 5.

Table 4: Split address households were least likely to return

<table>
<thead>
<tr>
<th>Dwelling type</th>
<th>n</th>
<th>Returned</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multi Residence</td>
<td>23</td>
<td>83%</td>
</tr>
<tr>
<td>Rural Delivery</td>
<td>172</td>
<td>76%</td>
</tr>
<tr>
<td>Post Centre</td>
<td>86</td>
<td>71%</td>
</tr>
<tr>
<td>Single Address</td>
<td>861</td>
<td>57%</td>
</tr>
<tr>
<td>Split Address</td>
<td>258</td>
<td>50%</td>
</tr>
<tr>
<td>Overall</td>
<td>1400</td>
<td>60%</td>
</tr>
</tbody>
</table>

Note: Single and Split Address households returned at significantly different rates to all others at the 90% level.

Table 5: Households in metro areas were least likely to return

<table>
<thead>
<tr>
<th>Location</th>
<th>n</th>
<th>Returned</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metropolitan</td>
<td>796</td>
<td>53%</td>
</tr>
<tr>
<td>Provincial</td>
<td>410</td>
<td>67%</td>
</tr>
<tr>
<td>Rural</td>
<td>194</td>
<td>72%</td>
</tr>
<tr>
<td>Overall</td>
<td>1400</td>
<td>60%</td>
</tr>
</tbody>
</table>

Note: While the Provincial and Rural addresses had similar return rates, the ‘Metropolitan’ households returned at a statistically different rate from the others at the 90% level.

The key finding presented in tables 4 and 5 is that Split Address and Metropolitan dwellings were least likely to return misaddressed mail, a result which reinforces the conclusions drawn from tables 2 and 3.

A Chi-Square analysis of all two-way combinations of the 4 variables presented above produced results that were consistently significant at the 99% level. Thus, although they measure different attributes, the household composition and location factors examined here all appear to be associated with a common underlying phenomenon which is in turn related to non-contact return response.

Specifically, households comprising individuals more likely to be itinerant and independent (younger, mixed inhabitant, split address, and metropolitan) appear less likely to return misaddressed survey invitations. Indeed, while the overall return rate for the present study was 60%, return rates for households associated with ‘independent itinerancy’ ranged from 30% (average age) to 53% (location).

Within the context of this study, we can only speculate why households that fit the ‘independent itinerancy’ profile under-return. However, a promising potential explanation can be found in the social psychology literature concerning altruistic behaviour. Guy & Patton, (1989) report that a number of studies have found a ‘diffusion of responsibility’ effect where an appeal for an altruistic behaviour is not clearly directed to a specific individual in a group situation. That is, it has consistently been found that assigning responsibility for an act to a group reduces its likelihood of occurring.

This explanation is appealing for a number of reasons. First, non-contact return can be classed as an altruistic behaviour because nothing is gained by those who undertake it. Second, the results of the study are consistent with what it would predict; ‘independent itinerancy’ is associated with private households containing multiple members living in a situation where responsibility for various household processes is shared. Conversely, it is not associated with cohabitation situations in which responsibility for household processes are likely to be clearly defined or institutionalised (e.g. halls of residence, rest homes, or private family residences).
Finally, the explanation is testable; future research can be directed at more clearly establishing that households associated with ‘independent literacy’ tend to have less clearly defined responsibility assignment relating to anomalous mail. Whatever the ultimate reason for the behavioural patterns found in this study, it is clear that they give weight to the assertion that non-contact is a much larger problem than typically exposed by returned mail.

Implications for response rate formulae
Although the implications of the non-contact return results for response rate calculation can be expressed here, more research is required to provide a generalisable empirical framework for addressing them. The framework for this work is discussed below.

Currently, a common formula for calculating response rates is:

$$RR = \frac{VR}{(VR + IR + R + US) - (I + NC)}$$

Where:
- RR = Response Rate
- VR = Valid Responses
- IR = Invalid Responses
- R = Refusals
- US = Unknown Status
- I = Identified Ineligibles
- NC = Identified Non-contacts

Often, an implicit assumption is made by researchers that those in the ‘Unknown Status’ category (i.e. respondents from whom nothing is heard) are all passive refusers. However, in light of the results of this study and others cited earlier, it is apparent that an adjustment is required to the ‘non-contact’ term (and, by implication, the US term) in the response rate formula to better reflect the true incidence of misaddressing. Indeed, although comment on the estimation of eligibles is beyond the scope of this paper, it is also possible that an adjustment should be made to the eligibles term to reflect unreported ineligibility.

Thus, a more appropriate formula for calculating response rates is:

$$RR = \frac{VR}{(VR + IR + R) + (US - (EI + ENC)) - (I + NC)}$$

Where:
- EI = Estimated Ineligibles
- ENC = Estimated Non-Contacts

The estimation of non-contacts requires knowledge of the expected proportion of incorrect addresses, or non-contacts ($p_{nc}$), in a specific sample. Once known, ($p_{nc}$) can be applied to the total sample size (TSS) and the identified non-contacts subtracted from it to give the number of estimated non-contacts.

$$ENC = (p_{nc} \times TSS) - NC$$

The expected proportion of non-contacts ($p_{nc}$) will be frame-dependent and empirical work is therefore required to build generalisable predictors of $p_{nc}$ for commonly employed frames. Unless a method is available for verifying every non-contact in selected mailings from a particular frame to establish such generalisations, this work will rely on a frame-independent means of extrapolating total non-contacts from those verified (i.e. returned) in mailing tests.

This study provides both a methodological and empirical foundation for such an extrapolation. For example, the results reported can be applied in a univariate fashion to studies in which one of the four variables examined here is known. Alternatively, multivariate models might be developed to combine variables found, by studies employing a similar methodology to that used here, to relate to return behaviour. Those models could then be applied to other datasets and frames as appropriate.

As stated earlier, despite the solid start provided by this and prior research, much work is still required in this under-investigated area to enable research-
ers to reliably estimate non-contacts. Specifically, further research is required to identify other return correlates, to explore the theoretical basis of return behaviour, and to generalise findings to different research environments.

In the absence of this additional research and dataset variables matching those examined here, practitioners can tentatively employ the following rule of thumb to estimate total non-contacts: identified non-contacts (i.e., ‘gone no address’ returns) to a single-shot mailing without an envelope message should be doubled.

One important caveat associated with this rule is that it assumes a random distribution of misaddressing (and, hence, non-contact) amongst the population. However, preliminary research by the authors not reported here suggests that ‘independent itinerants’ may be more likely to have mail misaddressed to them because of their higher frequency of movement. That is, it is possible that a form of ‘double jeopardy’ exists with respect to non-contact; those who are more likely to be non-contacts in a sample are also more likely to have resided in households that won’t notify you about it. Those applying the ‘double it’ rule presented above should therefore understand that it could lead to optimistic estimates of total non-contact.

More sophisticated estimates of total non-contact may need to take into account the occurrence of likely non-contact individuals in the sample as well as the propensity of the household types in which they live to return misaddressed mail.

Moving toward bias reduction mechanisms and improved database quality
Although exploratory, these findings are important because they relate to a systematic (i.e., survey and frame unspecific) form of nonresponse and link back to respondent information available in mailing frames such as an electoral roll. In doing so they illuminate the extent of the non-contact problem and provide a foundation for the development of methods for identifying and estimating the true proportion of respondent non-contact in a variety of mail survey contexts.

The utility of such knowledge is not limited to the calculation of response rates. Rather, this greater understanding of non-contact will form a critical input into efforts to create methods with the potential to mitigate the nonresponse error introduced where mail survey contacts and non-contacts differ on important attributes. One possible method would involve oversampling potential respondents identified as having a higher propensity to be non-contactable to ensure that the reached sample is more representative of the population.

Better knowledge relating to non-contact may also assist database maintenance efforts by enabling organisations that collect contact information to better focus resources on records more likely to be out-of-date by virtue of the individuals or households to which they relate.

Conclusions and future directions
An envelope message requesting return upon misaddressing was found to improve identification of mail survey non-contact by a factor of just under 30%. Such an improvement provides opportunities for respondent substitution where non-contact is identified, greater accuracy in response rate calculations, and a potentially substantial reduction in the number of reminders sent to non-contact addresses.

More research is required to identify whether variables such as envelope size, message content, or sender source mediate misaddressed invitation return rates. Future studies might also incorporate a survey to enable an assessment of the proportion of misaddressed questionnaires that will be returned completed by respondents other than the addressee, as examined by Esslemont & Lambourne.
Another important finding was that households of 'independent itinerant' individuals were found to return misaddressed mail at significantly lower rates than other households. Unpublished preliminary research by the authors suggests that this sub-group is also more likely to have mail misaddressed to them because of their higher frequency of movement. Hence, sophisticated estimates of true non-contact may need to take into account the occurrence of different types of individuals in the sample as well as the propensity of the households in which they typically reside to return misaddressed mail.

These results lay the foundation for methods to better identify and estimate the proportion of respondent non-contacts in mail surveys. In doing so, they also contribute to the development of sampling methods aimed at mitigating the potentially substantial nonresponse error introduced where mail survey contacts and non-contacts differ on important attributes.

The improved understanding of non-contact presented here and generated by future research in the area may also provide an important base from which organisations that collect address data for business or research purposes can begin to better focus their database maintenance resources.

References


Acknowledgements

The authors would like to thank Margaret Faulkner for comments on a conference presentation relating to this study which led to an examination of the literature regarding diffusion of responsibility discussed herein.
Book Review

The Art and Science of Interpreting Market Research Evidence

DVL Smith and JH Fletcher, John Wiley & Sons, 2004

Reviewed by Duncan Stuart FMRSNZ

This book has been written in response to the palpable drift that market research has taken away from its traditional reporting role and toward the decision-making cocoon of client organisations. The paradigm is one the authors refer to as ‘new’ market research and among its characteristics are:

- A shift toward decision facilitation.
- A more holistic view – employing quantitative as well as qualitative market evidence.
- Analytical frameworks that combine hard market research data with prior management knowledge and intuition.

Some researchers will question whether ‘new’ research is much different from the working style they’ve operated by over the past decade, but it is fair to say that the paradigm set out by Smith and Fletcher sits somewhere to the left of the way large research organisations operate; structured as they are with silos labelled qualitative and quantitative, and with a focus on research data only, rather than on multiple sources of evidence.

The beating heart of the text, and the hub of the argument presented by Smith and Fletcher, is the trio of chapters about Developing the Analysis Strategy, Organizing the Qualitative Data and Organizing the Quantitative Data. Here, the platform is laid down for a more strategic approach to analysis than many researchers may be accustomed to. In essence the authors set out a template for how to establish the main stories, and how to choose the best frameworks for the analysis. They are somewhat apologetic about treating qual and quant in separate chapters, but each side of the coin needs discussion, and the main point.

For this reviewer, however, the outstanding chapters sit at the back of the book: the trio of chapters that detail the presentation of research as a narrative, the decision-making process and the development of holistic analysis.

Market research books run the risk of being rather dry. In this case the voice of the authors is lively and logical and the tone has a welcome collegial ring about it. These are colleagues sharing their views: they don’t lecture. But here, in the last chapters the voice of the authors lifts a notch as they share their passion for the Big Picture: the wonderful process of synthesising the available information and weaving this into a clear story from which decisions can be made.

Smith and Fletcher aggressively attack the old building block approach that assembles research evidence in copious but separate piles but leaves to the client the act of drawing conclusions and making decisions.

When it comes to facilitating the decision making process, the authors effectively arm the reader with techniques to keep the researcher in the decision-making loop: an
antidote to the classic scenario where fine research ends up in the bottom drawer. The techniques include:

- Helping client focus on the right problem. Define the problem precisely.
- Weight or prioritize the decision criteria – don’t let little factors overshadow the decision.
- Review the alternative courses of action in a flexible, imaginative and innovative way. Here they provide a number of techniques to enable stakeholders with different operating styles to look at the options from different perspectives.
- Itemise the consequences of actions. Help the stakeholders consider the what-ifs.

Part of this chapter is devoted to the common errors or perceptual flaws in decision-making, for example the flawed self-knowledge of the decision-makers (the human tendency to remember our successes and forget our failures) as well the problem of lazy, cliché or stereotypic thinking which can overshadow the weight of good research evidence that has been laid out in front of the decision-maker.

One example I like is the “competition factor” which is the point that research, and the decisions that result, are usually conducted in a context where the competitors remain constant. But what will happen if our research leads to a decision that precipitates a big reaction from competitors in the market? How often do we consider the decision, but not the game plan?

The final main chapter, on the development of holistic data analysis raises the discussion to the meta-level: talking about how researchers can generally lift their game and, in consequence, become more pivotal within the decision-making process of clients.

This is nothing less than a gentle, professional call to arms: a call to stop describing the market and to work more closely with the people who wield the power in that market.

I could not help but wonder whether this thematic direction is a reflection of the authors’ own professional journeys as strategic researchers. Perhaps they too have seen fine research get shelved while wrong-headed decisions got made in client boardrooms. Perhaps they too have wondered why the dry financial analysts are granted Platinum admission to the power circle, while market researchers gain only a limited one-hour show and tell pass. They don’t say, but they do acknowledge the insight of their strategic research colleagues.

Whatever. In this volume they share a sense of empowerment that ought to rub off on readers.

In support of the logically organised text, the layout of the book is also mostly clear and navigable. Each chapter begins with a summary of key bullet points, and the chapters themselves develop like a series of stepping stones: a rare feat compared to the many research texts that zigzag around various techniques without ever quite following a thematic pathway. My only quibble here is with the placement of the (many) illustrations: some are several pages away from the text to which they refer. Luckily, what separates researchers from the lesser species is our ennobling ability to use the thumb as a bookmark while reading textbooks about analysis.

Does the book succeed on its own terms? If it seeks to contextualise the analytical
process within the framework of 'new' market research, the term employed in the opening notes of the preface, then I feel the authors haven't quite risen to their own challenge. For example they profess to take a Bayesian attitude ("we make no apology for giving the Bayesian approach prominence in this book") but then sideline this 200 year old school of statistics by concluding: "Bayesian methods, while commonplace in many academic and technical research circles are considered by most commercial market researchers to be far too inaccessible to be developed into everyday practice." As a result they give the subject less coverage than it deserves: a clear example of practical regression where the norms of the industry have moderated the mission of the authors.

Elsewhere, and for similar reasons, I felt the authors hadn't fully explored their options. The short section on how to present verbatim data skipped almost completely the existence of useful text-analysis tools that (like the available Bayesian softwares) are enabling researchers to integrate qualitative and quantitative data for more holistic thinking. Again, in just one line, the authors report low usage amongst commercial UK research companies of suitable softwares and then let this foreclose their discussion. Why let the laggardly adoption by commercial research companies crimp the discussion? In essence Smith and Fletcher are too willing to describe the current envelope, but aren't willing enough to push it to its potential.

I found this somewhat frustrating and a mild case of over-promise and under delivery in light of the 'new' research label they created. The authors admit in the forward that they have kept back on the drawing-board some of their more innovative ideas. That's a shame. As a consequence the book seldom challenges as hard as it could. It shies from the edges and stays in the middle ground.

Fortunately there is plenty of rich middle-ground to discover. The main argument of this book, and the practical how-to content, is well described and explained extremely clearly. For market research firms that wish to re-ignite internal discussion about how to make their analytical approaches more decision-focused and more holistic, this is an excellent place to start.

In fact it would make excellent standard issue for new employees and as a platform for internal training. The authors clearly have this in mind, and even support their text with a ten-unit lecture, available on PPT, to help educationalists and trainers teach along the guidelines they've so clearly set out.
Book Review

Developing Brands with Qualitative Market Research


(Book 5 of a 7 volume set titled Qualitative Market Research: Principle and Practice)

Reviewed by Peter Link – Senior Research Consultant with Chant Link & Associates, Melbourne.

Synopsis
This book on brands is one of two books in this series that address the content of typical MR projects rather than any specific processes (the other concerns advertising – Book 6). The book argues that a large proportion of commercial MR is about brands, and that qualitative MR is uniquely able to address brand issues.

Highlights
Ch 1: The Nature of Brands: This chapter defines and describes the notion of brands, from a historical perspective through to highlighting that the term “brands” is defined in so many ways that marketing as a discipline has not sufficiently standardised its thinking about it. All the various definitions though, have nine key themes:

- Brands as causal (impact behaviour),
- Brands as authentication,
- Brands as differentiators (seen as crucial by the author),
- Brands as coherent (a distinct meaning system),
- Brands as cultural icons (brands exist simultaneously at a cultural and individual level),
- Brands address needs (ie address consumers’ emotional and psychological needs as well as practical demands),
- Brands and orientation (tied up with processes of identity formation, expression and orientation),
- Brands as contract (consumers make implicit deals with brands that bind brand to delivery),
- Brands as charismatic (the allure or magic of a brand can solve issues that can’t be resolved rationally).

Each of these is explained in some detail, with examples. Accept them or not, the scope for qualitative MR in the brand space is already apparent. Chapter One concludes with a model of these brand effects, which is worth a look (page 22).

Ch 2: Brands and the Human Dimension
This chapter explores how brands exist in everyday life, their genesis and evolution within consumer culture. It starts by discussing some theory of psychology and the "Johari Window" – a heuristic device used to portray the existence of different levels of the mind (originally developed as a model for describing and managing interpersonal relations). Its basic propositions are that:

- There are some things people know and are happy to talk about.
- There are other things people know but try to hide.
- There are some things people don't consciously know, but would be happy to talk about if they did.
- Other things people don't know and would not be happy to talk about them if they did.

Other such psychology theories are also explored and general observations made about the nature of the mind and how these relate to the world of brands, eg:
- Brand associations are mostly unconscious.
- Brand associations often exist in the mind in the form of metaphors.
- The majority of brand associations are visual rather than verbal.
- Brand associations are also stored in the form of sensory impressions like smell, taste and sound.
- Brand associations also exist in the form of emotional responses.

Fragmentation of the Consumer: A discussion of brands in every day life explores culture and its manifestations. A theme that seems of high importance in this section is "fragmentation of the consumer". In effect, a single consumer is believed to display different personalities according to time, situation, peer group, stimuli and so on. For example, the following have been written about elsewhere:
- Portfolio consumers.
- Spider lives.
- Moments of identity.

Thus, when a brand resonates with an individual, it may have particular unique meanings to them. Thus, it is claimed, the challenge for Coke is to devise ways for the consumer to build and rebuild a positive vision of Coke.

It is further argued that marketing communication is all about attempting to develop in consumers a particular emotional orientation and subjective way of understanding a brand; it is not necessarily about selling it to them at all costs and in all circumstances.

ie. Brands are negotiated by consumers acting both individually and collectively, interpreting and behaving both consciously and unconsciously.

**Ch 3: Branding and Brand Modelling**

This is an interesting chapter as it explains various ways of exploring and describing brands, such as:
- Brand portraits (eg Brand Fingerprint, Brand Diamond, Brand Essence – brand benefits, values and personality).
- Brand Modelling – a tool that can be used to guide the development of brand communications. Brand positioning is a research oriented example. Examples of brand models dealt with include:
  - Brand platform (Feldwick & Bonnai) see page 50.
  - Central Bullseye (McNae) see page 51.
  - Brand cube (Context Research International) see p 51.

A so-called advanced brand model is presented to summarise consumer brand perceptions and value versus external market and marketing factors (see page 54).

Readers are entitled to be sceptical about the claimed differentiation between brand portraits and brand models – it’s all on a single spectrum of complexity or sophistication for me.
Ch 4: Why Use Qualitative Research to Look at Brands and Branding?
This chapter attempts to describe the defining characteristics and rationale for qualitative research and then makes the argument that qualitative research is great for addressing the above brand issues. So it claims qualitative research can address:
- The problem of the withheld.
- The problem of the unconscious.
- The problem of the taken for granted.
- The hermeneutic problem of meaning (language and mistranslation of ideas from language).
- The problem of prediction.
All of these are discussed from the perspective of using qualitative brand research.

Ch 5: Research and the Brand Development Process
This chapter lists the logical functions that need to be undertaken in a brand development process, and suggests key ways in which qualitative research can help. Unfortunately, the narrative is all too predictable: identify needs, explore existing marketplace, evaluate/refine brand/product, identify the meaning set, create the brand model or blueprint. The chapter also explores brand evaluation versus brand development.
A bland chapter...

Ch 6: The Structure of Brand Research
This is about groups versus depths, various timing and logistical issues, sampling issues and stimulus materials. One of the strongest arguments for using groups, it is argued, is they enable the researcher to use a wide array of techniques, such as projective and enabling techniques – to explore the hidden or obscured world – considered relevant for brand studies.

It is also argued that brands are fundamentally cultural in nature, and thus shared views can best be uncovered in a group setting.

It is argued that depth interviews are likely to be more helpful in evaluating how consumers interpret communications. (p 87).

The discussion of logistics, sampling and methods of showing stimulus materials is not very insightful or new.

Ch 7: The Content of Brand Research
Explores projective and enabling techniques in brand research. You can imagine the kind of thing here – word association, associative chains, laddering, bubble pictures (cartoon completions) and so on. Some good(!!) ones included guided fantasy and brand graffiti exercises. In the latter, consumers are given some representations of a brand and maybe some of its products and asked to deface these by scribbling graffiti over it. It is claimed:
- This can give useful insights into consumers' relationships with the brands and the paraphernalia associated with them.
- It can very useful in identifying some underlying passions within consumer culture and this can assist needs identification.
A very long list of such techniques is presented with explanations of how to deploy them in brand research. There was not enough time spent on the old problem with these techniques – how to interpret the results.

Ch 8 Developing Brands Through Qual Research: A summarising chapter.
Conclusion
Quite a useful book if you are looking at a branding research project. The strongest sections are early in the book – where the definitions of brands and the psychological and cultural issues surrounding brands are discussed.

The various ways of analysing a brand (brand portraits and models) is a useful summary, as is the comprehensive list and explanations of projective techniques and their harnessing in brand research (but only if you believe you can interpret the findings of such techniques).
Book Review

Developing Advertising with Qualitative Market Research

(Book 6 of a 7 volume set titled Qualitative Market Research: Principle and Practice)

Reviewed by Peter Link – Senior Research Consultant with Chant Link & Associates, Melbourne.

Synopsis
This book on advertising research is one of two books in the series that address the content of typical MR projects rather than any specific processes (the other concerns brands – Book 5). The book argues that one third of qualitative MR in the UK is advertising related, and thus, qual MR must be managed carefully. It argues that the researcher must understand how advertising works (that simple assumptions about this are often false, leading to worthless research).

Highlights
The first two chapters are Advertising 101 – but arguably necessary to cover all readers.

Ch 3: How Advertising Works
This is a better grounding for qual researchers to read. It outlines the various models of advertising as per Advertising 201:
- The famous but now most derided AIDA model (awareness, interest, desire, action) Lewis, 1900.
- DADGAR (defining advertising goals for measuring advertising results, Colley and others - late 1950s and 60s).
- USP (unique selling proposition) (Reeves, 1960s).
- The persuasion shift model (Schwirin, 1940s).

There is discussion of Fishbein (1975) who showed the relationship between attitudes and behaviour is complex, and the latest thinking of the 1990s (pulled together by Hall and Maclay) that four models have been influential:
- Sales response model.
- The persuasion model.
- The involvement model.
- Saliency – uses distinctive forms of advertising; consumers respond by talking spontaneously about the brand etc; puts the brand on the consideration list in the consumer’s mind. This model is thought to be the most favoured in the UK at present.

This chapter is quite interesting because it goes on to explore how “low involvement processing” works, how affective memory is more important to decision making than cognition, so our hearts as well as our heads are important in decision making. The consumer is seen as active rather than passive in processing advertising.

There is also a good description of the way in which various consumers think advertising works (eg the sophisticated critic, the uninhibited appreciator, the careful deliberator, the suspicious rejector) concluding that researchers must realise that some of these categories are people who will always reject a script etc. Outrageously (in my view) a further conclusion extends this: “Most researchers exclude people who have
strong negative attitudes towards advertising in the belief they will at best contribute little to any discussion about advertising and at worst, influence other respondents around them.

Ch 4: The Planning Process and the Role of Research
This chapter describes the process an ad agency goes through in creating and honing an ad campaign or individual ad. While mildly interesting, the whole issue of using market research to assist in formulating an advertising brief is missed.

Ch 5: The Research Process: From Briefing to Fieldwork
This is fairly ho hum, describing the research steps in basic language, and is very junior in target orientation. There is however a small section called The Semiotic Approach (page 52) in which it is claimed that semiotic analysis of advertising and markets is becoming increasingly important in the practice of UK qualitative research, and has been described as one of the “three grand pillars” of qualitative research, alongside traditional motivational research and ethnography. The terminology of semiotics is then explained in some depth (signifier, signified, sign, symbolic, codes, residual, dominant, emergent, discourse).

The text seems to promote and condone the use of research respondents as collaborators with advertising creative departments rather than using research as a scientific independent process of understanding the likely effects of an advertisement, and the reasons for such effects on an audience.

There is a claim that the UK is the only country in the world where qual research is not routinely conducted in viewing facilities. There is a hint that a debate is to be had about viewing/not viewing but this is left to other books in the series.

There is a useful list of “must talk-about”s on page 67, in advertising research: Four examples out of 10 include the following.
- Comprehension.
- Brand.
- Communication message.
- Brand imagery.

A list of suggested probes is provided for each of these, which could prove useful.

There is a brief discussion about conducting groups and depth interviews, which includes some discussion of NLP – neuro-linguistic programming (page 70) (eg “mirroring – useful if a group or respondent is stuck or over-excited, maybe with indifference. The interviewer/moderator “mirrors” the group/respondent by leaning forward when they are all leaning forward, talking quickly when they are talking quickly, giggling with them etc., then slowly changes to a more productive stance – the group will follow, and all will be well and certainly improved.”)

One wonders how many focus groups the authors have conducted, but maybe this is worth a try!!

Ch 6: The Research Process From Analysis and Interpretation to Presentation
This is quite thought provoking. A contextual onion to aid analysis (see page 79) points out how the researcher must seek to analyse the following in order to draw valid conclusions about a particular response:
- What the respondent said.
- Who the respondent was (eg brand lover, lapsed user, advertising hater).
- What was happening in the group.
- Environmental considerations.
- Cultural/moral/political considerations.

We are reminded that researchers approach the ad research project in three different consecutive frames of mind: subjective, objective and judgemental, and that different researchers will discover different insights and make different recommendations – all of which can be equally valid – something that is worthwhile remembering when we are jointly interpreting our findings in qual research projects!!!

The chapter spends time showing that part of the analysis and interpretation of ad research data involves understanding the structure of the ad, and the interlinking roles that the various elements have in that overall structure, as well as appreciating the unfolding narrative of any ad.

Likeability is considered by many as the best available test of predicting ad success. However, some ads are instantly liked (and could wear out quickly) while others are more likeable as they are seen several times. So, it is concluded that qualitative researchers should not place too much priority on likeability of ads.

Remaining Chapters
These deal with ad stimulus materials, story boards (OK), narrative tapes have problems, outdoor poster images great, mag mock-ups etc are fine. This is mildly interesting with reasonably useful hints about how a moderator should present ad materials to avoid bias or false results (Ch 7), and how to play the politics inside the agency and with the clients (Ch 8). This leads to an inevitable conclusion that there needs to be more teamwork between those on the apexes of the triangle - researcher, client and agency.

Ch 9 deals with international advertising research, focusing on the additional complexities and pitfalls of multiple cultures, and more political difficulties with multiple agencies and multiple "clients within the client" for the qualitative ad researcher to deal with.

Conclusion
This is a useful basic primer on advertising models and theories, with some useful ideas for qualitative researchers in this kind of work. However it displays patchy quality and depth for mine.