AUSTRALASIAN JOURNAL OF MARKET & SOCIAL RESEARCH

Editor
Professor Laster W. Johnson
Melbourne Business School

Editorial Advisory Board
Susan Ellis
Melbourne Business School
Mark Jessop
Disability Services Commission
V Kumar
University of Connecticut
Peter Oppenheim
Deakin University
Pascale Quaster
University of Adelaide
Liane Ringham
Inside Story
John Rossiter
University of Wollongong
Geoffrey Soutar
University of Western Australia
Don Stern
Washington State University
Jill Sweeney
University of Western Australia

AJMSR is published by the Australian Market & Social Research Society
Copyright 2006. ISSN 1832 7382

Australian Market & Social Research Society
Level 1
3 Queen Street
Glebe NSW 2037
Australia
Phone: +61 (0) 2 9556 3100
Fax: +61 (0) 2 9571 5944
Email: publications@amrsra.com.au
Web: www.amrsra.com.au
The Australasian Journal of Market & Social Research (AJMSR) is the official journal of the Australian Market & Social Research Society (AMRSR). It is published twice a year in June and December.

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:

**Australian Orders**
- $48.40 (inc GST) per single issue
- $82.50 (inc GST) per one year subscription

**International Orders**
- $53.00 (AUD) per single issue
- $94.50 (AUD) per one year subscription

**Submissions**
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 practicing market or social 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 Leicestler Street
Carlton VIC 3053 Australia
Email: L.Johnson@mbus.edu

**Subscriptions and advertising**
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 9565 3100
Fax: +61 (0) 2 9571 5944
Email: publications@amsrs.com.au

**Advertising rates**
Logo placement – outside front cover ($577.50 inc GST)
One page colour ad – inside front cover ($539.00 inc. GST)
One page colour ad – outside back cover page ($803.00 inc. GST)
One page black and white ad – inside ($269.00 inc. GST)

Copies of the articles in AJMSR 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.

AMRS recognises the contribution of Professor Lester Johnson (Editor) and the Melbourne Business School.
AJMSR – Editorial

The current issue of the journal contains four papers covering a range of topics of potential interest to the market and social research community. The first by Phil Gendall and Janet Hoek illustrates what can happen when questionnaire design principles are overlooked. They illustrate the problems that can result, using the 2003 International Social Survey Programme survey on national identity.

The second paper by Don Porritt discusses how an analyst might go beyond the usual cross-tabulations and multivariate techniques provided in standard statistical packages by theorising about the causal processes that underlie the relationships among the measured variables in a study. The result can provide much more insight into what is really going on. He illustrates with two examples, one on perceptions of company reputations, the second on customer satisfaction with the experience of buying a new car.

The next paper by Mathew Parackal and Ron Garland examines the contextual requirements of the 11-point Juster purchase probability scale. Three factors, namely question order, item order and question version, were examined. Results of their study suggest that the Juster scale does not require any additional contextual input and remains a satisfactory instrument for collecting purchase probability data.

The final paper in this issue by R. Susan Ellis examines the issue of information overload from both a psychological and a managerial perspective. Guidelines are provided to help researchers minimise information overload when preparing material and presentations for their clients.

Lester W Johnson, Editor
Australasian Journal of Market & Social Research
IMAGINATION Conference Review

The theme for the 2006 Australian Market & Social Research Society (AMSRS) conference was IMAGINATION - the ongoing challenge to develop ideas and new solutions that stay relevant and fresh.

Delegates were privileged to have three international keynote speakers, Dr Sheila Keegan and Gill Ereaut both from the United Kingdom and Professor Duncan Simister from Massachusetts Institute of Technology Sloan School of Management in the United States of America.

The conference organising committee, led by Dr Vicki Arbes, encouraged speakers to examine the industry’s place in the broader world in concert with other service providers, the wider business community and academia, and specifically how the creative craft of research can build business and communities.

Matt Balogh Wins Tony Wheeler Award

All submitted papers were eligible for the Tony Wheeler Best Paper Award, which is sponsored by Millward Brown and is named in recognition of its former CEO for Australia, New Zealand and Japan. A co-founder of Hoare Wheeler and Associates and then of Yann Campbell Hoare Wheeler (YCHW), now Millward Brown, Tony Wheeler was the Secretary of the Society’s National Council from 1985-87.

Matt Balogh’s paper “Why Dogs are Smarter Than Banks: a Crazy Multi-Dimensional Tale of Customer Loyalty” delved into dimensions and the space-time continuum to deliver a mind-bending look at how quantum physics can expand the thinking of market and social researchers. He drew parallels between quantum physics and qualitative research, in that both are generalisations based on observations that cannot be narrowed down to the specific, and challenged delegates to think about how qualitative research could get richer by making more generalisations.

‘Quantum physics is the market and social research of the physics world. It can measure behaviours and make predictions in certain quantities but when it is reduced to the molecular, atomic and sub-atomic level, it no longer works. For example the calculation of half lives of atoms can be very accurate, in fact that is how the most accurate clocks in the world work. But if we take a single atom of uranium, we can do no better than a weather forecaster, because we can make absolutely no prediction about whether it will lose its radioactivity in the next second, or in nine billion years time, or any time between now and then.’

Matt explored his thesis by looking at customer loyalty. After introducing delegates to Mr Square, who lives in a two dimensional world, he challenged researchers, every time they measure anything, to ask: ‘what if there was another dimension here?’

‘Quite often we are shown concepts that would enormously enrich our thinking if we accepted them as new dimensions, rather than reducing them back to our existing range of dimensions, such as Dimension A (loyalty),’ said Matt.
He proposed a hierarchical map of different dimensions that could translate into brand loyalty, comprising: earned loyalty and product attraction loyalty (both shared by word of mouth); accumulated loyalty, conscience loyalty and inherited loyalty (ingrained but not normally shared); message appeal loyalty and bought loyalty (both dependent on marketing); and trapped (or reluctant) loyalty.

"The right melting pot of accumulated loyalty, earned loyalty, a sprinkling of trapped loyalty and a dash of conscience loyalty can insure a brand or product against disasters," said Matt.

**Highly commended papers**

The judging panel highly commended three other papers.

Barry Elliott's paper "Using Psychological Theory in Social Research can Lead to Imaginative Behaviour Change Strategies: Speeding in WA" used the theorist workshop model (TW) of behavioural change which was extended to include variables based on behavourial science literature in relation to speeding and the authors own theoretical and experiential work in the area.

The results of the model suggest that three most significant influences on intentions not to speed are: (i) self efficiency or perceived behavioural control i.e. motorists who speed believe it is hard not to speed (ii) personal (moral) norm anticipated regret i.e. motorists who speed feel little or no moral pressure not to speed and have only slight (if any) feelings of regret for speeding and (iii) descriptive norm —motorists that speed hold the perception that every one/most drivers speed.

The application of the TW Model to speeding indicates that what is needed (over time) to influence decisions not to speed include (i) to challenge speeders self efficacy beliefs that is difficult to stay within the speed limit (ii) to develop a moral sanction against speeding and (iii) to challenge descriptive norm beliefs that everybody or the average driver is speeding.

Other important influences on speeding identified by the modelling process includes (i) personal identity (image of on oneself relative to others) (ii) response efficacy (the likelihood of positive outcomes of not speeding and (iii) habit.

**Degustation Menu of Ideas highly commended**

Tim Bock from Numbers and Toby Hill from the Leading Edge combined their efforts to present four ideas in helping to understand the complexity of brand image data.

Their paper titled "A Degustation Menu of Ideas for the Analysis of Brand Image Data" outlined creative approaches to comparing and summarizing brand attributes that are not as simple to compare such as the average of age of people in one group to another or popular colours.

The ideas presented were:
1. Synthesising brand association and market structure data. The approach demonstrated how brand association data and market structure data (e.g. brand attitude) could be analysed using traditional correspondence analysis in a way that generated a map reflecting insights in both the types of data.
(ii) Creating Moonplots. While correspondence analysis biplots have been in widespread use in the industry, Moonplots are much easier and accurate to interpret. The Moonplot diagram is circular with the attributes plotted around the perimeter of the map and the size of the fonts of the attributes proportional to their distance from the centre in the map. The map can therefore accurately describe the positioning of certain brands.

(iii) Imagelplex analysis helps to find the complex nature brand image held by different people and understand them, not create an average view. By following a process whereby 'cleaning' the initial data file then using tandem clustering (factor analyse the data and then cluster analyse the factor scores), brands are then profiled by 'Imagelexes'. The main advantage of this idea is that correspondence analysis fails to take into account how multiple attributes combine together in consumer's heads and hearts.

(iv) Plotting the Brandscape shows the diversity of views of brand image in a market via a map of Imagelexes. Techniques that can be used include (a) Brandscape contour plot, (b) labeled Brandscape to bring the map to life to describe the nature of the Imagelexes varies over the map and (c) plotting the frequency with which brands are associated with different spaces on the Brandscape map.

High Praise for the 3R's Paper
In her thought provoking paper “Introducing the 3R’s: Rebalance, Recreate and Reenergise”, Tiina Raikko, Consumer Insight Director at Unilever, invited delegates to consider the value of investing more time and resources into the research delivery process to better engage and gain buy-in from the research audience(s).

As suggested by Tiina, good communication is “as much about the delivery as content” presenting the application of a recent ideation project undertaken with the assistance of WhatIf and Two Blind Mice (referred as Project CrackerJack) where the aim was to generate ideas of ways in which to foster better information absorption and learning in the research delivery process.

Suggested by the paper is the need to focus on the 3R's of research, specifically spending more time at the outset on how to make research findings live and breath (i.e. ‘Rebalance’ our planning), using different interactive components to build learning over time (i.e. ‘Recreate’ the debrief) and capturing interest in the research to effectively compete for audience attention (i.e. “Reenergise”)

George Camakaris Award for Best Paper by a Young Researcher
Researchers under the age of 30, or in their first five years of their career, are eligible for the George Camakaris Best Paper Award. George Camakaris was the founder of Quantum Market Research, which sponsors the award and is now helmed by Adrian Goldsmith.

The Young Researcher Best Paper Award was inaugurated to encourage young researchers in the industry to develop and grow and to aspire to reach new heights attained by George Camakaris during his career. Jayson Chaplin from Galileo Kaleidoscope won the award for his paper 'Using Blogs in Commercial Market Research'.

Jayson defined blogs "...as an online self reflection of the blogger and his/her thoughts about themselves and/or the world around them. So rather than a collection of well writ-
ten articles or essays, what is written on a blog is meant to be the spontaneous, to the point, truthful and honest opinions of the blogger”. What separates blogs from simple online diaries is that bloggers “…continually reference other blogs and websites when writing about a particular subject matter, current affair or any other stimulus of interest…” effectively creating a collective consciousness. However, the plethora of blogs varying from the common to the most obscure has made blog research especially challenging for those new to the blogosphere. However, Jayson pointed out that “market researchers can rely on bloggers to be at the forefront of global marketing activity and consumer thought” and can effectively navigate the blogosphere by identifying the role and limitations two genres of blogs: personal and journalistic blogs.

Journalistic blogs are excellent filters of information and extremely fast to respond to news and current events from around the world. The use of multiple sources of information and the contribution of millions of bloggers as well as the overall connectivity of the blogosphere makes monitoring journalistic blogs not only entertaining, but important for researchers and clients that want to keep an ear to the ground. With most journalistic blogs being thematic in nature and commonly referencing other blogs in the same field, mining specific information becomes less arduous once a researcher becomes acquainted to the topic within the blogosphere. For a market researcher, the true utility of journalistic blogs rests in their content which constantly surprises, informs and entertains their readers. In summary, journalist blogs can be used as a powerful desk research tool that can be used at the proposal stage or to bring further insight and thought to the analysis process.

Personal blogs are a window to understanding “intrinsically truthful reflections of the blogger”. The blogosphere world is mainly dominated by teenagers and young adults (usually known as Generation Y) where they have populated websites such as www.blogger.com, www.livejournal.com and in particular www.myspace.com to create their own blogs mainly to express themselves, network and create friendship groups who understand them and their interests. Myspace.com in particular is a useful tool for understanding specific teenage tribes or youth groups. Other personal blogs in general are an excellent source of secondary data to help create “richer brand positionings” by using bloggers emotional language to further articulate consumer insights. Whilst bloggers are not a representative sample of the general online population, bloggers “…are never the less very articulate and can help researchers better articulate what previous respondents could not”. The challenge with personal blogs is that they “require researchers to be more discerning and imaginative when trying to use them for research purposes”.

**In closing**
In summary, the topic of Imagination was the catalyst for an upbeat conference showcasing the contribution market and social researchers can make to business, government and society.
Interested in becoming an AMSRS member?

The Australian Market & Social Research Society Ltd (AMSRS) is a not-for-profit professional membership organisation established to serve those practising or interested in market, economic and social research.

We have a national membership of over 2,000 researchers, marketing executives, market research managers and buyers, field managers, recruiters, interviewers, academics and students.

Being a member benefits those people:
- involved in or interested in market, economic, advertising or social research
- employed in the marketing or research sections of an organisation
- conducting qualitative and quantitative research
- who supervise field operations, interview or recruit participants for research projects
- who teach or study in marketing or social and market research

Membership entitlements include:
- discounted member rates to AMSRS and affiliate events
- a copy of the Annual Directory & Yearbook, and Research News
- regular updates on state and national industry information, developments and events
- a membership card and membership certificate
- Full members are entitled to use the postnominals MMSRS (Member of the Market & Social Research Society), vote in divisional and national elections of the AMSRS and receive the Society’s Journal AJMSR.

Note: Student members receive limited membership benefits so please see our website for full details.

All members are bound to observe the AMSRS Code of Professional Behaviour, which covers both ethics and standard conditions of conducting and reporting scientific research.

Contact us at:
Membership,
Australian Market & Social Research Society
Level 1, 3 Queen Street
Glebe NSW 2037

Phone: +61 (0) 2 9566 3100
Or check out Membership on our website at www.amsrs.com.au
A Problem In Questionnaire Design
Philip Gendall and Janet Hoek
Department of Marketing, Massey University
Palmerston North, New Zealand

Introduction
Respondents’ interpretation of survey questions depends not only on the words and sentences in them, but also on their formatting and presentation (see Dillman, 2000; Jenkins and Dillman 1997; Redline and Dillman 2002; Redline et al. 2003; Christian and Dillman 2004). Drawing on concepts from cognitive psychology and visual design, Dillman and his colleagues have developed principles for effective self-completion questionnaire design, including principles of question formatting and questionnaire layout. These principles, or ‘rules for revision’ as Dillman et al. (2005) have called them, are based on the natural reading and comprehension processes people use to interpret visual information. This research note illustrates what can happen when these questionnaire design principles are overlooked, something that occurred in the 2003 International Social Survey Programme survey on national identity.

The International Social Survey Programme (ISSP) involves leading academic institutions in 40 countries in an annual survey of economic and social policy issues. Each year the ISSP member countries carry out a 30-minute survey of the general public using the same source questionnaire. The data from these surveys are deposited in a central archive in Cologne, where they are freely available to all members. This collection of data enables researchers to examine similarities and differences between countries, and to monitor changes over time. (For more details on the ISSP see http://www.issp.org)

A drafting group, typically comprising members from five or six countries, develops the questionnaire for each ISSP module. The drafting group circulates a draft of the proposed questionnaire (written in British English) to the other ISSP members prior to an annual drafting meeting. At this meeting, members debate the content of the questionnaire, make appropriate changes, and agree on a final version of the questionnaire. This process of questionnaire design by committee produces questionnaires that are inevitably flawed (like all questionnaires), but which, nevertheless, are as successful as most. This reflects the careful development work done by drafting groups, subsequent adjustments after piloting, and the considerable experience of ISSP members in questionnaire design and survey research.

Most of the organisations in the ISSP administer their surveys face-to-face, only seven countries, including New Zealand, use self-completion questionnaires. Generally, this mode difference has few, if any, implications for the ISSP survey questionnaire. Interviewer-administered questions can readily be converted into self-completion format, and vice versa. However, in 2003 the design of one of the ISSP questions had unfortunate consequences, at least in New Zealand.

National Identity 2003
The ISSP module in 2003 was National Identity, a replication of a previous survey on the same topic conducted in 1996. Under ISSP rules, two-thirds of a replicated module must remain the same; the remaining third can include new questions. In 2003, the first question in the National Identity module was a new one, concerning the characteristics most important to respondents in defining themselves. This was a ranking question; respondents were asked to select, from a list of ten characteristics, the one most important in describing who they were, then the second most important, and then the third. The question is reproduced in Figure 1.
Figure 1
Question 1: ISSP National Identity 2003
We are all part of different groups. Some are more important to us than others when we think of ourselves.
1. In general, which in the following list is most important to you in describing who you are? And the second most important? And the third most important?

**PLEASE TICK ONE BOX IN EACH COLUMN**

<table>
<thead>
<tr>
<th>Most important</th>
<th>Most important</th>
<th>Second most important</th>
<th>Third most important</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Your current or previous occupation (or being a homemaker)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. Your race or ethnic background</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. Your gender (that is, being a man or woman)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. Your age group (that is, young, middle age, old)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. Your religion (or being agnostic or atheiest)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f. Your preferred political party, group, or movement</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g. Your nationality</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>h. Your family or marital status (that is, son/daughter, mother/father, grandfather/grandmother, husband/wife, widower/widow, not married, or other similar)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>i. Your social class (that is upper, middle, lower, working, or similar)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>j. The part of New Zealand that you live in</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Of the 1,034 New Zealanders who returned valid questionnaires only 358 answered this question correctly. That is, only 35% of the 1,034 respondents who completed the rest of the questionnaire correctly followed the instructions for this question. Many respondents ticked more than one box in each column, some ticked several boxes in two columns and none in the third (and the boxes ticked were not necessarily in the first two columns or even in the first column), and a few respondents ticked all of the boxes in the first column and none in the second or third. In all of these cases, the responses had to be recorded as 'missing' (because there was no way of determining which single attribute respondents considered most important, second most important, and so on).
In other ISSP countries that use self-administered questionnaires response to the first question in National identity was mixed. In Sweden, 53% of responses were coded as non-valid or missing. In Norway, the question was modified by including only a single column of response boxes in which respondents were asked to write in the numbers 1, 2 and 3 to indicate the most important, second most important and third most important characteristics. This format reduced the proportion of incorrect answers to around 20%.
Figure 2

Australian Version of Question 1: ISSP National Identity 2003

We are all part of different groups. Some are more important to us than others when we think of ourselves.

1. In general which in the following list is most important to you in describing who you are? And second most important? And the third most important?

PUT THE LETTER OF THE STATEMENT IN THE APPROPRIATE BOX BELOW

A. Your current or previous occupation (or being a homemaker)
B. Your race or ethnic background
C. Your gender (that is, being a man or woman)
D. Your age group (that is, young, middle age, old)
E. Your religion (or being agnostic or atheist)
F. Your preferred political party, group, or movement
G. Your nationality
H. Your family or marital status (that is, son/daughter, mother/father, grandfather/grandmother, husband/wife, widower/widow, not married, or other similar)
I. Your social class (that is upper, middle, lower, working, or similar)
J. The part of Australia that you live in

Most important
Second most important
Third most important

In Australia, the format of the question was changed in that country's survey by including three response boxes below the list of items, labelling the items A, B, C, ... J, and asking respondents to write the letter referring to the most important characteristic in the first box, and so on. This question had only 3% to 5% missing or non-valid responses. The question is reproduced in Figure 2.

What happened in New Zealand?
Analysis of the 676 'missing' cases in the New Zealand survey revealed that only eleven respondents (less than 1% of the whole sample) had failed to answer the question. The large majority of 'missing' cases (90%) were actually cases in which respondents had chosen different combinations of all 10 responses. Another 6% had chosen combinations of eight or nine of the 10 response categories. In other words, virtually all the cases classified as 'missing' were not missing at all.

Logistic regression analysis was used to examine the characteristics of respondents who answered the question correctly and incorrectly. The dependent variable had the value of 0 if the respondent did not answer the question correctly and 1 if the question was answered correctly; the independent variables were age, sex and years of education. The results are shown in Table 1. The only significant correlates were years of education and age. The sex of the respondent had no effect on whether the question was answered correctly, and none of the interactions between age, sex, and years of education were significant.

Years of education had the largest effect on whether respondents answered the question correctly, and the odds of correctly answering the question were 24% higher...
Table 1. Logistic Regression Results

<table>
<thead>
<tr>
<th>Predictors</th>
<th>B</th>
<th>S.E. B</th>
<th>Wald</th>
<th>Sig</th>
<th>Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-0.15</td>
<td>0.04</td>
<td>12.1</td>
<td>.001</td>
<td>0.985</td>
</tr>
<tr>
<td>Education</td>
<td>0.217</td>
<td>0.025</td>
<td>73.2</td>
<td>.001</td>
<td>1.242</td>
</tr>
<tr>
<td>Constant</td>
<td>-2.780</td>
<td>0.459</td>
<td>36.7</td>
<td>.001</td>
<td>0.062</td>
</tr>
<tr>
<td>-2 log likelihood</td>
<td>1199.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z</td>
<td>126.3</td>
<td>2 d.f., p=.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nagelkerke R2</td>
<td>.16</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample size</td>
<td>1011</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

for every year of education. Conversely, a one-year increase in age decreased the odds of answering the question correctly by 1%. These relationships are reflected in the median values of years of education and age for the two groups of respondents. Median years of education and age for those who answered the question correctly were 16 and 45, respectively, compared to 13 and 53 for those who did not; that is, older respondents and those with less education were more likely to answer the question incorrectly. One interpretation of these results is that the source of the problem was failure to understand the instructions, based on the assumption that this is more likely among older people and those with less education (a finding reported by Redline and Dillman 2002).

Cognitive Testing
To examine this possibility, depth interviews were conducted with a small convenience sample of older members of the general public, recruited in a local shopping plaza. Each respondent was given a shortened version of the original questionnaire (reduced to the first two pages of questions and a page of demographics). Participants were asked to complete the questionnaire and then, if they incorrectly answered the first question, asked why they answered as they had and whether they had read and understood the instructions.

All of the participants interviewed answered the ‘problem’ question incorrectly, and all answered in the same way; namely, by ticking one responses box in each row. Discussions with these respondents revealed a consistent explanation: the respondents looked at the question and assumed they were supposed to treat each statement as a stand-alone question. Most did not read the instruction to "PLEASE TICK ONE BOX IN EACH COLUMN", or, if they did, they ignored it because they expected to have to tick a box in each row.

The following quotes are indicative of participants’ comments:

“I expected to put one answer in each row because this is how questions are normally asked. This is what the rest of the questionnaire was like. I thought this is what I needed to do.”

“I thought this was the best way to answer the question. I treated each statement as one question, not the whole set of statements as one question.”

1. Though both education and age are measured in years, the ranges and standard deviations of the two variables differ. To make the coefficients for the variables more comparable the percentage change in odds for a one standard deviation change in the variable concerned can be calculated. The resulting coefficients show that a standard deviation unit increase in education increases the odds of answering the question correctly by approximately 100%. For age, a standard deviation unit increase decreases the odds of answering the question correctly by approximately 20%.
"I read each statement as a stand-alone question. This is how I expect to see questions. I didn’t notice the instruction that was there and I’m not sure how the question would work otherwise."

"I read each question as a separate question. Each one looks like a question."

This behaviour by respondents is an example of pre-attentive processing (or ‘top down’ processing), which involves the automatic interpretation of visual features and patterns based on expectations and knowledge of the way in which the world works. One manifestation of this phenomenon of pre-attentive processing is that survey respondents often assume they are supposed to answer every question in a questionnaire (Jenkins and Dillman 1997).

In this case, the question concerned occupied most of the first page of the questionnaire. At the bottom of the page was another question that asked respondents "How close you feel ...?" to four different areas (their town or city, province or region, New Zealand, or the South Pacific). For this question respondents were expected to rate each region on a four-point 'closeness' scale. It appears that many respondents to the National Identity survey looked at the first page of the questionnaire, assumed that both questions on it were rating questions, since that is what they both looked like, and answered both accordingly. In retrospect, it is relatively easy to see how the first question, with its ten statements and three-column, matrix format could have been misinterpreted in this way.

This explanation means most respondents must have grouped the response boxes for the first question horizontally rather than vertically and ignored the question's specific instructions. The Grouping Law of Proximity, which states that when similar figures are located close to each other we tend to see them belonging to the same group, provides an explanation for the former (see Dillman and Jenkins 1997, pp. 173 - 174), while, as Dillman notes, the latter is not uncommon:

"...I am convinced that many respondents do not read the content of questionnaires in a thoughtful way. Respondents take cues from the layout about what must be read and what can be safely ignored, and some respondents skip many words, with the frequent result that questions get misinterpreted".

(Dillman 2000, p. 81)

Dillman's explanation for this behaviour is based on an understanding of the nature of visual perception (see Jenkins and Dillman 1997). When respondents perceive a questionnaire two processes are involved: 'bottom up' processing and 'top down processing'. First, respondents see an external stimulus in the form of a questionnaire page (bottom up processing), then they interpret what they see in the context of what they already know about questionnaires or forms, and their expectations of what this means (top down processing). If respondents expect to answer every question, as cognitive testing suggests many of them do (Jenkins and Dillman 1997), or expect to tick a box on every row, this is what they will often do, regardless of any instructions to the contrary.

An interesting speculation is whether this problem in questionnaire design could have been solved without changing the basic format of the question, by manipulating the verbal, numerical, graphical and symbolic language used to present it. Several studies have shown that the visual layout of questionnaires influences whether people respond correctly to instructions (for example, Christian and Dillman 2004; Redline et al. 2003). Thus, there may be other ways, apart from the
radical reformatting done in Australia, of successfully manipulating elements of the question so that most, if not all, respondents would treat it as a ranking question as intended. These modifications would have to be tested empirically, but what we know about pre-attentive processing suggests they might still be less effective than the unambiguous formatting of the Australian version.

There is one other issue that is raised by this 'problem' question. In Table 2, the responses to the question of which is the most important characteristic in defining the respondent are compared for those who answered the question correctly and those who did not. Because respondents in the latter sample usually ticked more than one box in each column, two measures are reported for them: percent of cases and percent of responses.

The different ways in which respondents answered this question inevitably paint slightly different pictures about the characteristics that are important in defining people. The 'select one' interpretation emphasises family or marital status as the most important characteristic; nearly 50% of respondents chose this as most important when only one characteristic was selected. Then there is a large gap to occupation, gender and nationality as the other main determinants of identity.

By contrast, in the alternative interpretation, while family or marital status is still clearly ranked first, nationality is ranked second, ahead of occupation and gender, and race or ethnicity is elevated to fifth, ahead of religion or age group. Because the two samples may not be equivalent, the only conclusion that can be drawn is that the results are different; however, it seems likely that even with identical samples the two questions would produce different answers.

The relevance of this is that even if questions 'work' as intended, this does not guarantee the assumptions embodied in them are valid. While it may be convenient for researchers to assume there is one main determinant of self identity, for example, and respondents may be

---

**Table 2. Comparison of Responses by Question Interpretation:**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Intended Interpretation Most Important (n=358)</th>
<th>Alternative Interpretation Most Important (n=645)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% cases</td>
<td>% cases</td>
</tr>
<tr>
<td>Family or marital status</td>
<td>46.9</td>
<td>85.6</td>
</tr>
<tr>
<td>Occupation</td>
<td>13.7</td>
<td>59.4</td>
</tr>
<tr>
<td>Gender</td>
<td>12.0</td>
<td>58.4</td>
</tr>
<tr>
<td>Nationality</td>
<td>10.1</td>
<td>67.4</td>
</tr>
<tr>
<td>Religion</td>
<td>7.3</td>
<td>27.3</td>
</tr>
<tr>
<td>Age group</td>
<td>3.1</td>
<td>25.6</td>
</tr>
<tr>
<td>Race or ethnicity</td>
<td>2.8</td>
<td>52.9</td>
</tr>
<tr>
<td>Location in New Zealand</td>
<td>2.2</td>
<td>46.5</td>
</tr>
<tr>
<td>Social class</td>
<td>1.1</td>
<td>16.9</td>
</tr>
<tr>
<td>Preferred political party</td>
<td>0.8</td>
<td>16.9</td>
</tr>
</tbody>
</table>
willing to provide it, this does not necessarily mean this is how respondents see themselves. This is a general observation, not one confined to this particular question, but it is illustrated by the fact that the question concerned was interpreted in two different ways, each of which produced plausible, but different, answers.

Conclusions
Over nearly 20 years the questionnaires used in the ISSP have proved to be reliable, robust instruments for data collection. Like all questionnaires there are sometimes problems with individual questions; the first question in the 2003 National Identity module is a case in point. In its unmodified form this question was incorrectly answered by large numbers of respondents to self-administered versions of the ISSP questionnaire.

In New Zealand the problem was not that respondents failed to answer this question, but rather that most of them did not answer the question as intended. Instead, these respondents treated it as a rating task and answered a different question; “How important is this particular characteristic in defining who you are?” They did this in spite of verbal instructions asking them to rank three of the ten characteristics.

The cause of the problem in questionnaire design in the 2003 ISSP survey was the failure by many respondents to comprehend and comply with the question instructions. It seems almost certain that these respondents misinterpreted the non-verbal languages of the questionnaire and ignored, or did not even read, the verbal instructions. This is consistent with an increasing body of research that clearly shows that whether questions are read, the order in which they are read, and the meaning they convey to respondents depends at least as much on how questions are displayed as on the words in them. The redesigned version of the question used in the Australian survey demonstrates that changing the format of the question, by restricting the response boxes to only three (and thereby clearly indicating to respondents that was this was not a rating question) would have solved this problem.

However, another principle of questionnaire design is to let respondents tell you what is in their minds and not to impose your own values or perceptions on them. The fact that the high levels of invalid responses to the National Identity question seen in the New Zealand and Swedish surveys were not manifested in the Norwegian and Australian versions does not contradict this maxim. Imposing a view of the world on respondents does not mean they necessarily accept it or that it reflects their reality. In other words, modifying the format of this troublesome question may have only concealed a more fundamental problem with the question.

The failure of the first question in the 2003 ISSP survey in New Zealand emphasises the importance of visual elements in self-completion questionnaire design. Written instructions may be clear, but there is no guarantee that respondents will read them; they are just as likely to draw their cues about what is expected of them from other elements of design—numbers, symbols, brightness, colour, shape, and location of information—assumptions based on their past experience, and a quick assessment of what they see. The task of questionnaire designers is to manipulate these non-verbal questionnaire languages, the design elements under their control, to create cues to elicit the desired behaviour from respondents and avoid experiences like the one reported here.
REFERENCES


Seeking Insight Beyond Packaged Analysis

Associate Professor Don Porritt
Senior Fellow
IMIA Centre for Strategic Business Studies
An Affiliate Institute of Charles Sturt University

Abstract
Most market research analysis is limited to exploring relationships between pairs of variables, and sometimes looks further at the joint impact of two or three variables on an outcome measure. When more sophisticated multivariate techniques are used, these are mostly limited to models that assume linear relationships, or test for simple interactions. Additional insights can often be achieved using quite simple cross-tabulation methods, but applying them in unusual ways guided by an emerging theory about what is behind the data. Two examples are presented demonstrating the additional insight that can be achieved by thinking beyond the analyses provided by standard cross-tabulation and statistical packages – one from analysis of data on perceptions of company reputations, the other from analysis of customer satisfaction with the experience of buying a new car. The analysis of reputation data provided quantitative confirmation of a theory based on qualitative work suggesting that what some see as excessive profits, gained at the expense of customer, staff and communities, produces a reputational backlash. The analysis of satisfaction with the purchase experience demonstrated that the primary determinant of satisfaction was how the purchaser was treated by the initial sales person. Service problems experienced after a satisfactory initial experience reduced satisfaction but rarely produced outright dissatisfaction. However, even heroic efforts to recover from an initially unsatisfactory experience at best somewhat ameliorated the dissatisfaction of the customer. None of these effects would have emerged without some theorizing about the causal processes behind the data, theories that then guided the subsequent analyses. The paper concludes by drawing a number of lessons about effective data analysis for achieving insight beyond the packaged procedures of stub-and-bag cross-tabulations and more sophisticated multivariate analyses.

Study 1: The Bottom Line Backlash Effect
Australian research (Porritt, 2005) has revealed that a company reputation for financial success can adversely affect overall reputation. This "bottom line backlash" can result in additional hostility to companies that are seen as making large profits at the expense of other stakeholders, especially where these other stakeholders are seen as having limited choice of supplier.

The research, conducted in three large general population samples and a fourth sample of small business managers, found strong support for the following propositions about company reputations:

1. The overwhelming 'driver' of overall reputation is the Emotional Appeal of a company, a combination of liking for, trust in and admiration and respect for the company.

2. Other more specific aspects of company reputation define two major dimensions:
   Relationship Reputation (RR), a composite of items covering how well the company is seen as meeting the requirements of its customers (covering product quality and value, service, and reliability), its staff, and of the wider community (including environmental responsibility and effects on local communities where the company operates); AND

Bottom Line Reputation (BLR), centering on a reputation for sustained financial performance, and a range of attributes that are seen as contributing to such success, including exploiting opportunities, and having a clear future vision – the characteristics that allow a company to directly satisfy requirements of capital markets.
3. The better a company's Relational Reputation, the stronger its Emotional Appeal and overall reputation.

4. The effect of a company's Bottom Line Reputation on its Emotional Appeal and overall reputation is more complex, and depends on the strength of its Relational Reputation. Figure 1 illustrates the pattern.

As Figure 1 shows, if RR (Relational Reputation) is at least average, then the stronger the company's BLR (Bottom Line Reputation) the stronger its Emotional Appeal (and consequently, its overall reputation).

What happens if the company's Relationship Reputation is below average? If Bottom Line Reputation is very low, then the company will have very low Emotional Appeal and overall reputation. If BLR is slightly better, this improves its Emotional Appeal. However, further gains in Bottom Line Reputation do not improve Emotional Appeal or overall reputation, and can even reduce the company's appeal (as in the RQsm Australia 2001 data shown in figure 1).

This effect was originally found in the RQsm Australia 2001 general population study then confirmed by reanalysing data from the RQsm Australia 2000, and replicated in the RQsm Australia 2002 and Small Business RQsm Australia 2002 studies.¹

The Analysis Process
This brief summary of the effect hides the actual process by which it was discovered. Although Figure 1 makes the effect very clear, using a very simple tabular approach, finding the effect was not so simple.

The initial attempts to make sense of the "drivers" of perceived reputation applied multiple regression techniques. The established 20 items of the RQsm survey had been grouped, based on qualitative work, into six conceptual dimensions – Emotional Appeal, Products and

**Figure 1: Illustration of the Bottom Line Backlash Effect**

![Diagram](image-url)

(Adapted from Porritt, 2005, Figure 2)
Services, Workplace Environment, Social Responsibility, Financial Performance, and Vision and Leadership. Each of the conceptually-based dimensions was measured by a composite of three to five rating items. Exploration of US data had developed a model in which the primary driver of overall rated reputation (assessed on a single, 7-point rating scale) was clearly the Emotional Appeal dimension, formed of a composite of three ratings (trust in, admiration and respect for, and liking of the company rated). Further attempts to understand what produced differences in perceived reputation then had used multiple linear regression methods with either the conceptual dimension composite scores, or the primary individual rating items.

Both in US data, and in the RQsm Australia data for 2000 and 2001, regression analysis produced clear enough models for the major drivers – items to do with the performance of products and services, followed by items tapping aspects of social responsibility appeared to be the primary drivers of Emotional Appeal. However, the items included in the Financial Performance and Vision and Leadership dimensions produced inconsistent results. They clearly added to the variance accounted for, but every data set produced a different combination of items as the "best predictors". Further, there were always one or two that, despite having positive correlations with Emotional Appeal, showed significant negative regression weights.

It appeared that there might be some inconsistencies produced by collinearity between the items. However, qualitative inputs suggested that something more interesting and complex might be behind the confusing linear regression results.

The first attempt to resolve the issue with the Australian data explored the structure of the predictor items using factor analysis. This revealed the two reputation factors with very similar structure in both the 2000 and 2001 samples, separating perceptions of relationships with customers, staff and communities on the one hand, from business outcomes on the other. Simple linear regression models using estimated factor scores showed clear cut results – Relational Reputation was the primary driver of Emotional Appeal and hence of rated reputation, with a significantly but relatively small additional contribution from Bottom Line Reputation.

Comments from respondents on the reputation of the company rated suggested that there might be a more complex process at work – that there was an expectation that companies should not profit at the expense of customers, staff and communities, and that doing so could damage rather than enhance reputation.

The first attempt to test this with the quantitative data assumed that the "natural" structure was for all aspects of a company's reputation to be relatively consistent. This could be assessed by using a measure that simply summed the scores on the two factors, to give a "total positive reputation" score. Inconsistency in reputation was then measured by taking the difference between the scores on the two factors. In effect, the factors were rotated by 450. Non-linear regression was then applied for each of the derived scores (Total Positive Reputation and Inconsistency of Reputation), and linear, quadratic and cubic functions tested. In both the RQsm 2000 and 2001 samples, a strong and purely linear relationship was found to Total Positive Reputation, and a highly significant quadratic relationship emerged for the Inconsistency of Reputation composites.

Further thought suggested that this rather complex analysis might be revealing something that could be demonstrated much more simply, by taking into account the possibility that scores on Relationship Reputation might act as a moderator variable, determining the form of the relationship between Bottom Line
Reputation and Emotional Appeal. This lead to the analysis shown in Figure One – essentially a simple cross-tabulation combining groupings on the two independent variables (Relationship Reputation and Bottom Line Reputation).

No pre-packaged analytical technique would have been likely to discover this pattern. The analysis using Total Positive Reputation and Inconsistency of Reputation came close, but did not detect the interaction. What lead to identifying the form of the interaction was reflecting on the causal processes that might produce the response to “inconsistent” reputational performance – that doing well financially could become a reputational negative if achieved, as seen by the respondent, at the expense of stakeholders the respondent would identify with, including customers, staff, and communities. While identifying how to demonstrate this effect took many different attempts at analysis, the final step was relatively simple.

Fombrun and van Riel (2003) report results from structural equation modelling on RQ data from the USA which essentially assumes that relationships are linear and do not interact. On this basis, the Vision and Leadership and Financial Performance dimensions appeared to add little or nothing to the variation accounted for by the relational dimensions. We strongly suspect that the Bottom Line Backlash effect is not unique to Australia. However, until data from other cultures is analysed using the simple, but somewhat unconventional steps developed to understand the Australian data, it seems unlikely that the effect will be replicated. The combination of a moderator variable (Relational Reputation, RR) controlling the form of the relationship between Bottom Line Reputation and Emotional Appeal, with the appearance of a linear relationship when RR is above average, and a curvilinear relationship when it is below average simply cannot be found by any ‘standard’ analysis. The analysis has to be guided by some theory about the underlying processes, and tailored to test the theory. In practice, the theory did not emerge full-blown and ready for the crucial test against data – the process, as described above, was one of trying a variety of analysis methods, returning to think some more about the underlying processes, and then trying a different analytical approach.

Did this process reveal useful insight into the data and the real-world processes that the data are intended to measure? We believe it did. Instead of finding that, for the general public (and even for managers of small businesses) reputation for business performance makes little difference to a company’s Emotional Appeal and perceived overall reputation, we found that it could substantially enhance both, so long as it was achieved by meeting the needs of customers, staff and communities. We further found that a poor business performance reputation could further damage the Emotional Appeal and overall reputation of a company that was not meeting the requirements of other stakeholders, and that a reputation for high business performance could inhibit Emotional Appeal if seen as being gained at the expense of other stakeholders. This adds to understanding of both the data and the causal processes. Bottom Line Reputation matters much more than linear models that ignore interactions between drivers can reveal. To further demonstrate the value that can be gained from these insights into the data, consider briefly some of the implications.

**Action Implications**

The Bottom Line Backlash effect raises a number of questions including:
- how does recognition of sustained financial strength damage the attractiveness of a company to the general population, and even to small business managers?
• how much does reputation, and Emotional Appeal in particular matter? Should boards be concerned if their company has a poor reputation in the general population, if its financial strength is recognised by shareholders and financial markets?

Media discussion of large profit announcements reveal a widespread disapproval of profit that is seen to emerge at the expense of customers, staff and communities. The impact is especially evident for companies that provide what many regard as “essential services”, such as banking and telecommunications. This suggests how the “bottom line backlash” adversely affects overall company reputation:
• high profits are seen as acceptable if achieved by meeting the needs of all stakeholders;
• high profits are seen as unfair if the profits are seen as being achieved at the expense of customers, staff and communities, and so produce no emotional benefit or a reduction in approval;
• many consumers and small business managers accept high profits if these are seen as the result of meeting the requirements of customers, staff and communities, but they believe that such profits should not be achieved by cutting quality, convenience, value, reliability or service to customers, treating staff badly, or acting in ways that disadvantage the environment and the community in general.

Another process might also contribute to the backlash effect. Where consumers see themselves as having little choice, either because a company is a near monopoly or dominates its market (as was the case of Telstra), or because the major suppliers are seen as all following similar unfair policies (as with the major banks), the sense that there is “no way out” might produce greater sensitivity to the way the company is seen as achieving its profits.

Does the backlash matter? Should boards faced with such a backlash seek to reverse it?

Research has not yet settled the vexed issue of the relationship between social responsibility and corporate financial performance (e.g., Griffin and Mahon 1997; Roman et al. 1999; Sabate and Quente 2003; Ullman 1995). However, social responsibility continues to be a key aspect of corporate reputation in the public mind. Further, there is evidence that corporate reputation is an important corporate asset. For example, Fombrun and van Riel (2003) have shown that companies with a stronger reputation:
• are better able to attract customers, investors and quality employees;
• are better able to survive crises that would destroy weaker firms.

“In fact, reputation and financial value are related in three ways. First, reputation affects the operating performance of a company and therefore its profitability. Second, profitability affects market perceptions of the company’s future prospects – and so influences the level of demand for a public company’s shares, that is its market capitalization. Third, the company’s operating activities themselves contribute to building ‘reputation capital’ – a shadow asset whose value encompasses the intangible equity hidden in both a company’s product brands and corporate brand, and that describes the positive regard in which it is held by all of the company’s stakeholders. Positive regard, in turn, attracts people to work for and invest in the company – and so increases profitability.” (Fombrun and van Riel 2003, p26). Similar conclusions are reached by Alsop (2004) and Jackson (2003) from analysing exemplary cases and by Dowling

When major competitors are absent, are weak, or suffer the same reputational problems, a board can perhaps ignore these processes – at least until competitors emerge. While there are those like Milton Friedman who argue that it is ethically legitimate, and indeed morally required to disadvantage other stakeholders if this delivers profitability to stakeholders (Phillips 2003), many others challenge the ethical integrity of such a position (eg, Jackson 2003; Phillips 2003). Even ignoring the ethical issues, considering the growth of competition for the major banks from smaller (but fast growing) regional banks, the entry of overseas financial institutions like GE Money and ING, and Bendigo Bank’s profitable entry into exactly those local communities that the major banks are seen to have abandoned and let down, it would appear unwise for their boards to not pay attention to these issues. The speed at which Telstra is now losing market share in some of its markets might be in part due to limitations on its capacity to compete on price; however, as its effective monopoly position is weakened, the company’s board and management must face increasing demands to raise service levels, and to provide and ever-expanding range of ‘essential services’ to customers as mandated by the Government’s ‘local presence plan’ (Kohler 2005).

How these companies can regain reputations for meeting customer, staff and community requirements without damage to their bottom lines is not clear. The need to address the issues is, however, becoming ever more acute. All companies must also face one additional unpalatable fact: it takes much longer to rebuild a reputation of caring for non-financial stakeholders than it does to lose or damage such a reputation (Fombrun and van Riel 2003).

Study 2

This study was based on analysis of proprietary customer satisfaction data collected by a major automotive manufacturer in Australia. The exact quantitative data are not reported here, but the process of analysis and some general indications of the findings further illustrate the value of going beyond standard approaches to analysis. Again, the final findings can be presented using relatively simply cross-tabulation analyses, but were discovered by reflecting on anomalies in more complex multiple regression analyses, following application of factor analysis to some of the data.

The data were very large samples of responses to a customer satisfaction questionnaire completed three months after purchase of a new car by personal buyers. The questionnaire included:

- a checklist of initial sales representative behaviour, based on the manufacturer’s standards for desirable sales representative behaviour in greeting, listening to and responding to a prospective customer’s requirements
- ratings of overall satisfaction with the vehicle use experience.
- whether problems had been experienced with the vehicle and if so, whether these had been fixed the first time, after two or more service visits, or remained uncorrected.
- whether the dealership had contacted the buyer to check on their experience of the vehicle and on whether there had been problems, and the outcome.
- ratings of overall satisfaction with the purchase and ownership experience to date (3 months after purchase), willingness to purchase the same make of car, and to continue use of the same dealership in making future purchases.

The purpose of the analysis was to discover what were the drivers of satisfaction and intention to consider the manufacturer and the dealer for future purchases.
The first step was to check the correlations between the items in the checklist of sales representative behaviors. These were all similar and relatively high for simple dichotomies, so a factor analysis was conducted. This showed that the items formed a single factor, with all items loading substantially (over 0.40). Consequently, a simple scale formed by counting how many items were endorsed showed very high internal consistency as measured by the coefficient alpha statistic.

This suggested that the basic experience of interaction with the sales representative was a fairly undifferentiated one leading to a more or less positive overall attitude. The distribution of scores showed that most of the checklist items were endorsed at quite high levels, and that it was rare for more than three items to not be endorsed.

Regression analysis was then used to test how the Sales Representative Performance score and satisfaction with the vehicle, the dealership, any experience of problems, and the response, contributed to willingness to consider a vehicle from the manufacturer, and to consider purchasing through that dealership in the future. Not surprisingly, some quite strong relationships were found. By far the strongest contribution was made by the Sales Representative's performance. This might partly be because problems with the vehicle were relatively rare, and when experienced were usually promptly fixed to the buyer's satisfaction.

However, it was noted that the ratings were highly skewed. As usual with customer satisfaction data, the vast bulk of respondents were satisfied, with a substantial minority highly satisfied on the five point rating scales used. Even the scores on the Sales Representative Performance scale were skewed, with the majority being described as performing all, or all but one of the designated behaviours, and with a small proportion missing four or more.

This suggested that linear regression might be inappropriate for the data. It was also thought desirable to be able to demonstrate relationships within the data using simpler cross-tabulation analyses that the client, and dealers, would be better able to understand.

The sample was thus split into levels on the Sales Representative Performance scale, and relationships of the satisfaction ratings, experience of vehicle problems, outcomes of any problems, and contact by the dealer to the intentions ratings were explored separately and in combination.

Some very powerful non-linear effects emerged.

In summary, where the Sales Representative was described as performing all, or all but one, of the designated desired behaviours, almost all buyers were satisfied with the ownership experience, willing to consider the make and the dealer for future purchases.

The picture was remarkably different if the initial Sales Representative had not performed so well. If three or more of the designated behaviors were not endorsed, then satisfaction and willingness to consider the make and the dealer for future purchases were generally much lower, with very few ratings of "very satisfied" or "very likely to use". This remained true, even if there were no problems with the vehicle, and if the dealer had followed up to ensure that the customer was happy. If there were any problems, even if these were promptly resolved to the buyer's satisfaction, overall satisfaction was rarely above the "quite satisfied" level, and willingness to consider the make and dealer for future purchases was rarely above the "might use" level. If vehicle problems were not resolved, ratings were extremely negative.
These patterns are summarized in Figure 2, using a combined “outcome index” based on the overall satisfaction and willingness to use in future scales.

What was the effect of problems with the vehicle, and whether these were fixed first time, and of contact from the dealer?

Problems with the vehicle, especially if not fixed first time, reduced satisfaction with the ownership experience, and the overall outcome. However, for those who reported a good initial contact experience, there had to be serious and unresolved problems with the vehicle to drop their outcomes into the “bad” range. Conversely, lack of problems, or prompt correction of problems at best moved outcomes into the neutral/mixed range, and rarely produced good outcomes if the initial contact was below standard. A similar pattern was found for contact by the dealer.

The implication were clear. It appeared that the initial experience shaped how buyers responded to everything that happened thereafter. If this initial experience was relatively poor, using a quite “objective” checklist of Sales Representative behaviour, even a trouble free experience of the vehicle and efforts to keep in contact with the buyer and ensure that all was well could not fully recover the relationship. If the initial experience was good, problems with the vehicle did not have so severe an impact on outcomes, and even if problems emerged, outcomes rarely became poor.

Again, these effects were only detected because analysis went beyond standard approaches. Close exploration of how the data combined to provide a picture of the buyer’s experience over time, and awareness that responses to items about early stages of the experience could alter the final impact of what
happened later in terms of problems with the vehicle, and further contact with the buyer was required. Apart from the factor analysis that justified scoring the Sales Representative behaviour checklist as a single variable, the final analysis actually used simple cross-tabulations, guided by an exploration using CHAID (chi-square automatic interaction detection). Linear multiple regression analysis did not detect the interaction effects. As with the reputation data, the analysis was guided by, and influenced thinking about, an emerging theory of the role of buyer expectations and attitudes.

It is of course possible with retrospective data that buyers who were unhappy were revising their memory of Sales Representative behaviour to be consistent with their subsequent experience and attitude. However, it appears more reasonable to conclude that this initial experience creates either a buffer of relationship capital that can cushion the relationship against some subsequent problems, or a disposition to find fault and be unhappy which translates into intense dissatisfaction if something further goes wrong. “Doing the right thing” at a later point can move an initially satisfied customer to being a delighted, committed advocate, but for a customer who is initially unhappy it can at best recover some ground to produce a customer who is grudgingly willing to accept offers of additional exemplary service.

Conclusions
The two studies that have been described, and reflection on the experience of developing the final analyses, illustrate a number of key points.

First, the “truth” that can be revealed by data analysis might not be immediately obvious, and can often be obscured by both standardized cross-tabulations, and by routine application of more sophisticated data analysis techniques.

Second, developing a theory about the causal processes that have produced the data can lead to more incisive analysis that can allow “truth” to emerge that can be demonstrated by relatively simple, but non-standard analyses. The emerging truth will often reveal non-linear interactions between variables, and relationships that do not follow straight lines.

Third, the process of theory development and data analysis is, in practice, iterative. The relevant and applicable theory will often not be obvious in advance of the data. The appropriate form of data analysis will often emerge in steps, perhaps motivated by dissatisfaction with the results or complexity of interim analyses, and being forced to reflect on the meaning of the data and the processes that the data emerge from.

Fourth, while the final presentation of the data might be relatively simple (as illustrated in Figure 1), it takes time, repeated revisiting of the data, and refinement of relevant theories about what is happening, to arrive at such relative simplicity.

Fifth, along the way, always reconsider your assumptions about such matters as the temporal sequence and causal linkages revealed in what might appear to be cross-sectional, single occasion data; consider whether the combined effect of two variables might be quite different from the simple sum of their separate effects; and think about whether there might be relationships that do not follow simple straight lines.
REFERENCES


ACKNOWLEDGEMENTS

The ROsm Australia studies are conducted by AMRinteractive Pty Ltd under licence from Harris Interactive. Study 1 data used in this paper were supplied by AMRinteractive Pty Ltd.

ROsm is a registered Service Mark of Harris Interactive.

For further information about the ROsm Australia studies contact: insights@amrinteractive.com
Investigating the Contextual Requirements of the Juster Scale

Mathew Parackal and Ron Garland
Dr Mathew Parackal is Lecturer in the Department of Marketing, University of Otago, Dunedin, New Zealand
Dr Ron Garland is Associate Professor in the Department of Marketing, Waikato Management School, University of Waikato, Hamilton, New Zealand
All correspondence to be directed to Dr Mathew Parackal. Postal Address: P.O. Box 56, Department of Marketing, University of Otago, Dunedin 9051, New Zealand.
Email: mparackal@business.otago.ac.nz
Telephone: 0064 3 479 7686

Abstract
Researchers have employed the Juster Scale to collect purchase probability data with notable success. Reviewing the Juster Scale studies, however, has revealed that there is considerable variation in its performance. Some of these variations appeared to be caused by the context in which the Juster Scale has been presented to respondents. This paper discusses three factors that influence the context of the Juster Scale and reports the results of a study that attempted to standardise its contextual requirements.

The results substantiate further the Juster Scale's satisfactory performance in collecting purchase probability data.

1. Introduction
In survey research, the poor performance of intention scales to forecast purchase behaviour has led researchers to test probability scales. The earlier studies that compared probability scales with intention scales found the former to produce more accurate forecasts (Day et al. 1991; Gan, Esslemont and Gendall 1985; Pickering and Isherwood 1974; Clawson 1971; Gruber 1970; Juster 1966; Ferber and Piskie 1965). Since then, investigators have tested an eleven-point scale pioneered by Juster (1966) and found it to produce satisfactory results (Gan et al. 1985; Gruber 1970; Juster 1966). The Juster Scale (Figure 3 & 4) as it is known in the academic literature has been customised for use in self-completion questionnaires (Gendall, Esslemont and Day 1991); telephone surveys (Brennan, Esslemont and Hiri 1995) and Internet-based surveys (Parackal and Brennan 1999).

Versions of this scale have been successfully tested to forecast purchase rates (Gendall et al. 1991; Gan et al. 1985; Gabor and Granger 1972; Clawson 1971; Juster 1966), purchase levels (Brennan et al. 1995; Brennan, Esslemont and U 1995; Brennan and Esslemont 1994; Seymour, Brennan and Esslemont 1994; Hamilton-Gibbs, Esslemont and McGuinness 1992), mutually exclusive behaviours (Parackal and Brennan, 1999; Hoek and Gendall, 1997), demand schedules (Brennan 1995) and customer loyalty (Garland, 2002; Danenberg and Sharp, 1999; 1996a; 1996b). In recent years, the scale has been employed in choice modelling studies (Rungie and Danenberg 1998) and in modelling repeat purchases using the Dirichlet model (Wright, Sharp and Sharp 2002).

2. Literature review
The Juster Scale studies cited above were successful in achieving their respective objectives. Nonetheless, there were considerable variations in the accuracy of the scale across the studies mentioned. For example, forecasts of automobiles made on the Juster Scale were reasonably accurate (Juster 1966; Stapel 1988; Pickering and Isherwood 1975; Gan et al. 1985) but forecasts of other durables were not so accurate (Brennan et al. 1995; Pickering and Isherwood 1975; Clawson 1971; Heal 1970; Juster 1966). A critical examination of the Juster Scale literature revealed that some of the variations could have been caused by the context in which
the scale was set. In the following section the Juster Scale literature is reviewed from such a viewpoint.

2.1 Question-order effect
The literature on context suggests that questions asked previously influence the contexts of those asked subsequently, and in turn influence the responses collected (Lorenz, Robertson and Lesser 2001; Schwarz and Hippler 1995; Schuman 1992; Johnson, O’Rourke and Severns 1998; Schuman and Presser 1981). Such contextual effects appear to have occurred in the earliest Juster Scale studies. These studies were carried out by including the scale in omnibus surveys (Brennan 1995; Brennan and Esslemont 1994; Gan et al. 1986) and by piggybacking on other studies (QIS studies by Juster 1966; Byres 1964). Consequently, investigators have had very little control over where the Juster Scale and its question appeared in the questionnaire. The placement of the Juster Scale in a questionnaire was not a concern (see Kalton and Schuman 1982) for a discussion on placement of questions in a questionnaire, but the questions asked before the Juster Scale were a concern, as they could influence the context and consequently the responses collected.

Brennan (1995) attributed such contextual effect to the illogical estimates made for an innovation in two successive omnibus surveys. The Juster Scale was included in the 1994 and 1995 Household Omnibus Survey conducted by the Department of Marketing at Massey University, New Zealand to collect purchase probability data for laser disc players, a product for viewing digital video discs (DVDs). The product was introduced in 1994; one would expect its adoption to have increased in the following year (1995). This, however, was not reflected in the adoption rates estimated in the two omnibus surveys. The estimates for the two years were in the reverse order, with adoption in 1994 being greater than in 1995. In the 1994 questionnaire, the question about laser disc players was placed after a series of questions about video and video stores, whereas in the 1995 questionnaire there was only one item on this topic. It appears that the questions on videos and video stores gave the 1994 questionnaire a “video hiring and viewing context”. The question on laser disc players was the same in the two questionnaires except for the overshadowing context of video hiring and viewing. The illogical adoption rate estimated for this product was attributed to the context (Brennan 1995).

Similar contextual effects (in this case a question-order effect) could be present in the results of studies implemented in other omnibus surveys (Brennan 1995; Brennan and Esslemont 1994; Brennan et al. 1994; Gan et al. 1986) and piggybacked on other studies (Juster 1966; Byres 1964). Thus, some of the variations in the performance of the Juster Scale observed in those studies could be due to contextual variation arising from question-order effects.

2.2 Item-order effect
It has been a common practice amongst investigators to test the Juster Scale concurrently on different product categories (e.g. Day et al. 1991; Brennan and Esslemont 1994; Brennan et al. 1995; Hamilton-Gibbs et al. 1992; Seymour et al. 1994; Clawson 1971; Gabor and Granger 1972; Gan et al. 1986; Gruber 1970; Heald 1970; Juster 1966). While there is no contention as to the purpose of this practice, listing test items of different categories one below the other could give rise to order effects. To differentiate this type of order effect from the previous one (question-order effect), it has been labelled “item-order effect”.

Test items listed at the start could set the context and tone of response for those listed lower down in the order. Consequently, responses to items listed lower in the order could suffer item-order effect and may not be reflective of the corresponding behaviours. This may have been the case in three studies that reported forecasting errors based on recall data (Day et al. 1991; Gan et al. 1986; Clawson 1971).
In these studies, the Juster Scale was concurrently tested on three product categories (durables, services and fast moving products). Forecasting errors reported across these studies ranged from -17% to +245%. It was not possible to conclusively establish whether the variations observed were because of item-order effect. All the same, it is plausible to argue that item-order effect was responsible for some of the forecasting errors observed in these three studies.

Similar to item-order effect is response-order effect (Redline et al. 2006; Krosnick 1999; Ayidiya and McClendon 1990; Krosnick and Alwin 1987; Shuman and Presser 1981). It refers to respondents’ tendency to select the first listed response category, known as primacy effect, or the last listed response category, known as recency effect. The literature on these two types of response-order effects reveals that they are associated with the way response categories are presented, either visually (as in the case of a self-completion questionnaire) or orally (as in the case of telephone and face-to-face interviews), to respondents. Primacy effects have been observed in visual presentation whereas recency effects were observed in oral presentation (Krosnick 1999; Krosnick and Alwin 1987). The theory behind primacy effect is that the first response category on a list establishes a cognitive framework of comparison that directs the interpretation of subsequent categories on the list (Krosnick and Alwin 1987). Recency effect results from the availability of more cognitive processing time for the last response category on the list (Krosnick and Alwin 1987). When response categories are presented orally one after the other, the presentation of a subsequent category terminates the cognitive processing of the previous one. In general after the last category is presented, respondents tend to receive more processing time as there is a longer pause to receive a response before moving on to the next question. In this case the cognitive framework tends to be around the last items. The same sorts of ramifications are plausible in the three Juster Scale studies (Day et al. 1991; Gan et al. 1986; Clawson 1971) discussed above, except they are for test items listed one after the other. Thus, some of the errors observed in the three studies may be because of item-order effect.

2.3 Question Wording
In Juster Scale studies, respondents are encouraged to consider all factors that influence their purchase behaviour before indicating their purchase probability scores. For this purpose, many investigators have adopted the question formulated by Juster (1966) (“Taking everything into consideration what are the chances of you buying a < >?”). Other question versions were also used in conjunction with the Juster Scale in the literature (Brennan et al. 1995; Urban et al. 1996; Day et al. 1991; Gan et al. 1885). Forecasting errors across the studies cited ranged from -5% to +245%. The studies mentioned above employed the same approach of getting respondents to indicate their purchase probability data using the Juster Scale. They, however, have used different question versions and some of variations in the performance of the Juster Scale could be because of the versions used.

Studies that investigated question versions have observed considerable variation in the interpretations given (Gendall 1998; Schuman and Presser 1977). Turner and Krauss (1978) observed that interpretations influence the context of the question and the response collected. In survey research, respondents have to understand and interpret the questions appropriately. For this reason, questions are pre tested to ensure there is reasonable consistency in the way respondents understand and interpret them (Gendall and Hoek 1990; Belson 1986). While investigators testing the Juster Scale may have pre tested their versions, because of differences in the wordings there may have been differences in interpretation between the Juster Scale studies mentioned above. Contexts could also vary with interpretations and consequently result in these studies not being strictly comparable.
Discussions so far have raised question-order effect, item-order effect and question versions as potential factors that influence the context and the responses collected on the Juster Scale. These factors seem to exist in most Juster Scale studies and their influence on the context and the forecast obtained is plausible. If this is the case, then Juster Scale contexts could be varying across the studies and results may not be comparable as was seen in the mainstream survey contextual literature (Schuman and Presser 1981). The Juster Scale would require fresh testing to produce comparable results to assess its reliability to forecast purchase behaviours. The seriousness of this matter warrants a systematic investigation of the contextual factors raised in this literature review. First it was seen necessary to standardise the contextual requirements of the Juster Scale. The following sections outline the research objective, describe the methodology, and provide a discussion on the results.

3. Research Objective
The objective of the research reported in this paper was to find out whether the Juster Scale required standardisation of context to collect purchase probability data. This objective was formulated into the following hypothesis for statistical testing:

H1: The Juster Scale implemented without providing contextual information produced purchase probability data that was similar to when implemented after providing contextual information. (Mean purchase probability scores obtained in the versions were similar).

To test the above hypothesis, quantitative data were collected through questionnaire versions implemented in separate treatments. The versions were made different by presenting the Juster Scale with and without providing contextual information. Internet technology was used to supply contextual information in the relevant version; hence, the questionnaire was applied through two Internet-based surveys. Mean scores obtained on the Juster Scale in the versions were compared for statistical difference. The hypotheses were tested in two separate samples (one for a sample of business clients, one for a sample of the general public, labelled respectively as Business and National survey). In the National survey, information on search behaviour was collected allowing additional analysis.

4. Questionnaire versions
A questionnaire on Wireless Application Protocol (WAP-capable) mobile phones, new generation mobile phones introduced in the market at the time of this research, was developed and used in this research. This questionnaire was made into separate versions to generate the required comparisons. In the first version, the Juster Scale was presented without providing any information about the test product. This version was the standard against which comparisons were made, therefore will be referred to as “Standard” in this paper. Respondents were given the following simple descriptive of the test product before being presented with the Juster Scale questions:

“WAP-capable mobile phones can access the Internet and download emails and Web pages. We would like to find out what you think about this idea of WAP. Click the "Next" button when you are ready...”

In the second version, the Juster Scale was implemented after respondents had had the opportunity to view information about WAP capable mobile phones. Information was supplied via a web page by listing hyperlinks to relevant websites. By clicking on the hyperlinks, the corresponding web pages loaded onto respondents' computers ready for viewing (Figure 1). When respondents had viewed the information, they were presented with the twelve and six months Juster Scale questions shown in Figure 3 and 4. This version is called “Point & Click” for the sake of simplicity.
Figure 1: Point & Click

Information about WAP mobile phones

Please,

- click on the side arrow (↑) to expand each heading to expand the list.
- select items from the expanded list by double clicking the links for viewing.

When you've finished viewing the information please return to the questionnaire via the "Continue Survey" link placed on top of the list.

PS
- If you are using Opera to view this page, please click here to take you to a page that is compatible with Opera.

Figure 2: Search Engine

Information about WAP-capable mobile phones

Please type in a keyword into the text box and click the "Search" button on your left.
- Please select items from the search results that will be listed below the search engine for viewing by double clicking the links.

When you've finished viewing the information, please return to the questionnaire via the "Continue Survey" link placed above the search engine.
Point & Click was an adaptation of the approach used by Urban et al. (1997), Urban, Weinberg and Hauser (1996) and Urban, Hauser and Roberts (1990) to provide contextual information about test products. In their approach, respondents were brought into a lab and provided with computers to view contextual information. Information items were listed on the computer screen and respondents selected items for viewing using a mouse. On completion of viewing, respondents were presented with the Jister Scale and the associated questions. The same format was used in the current research with the adaptations made for an online environment.

To verify whether listing information items (see Figure 1) had any effect on the purchase probability scores given, a third version of the questionnaire was included in the comparison. In this version, information was supplied via a search engine (Brucke 1985). Respondents were required to enter keywords (e.g., brand, price) into the search engine which then produced a list of hyperlinks to web sites containing the corresponding information. This version has been called “Search Engine” in the discussion that follows.

In all three questionnaire versions, purchase probability data were collected for two time horizons - twelve and six months. The questions accompanying the Jister Scale for the two time horizons were as follows:

- Taking everything into account, what are the chances that you would replace your present mobile phone with a WAP-capable one within the next TWELVE MONTHS, that is up to the end of < > ? (see Figure 3)

- Taking everything into account, what are the chances that you would replace your present mobile phone with a WAP-capable one within the next SIX MONTHS, that is up to the end of < > ? (see Figure 4)
5. Survey approach
As the Internet was used to provide contextual information in the relevant versions, the survey was carried out over the Internet. The research used an approach recommended by Kingsley & Anderson (1998) to undertake probability based Internet surveys. The rationale behind the approach is that Internet users tend to be normally distributed within most general populations. Hence, a simple random sample drawn from a general population would have Internet users in the same proportion and make-up they are found in that population. Approaching respondents by their postal address with a request to participate in an Internet-based survey achieves the recruitment of Internet users into the sample. The approach relies on the innovativeness of individuals to engage in new Internet tasks (Eastlick & Lotz 1999; Well & Chen 1999; Citrin et al., 2000; Goldsmith 2001; Mai & Mai 2002) to participate in the research. As the initial sample was a random selection, the final sample possessed sufficient randomness, and ensured a normally distributed Internet user population. The current research employed this research approach to implement the questionnaire versions in two separate surveys. One was on a simple random sample from a client list of a communication business (Business survey) and the other from the electoral roll of New Zealand (National survey).

The Business survey was carried out on a random sample of 3,400 respondents selected from the client list of a business. The survey produced 460 useable responses. As the survey was on the Internet population, response rate was weighted to the size of this population (52% incidence of households with Internet (Ministry of Economic Development 2003) and was 30%, after the initial number of contacts was adjusted for refusals and ineligibles. This response rate was within the range (9% to 44%) reported for similar surveys in the literature (Schonlau et al. 2001). Subjects were randomly assigned to the three questionnaire versions (n = 167 for Standard; n = 139 for Point & Click; n=154 for Search Engine).

The National survey was carried out on a random sample of 3,000 respondents selected from the electoral roll. In this survey, non-Internet users were offered the option of completing the survey using a paper version of the questionnaire (Quigley et al. 2000). The number of respondents who completed the survey over the Internet was 403 (55% of the total sample) and those completing using the paper version was 326 (45%). The proportion of Internet participants (55%) was comparable to the percentage of households (52%) with Internet connection in New Zealand at the time of the research (Ministry of Economic Development 2003). This observation confirmed that the approach used (perhaps) was successful in reaching the Internet population under investigation. The analysis required for this research was, however, restricted to those who participated on the internet.

As the sample was drawn from the general population it included both mobile phone users and non-mobile phone users. For this research, only the responses of mobile phone users were used. Compliance of respondents to perform the required antecedent activities, that is, to view contextual information, was recorded in the versions that provided this (Point & Click and Search Engine).

6. Results
6.1 Business survey
The mean purchase probability scores collected in the three questionnaire versions exhibited logical progression over the two horizons, that is, adoption rate increased over time (see Table 1). Analysis of variance tests were used to investigate whether the mean purchase probability scores were statistically different in the three versions.
Table 1: Mean probability scores of the 12 and 6 months time horizons

<table>
<thead>
<tr>
<th></th>
<th>12 months mean</th>
<th>6 months mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>0.41</td>
<td>0.26</td>
</tr>
<tr>
<td>Point &amp; Click</td>
<td>0.39</td>
<td>0.24</td>
</tr>
<tr>
<td>Search Engine</td>
<td>0.41</td>
<td>0.24</td>
</tr>
</tbody>
</table>

F = 0.17, p = 0.84
F = 0.22, p = 0.81

Table 2: Mean probability scores of the 12 and 6 months time horizons

<table>
<thead>
<tr>
<th></th>
<th>12 months mean</th>
<th>6 months mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>0.33</td>
<td>0.22</td>
</tr>
<tr>
<td>Point &amp; Click</td>
<td>0.27</td>
<td>0.16</td>
</tr>
<tr>
<td>Search Engine</td>
<td>0.27</td>
<td>0.17</td>
</tr>
</tbody>
</table>

F = 0.96, p = 0.39
F = 1.52, p = 0.22

Table 3: Mean probability scores of the 12 and 6 months time horizons after applying the information viewship criterion

<table>
<thead>
<tr>
<th></th>
<th>12 months mean</th>
<th>6 months mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>0.33</td>
<td>0.22</td>
</tr>
<tr>
<td>Point &amp; Click</td>
<td>0.28</td>
<td>0.17</td>
</tr>
<tr>
<td>Search Engine</td>
<td>0.29</td>
<td>0.20</td>
</tr>
</tbody>
</table>

F = 0.45, p = 0.64
F = 0.62, p = 0.54

The analysis of the twelve months probability data produced an F value of 0.17 with an associated p value of 0.84. The analysis of the six months probability data produced an F value of 0.22 with an associated p value of 0.81. In both cases, the p values did not reach the rejection level (p = 0.05), hence, the evidence collected was not sufficient to reject the hypothesis. Mean purchase probability scores produced in the three versions were similar with differences attributed to sampling error.

6.2 National survey
The mean purchase probability scores collected for the National survey respondents exhibited logical progression over the two horizons that were consistent with those seen in the Business survey (see Table 2). Results of ANOVA for the twelve month-probability data produced a F statistic of 0.96 with a p value of 0.36 (see Table 2) and that for the six month-probability data produced a F statistic of 1.52 with a p value of 0.22 (see Table 2). In both cases, the p values did not reach the rejection level (p = 0.05), thus, the evidence was not sufficient to reject the hypothesis (H1). These results concurred with those of the Business survey.

Compliance status recorded in the National survey revealed that 62% (n = 53) in the Point & Click version and 37% (n = 39) in the Search Engine version performed the precursory task of searching and viewing information. The number of items viewed by these respondents ranged from one to fourteen. Mean purchase probability scores of those who viewed and did not view information within each questionnaire version were compared. For the Point & Click version, an independent sample
test revealed that the means were similar with the p value not reaching the rejection level for both time horizons (12 months: t = 0.09, df = 82, p = 0.93; 6 months: t = 0.34, df = 82, p = 0.73). The same results were observed for the Search Engine version (12 months: t = 0.79, df = 86, p = 0.43; 6 months: t = 1.15, df = 86, p = 0.25). These results confirmed those reported in Table 1 and 2.

Analyses were carried out after removing the responses of respondents who failed to view information in the two experimental groups (Point & Click and Search Engine). The ANOVA for the twelve month-probability data produced a F statistic of 0.45 with an associated p value of 0.64 (see Table 3) and that for the six months-probability data produced a F statistic of 0.62 with an associated p value of 0.54 (see Table 3). These results concur with the earlier observations.

7. Discussion
The research reported in this paper compared the Juster Scale with and without providing contextual information. The variations were made in three separate versions of a questionnaire on WAP-capable mobile phones which collected purchase probability data from two separate samples. Purchase probability data were collected for twelve and six-months time horizons. In total, the research produced six separate comparisons based on analysis of variance (ANOVA) to investigate for statistical differences between purchase probability scores.

All six ANOVA tests produced non-significant results; hence, the evidence was not sufficient to reject the hypothesis (H1). From this, it is concluded that the Juster Scale implemented without contextual information produces forecasts similar to those produced when the scale is implemented with contextual information. This observation implies that the Juster Scale does not require any additional contextual input.

An earlier study enquiring into the extensiveness of describing a new service when collecting purchase probability using the Juster Scale observed similar results (Armstrong 1971). In that study, a brief description utilising one-way communication was compared with a comprehensive description that encouraged two-way communication. Comparisons were made for estimates of demand at various prices, price elasticities, and segment creation. For all comparisons, estimates produced were similar; hence, that study concluded that there was no difference derived from the amount of description provided. The product used in the above mentioned study was a new and expensive transportation service.

The observation made from the research reported in this paper was based on tests carried out on a single test item (WAP-capable mobile phones) that belongs to an Information Communication Technology (ICT) category. The observation, however, was in line with Armstrong's (1971) study. At least now there are two test situations in which comparable results have been obtained. It would be worthwhile seeing if the results hold across other products and product categories.

8. Future direction
As for the contextual issues raised in this paper, investigation can now be carried out using WAP-capable mobile phones as the test product. Question-order effect could be investigated further by comparing questionnaire versions in which the position of the Juster Scale question for WAP-capable mobile phones is varied. Item-order effect could be investigated by included WAP-capable mobile phones along with other test items. The order of the test items could be varied in versions with WAP-capable mobile phones placed at the top and bottom of the list. A control version in which the WAP-capable mobile phone is included on its own would help in establishing the extent to which item-order effect influences the Juster Scale forecasts. Question forms used with the Juster Scale could be investigated by including them in separate questionnaire versions. In each of these versions prob-
ability data could be collected for WAP-capable mobile phones. Such future studies would help in establishing best practices for using the Juster Scale to forecast future purchase behaviour.

9. Conclusion
In this paper, three factors (question-order effect, item-order effect and question version) that exhibited strong tendencies to influence the context of the Juster Scale were raised. They are pertinent to most Juster Scale studies and hence warranted future investigation. To systematically investigate them, it was first necessary to standardise the context of the Juster Scale. The results suggest that the Juster Scale does not require any additional contextual input and remains a satisfactory instrument for collecting purchase probability data. This result has provided a starting point for commencing a systematic investigation into the three contextual factors mentioned above. The WAP-capable mobile phone is recommended for the future investigation into the Juster Scale’s question-order effect, item-order effect and question version.

REFERENCES


Dealing with Information Overload

R. Susan Ellis  
Melbourne Business School  
The University of Melbourne

Abstract
Market researchers both affect and are affected by information overload. This article provides an overview of information overload from both a psychological and managerial perspective. It also links information overload with continuous partial attention (CPA). Finally, guidelines are provided to help market researchers minimize information overload when preparing material and presentations for their clients. Seven specific topics, including non-obvious caveats, are discussed: data summaries, note taking, expertise, repetition, images and sounds, speed control and interaction.

Acknowledgments
Special thanks to Claude Alexidis for raising the topic and some of the specific issues addressed in this paper.

During the [on-stage] presentations the faces of at least half the crowd were lit with the spooky reflection of the laptops open before them. Those without computers would periodically bow their heads to the palmtop shrine of the Blackberry. Every speaker was competing with the distractions of e-mail, instant messaging, Web surfing, online bill paying, blogging and an Internet chat "back channel" where conferences supplied snarky commentary on the speakers. (Levy 2006)

How easy it is for us to access information! And how challenging it is to process and utilise all that we access. Consider these facts.
- The total amount of printed information worldwide is currently estimated at five exabytes (Wikipedia 2006). An exabyte is approximately 1018 bytes.
- The Internet is growing by between one and two exabytes of information per year (Herman 2002).
- The typical person interacts with one-millionth of an exabyte of information each day (Lord and Macdonald 2004).
- The limit of human processing capacity appears to be the concomitant interaction of only four independent variables. A five-way interaction gives performance at chance level (Halford, Baker, McCredden, and Bain 2005).

The foundation of market researchers' contribution to organisations is, of course, information. In effect, ours is a profession that has the potential to make a major contribution to information overload. It behooves us therefore to understand the phenomenon so that we have a chance to manage the situation as best we can to the benefit of our clients and ourselves.

What constitutes information overload?
Information overload is an inability to process all of the information to which we are exposed, because of natural limits on our capability (e.g., Miller 1956, Jacoby 1977). While a precise definition of "information" is context dependent (e.g., Jacoby 1977), one unit of information is loosely parallel to the human equivalent of a computer "bit" (binary digit) which is a 1-0, on-off indication of a switch and corresponds to the presence or absence of a fact. Miller (1956) describes a bit as "the amount of information that we need to make a decision between two equally likely alternatives" (p. 83).

Information overload has been variously referred to as "data asphyxiation, data clutter, data smog, data delirium, data dredge, data swamp, info glut and info pollution" (White and Dorman 2000), all of which emphasise its potentially chok-
ing, chaos-causing capability. There are even disease-like names for it such as information fatigue syndrome, information obesity, and fragmentia (Akin 1998).

The concept of information overload has been discussed in the marketing literature for a number of years, particularly as regards consumer capability to process information (e.g., Jacoby 1977; Malhotra 1984; Hahn, Lawson and Lee 1992; Winzar and Savik 2002; Lee and Lee 2004). While its degree of presence in a particular situation has been contested depending upon the definitions of "information", "overload" and resulting decision "effectiveness", information overload as a general construct is considered valid.

Whether the overload occurs at the time the information is encoded into memory or when it is retrieved has been the subject of some discussion in the psychological literature. It now appears that this is to some extent task dependent (e.g., Yantis and Johnston 1990). [For an excellent interdisciplinary review of the information overload literature, combining research from psychology and information science as well as from various business disciplines, see Eppler and Mengis 2004.]

More immediately relevant in the marketing and market research environments is the fact that, irrespective of when information overload occurs, it can lead to poorer decision-making, because information overload can make it difficult to determine what information is relevant, let alone identify the most important information (Elliott 1988). Not surprisingly, given the volume of market research information available to marketing managers, the area of marketing management has itself been found to have far more information than a decision maker can reasonably be expected to process (Meyer 1998).

Information overload has also been studied in the banking and finance sector, where decisions are more frequently made on the basis of a large quantity of information (Kanaan 1993) and can have potentially substantial financial consequences for the decision maker and/or his or her organisation. Here again, the correlation between information overload and decision quality degradation has been identified. For example, Casey (1980) found, using an experimental design, that bank loan officers who were asked to assess bankruptcy potential of firms took significantly longer to mentally process the information and did not make any better predictions about those bankruptcies when they had more information. Paquette and Kida (1988) found that professional accountants had decreased accuracy in selecting from a set of firms the company with the highest bond rating as the number of firms to evaluate increased.

These findings are consistent with a large body of psychological research that shows that all of us have limits to our mental processing capabilities (e.g., Miller 1956; Halford, Baker, McCredden, and Bain 2005). Too much information can in fact act as a stressor and degrade decision-making, sometimes even resulting in random responses (Franco-Watkins, Pashler and Rickard 2006). To add to that, when the information is presented in a bilingual individual's non-native but still fluent language, information overload tends to occur sooner, and decisions degrade faster (Dolinsky and Feinberg 1986).

**How much information does it take to overload?**

Miller (1956) gave us the "magical number seven, plus or minus two" (p. 81), suggesting that in our short-term memory we can hold a maximum of seven items of information in order to be able to make absolute judgments about them. Because those items can equally be groups or, as he calls them, chunks of items, the actual amount of information that can be held in the short-term store can be significantly greater than seven. Winzar and Savik (2002), in reviewing the
consumer marketing literature, identified previously published research results that showed that consumers could manage somewhere between 4 and 10 brands, but at 11 or 12 brands there was information overload. Similarly, for product attributes (e.g., size, colour, flavour), somewhere between 9 and 16 attributes created overload, depending on the particular situation. Hahn, Lawson and Lee (1992) found that it was time pressures that created the information overload effect. When time was not a critical element, they did not find the same degradation in decision quality.

But the reality is that consumers, customers, clients, managers and market researchers are all time poor. With time being of the essence, it does not appear to take much to trigger information overload.

Is it data or is it information?

Devin (2001) describes data as "the raw material out of which information is obtained" (p. 21). More completely, the definition in SQLServer.com (2005) is that information, which is any meaningful stimulus, becomes data once it is put into a computer, and then it re-emerges from the computer as information.

And this leads to a bit of a debate about the existence of information overload. Bill Gates says that there is no such thing as information overload and that in fact people and organisations are suffering from information underload in critical areas such as sales and finance (Fried 2005). What he is really saying is that there may be enough raw data but not a system that easily and centrally makes it available as information. (And of course he is championing Microsoft to provide that system.)

At the same time there has been rise in attention deficit disorder in adults (Newman and Parker 2001). There is some speculation that the dramatic rise in available information and connectedness to the world has been at least partly to blame.

How can you test for information overload?

In a seminal paper on human information processing, Miller (1956) analysed data from a series of experiments requiring absolute judgments about physical phenomena. The researcher kept giving the subject additional information until, as he described it, the subject began to exhibit confusion. At that point, Miller postulated, the subject was getting close to reaching his capacity for information.

Later researchers have refined the process, now looking for an inverted U-shaped curve of decision quality rather than "confusion". Now established in the literature is what is called the information overload effect, the situation in which decision quality first increases with increasing information and then decreases as information continues to increase (Hahn, Lawson and Lee 1992). This means you can actually examine a series of decisions to see if their quality first increases and then decreases as information continues to increase, that is, if decision quality follows that inverted U-shaped curve. If it does not, that is, if it decreases monotonically as information increases, then you may not be dealing with an information overload effect but rather with task difficulty or complexity.

It has generally been understood that "only a subset of simultaneously occurring events capture our attention at one time" (Bernays and Wcisto 1994). However, there now exists the phenomenon of continuous partial attention (CPA), a phrase coined by Linda Stone of Microsoft in the 1990s. While the term was originally intended to apply to people's ability to be doing a number of technical activities all at once, the term has come to represent a generational ability to scan a large amount of the environment but in a fairly shallow manner. Stone says, "It's not the same as multitasking; that's about trying to accomplish several things at once. With continuous partial attention,
we’re scanning incoming alerts for the one best thing to seize upon: “How can I tune in in a way that helps me sync up with the most interesting, or important, opportunity?” (Maxwell 2002). It is semi-tasking (Kirwan-Taylor 2006) rather than multi-tasking. Friedman (2001) says CPA occurs because “we are constantly scanning the world for opportunities and we are constantly in fear of missing something better”. Lord Maurice Saatchi (2003) gives CPA, and its distraction from paying full attention to the television when it is on, credit for reducing television commercial day-after recall by more than two-thirds since the 1960s.

If CPA is an attempt by people to forestall information overload, then it is possible that as CPA activities increase the individual is moving toward information overload. When even CPA is no longer sufficient to manage the environment, then overload can certainly be said to have been reached.

How can we deal more effectively with potential information overload?

Market researchers have a number of possible ways to address information overload. Individual researchers who find themselves overloaded with information could, for example, work to improve their time management skills, including setting and maintaining of priorities. At the firm level, market research departments or market research firms could work on improved technological solutions such as integrated databases, intelligent agents, “what if” software, and easy to access on-line libraries.

But where market researchers can make at least as great an impact is with their clients, whether internal or external. The clients, being either decision makers or closer in line to the decision makers, may find that they are faced with an increasing number of non-routine, complex and ill defined tasks. It is precisely in those situations that information overload can most easily occur. Market researchers can help ease the overload burden by creating structure, focus and aggregation of information.

There are seven major ways that the potential for information overload can be managed. While a number of these are more obvious because they are already being used by market researchers, some are less so. And a number of them have caveats, or “gotcha’s”, of which to be aware. Dealing with information overload is neither straightforward nor necessarily intuitively obvious.

1. Data summaries
2. Note taking
3. Expertise
4. Repetition
5. Images and sounds
6. Speed control
7. Interaction

Data summaries. Probably the most commonly used technique by market researchers is to provide summaries of data rather than (just) the raw data. Cherny and Dickson (1974) found that decision makers who utilised summarised data (descriptive statistics but without any interpretation) made better decisions than those who utilised raw data. However, and this is an important “however” - those same individuals who utilised summarised data took longer and had less confidence in their decisions. Thus there appears to be a trade-off between ease of working with the information and speed/confidence.

Under those circumstances, how might decision makers try to increase their confidence? They might do so by asking their market researchers either for additional summarised data or by asking to see the raw data so they could perform their own analyses. Certainly we have all seen market research clients do that on occasion. What can result then is what we might term overload seeking
behaviour to improve confidence and/or lessen perceived risk, even if it results in impoverished decisions.

You may find that, under certain circumstances where a future decision is perceived to be riskier than usual, certain of your colleagues or clients – or even yourself – may seek far more information than is really needed to make an appropriate decision. And the decision may be no better – in fact, it may be worse – than with less information. Being aware of this phenomenon is part of the solution. Client risk aversion and internal politics however may interfere with the ability to fully put such knowledge into action.

Note taking. Another fairly obvious way to minimise information overload is through note taking. Horowitz and ForsterLee (2001) found that mock juries that were allowed to take notes throughout the trial – and allowed to refer to those notes during jury deliberations – regarding a complex case on chemical dumping were better able to match appropriate awards to worthy plaintiffs. Further the jury members believed that overall their jury operated more efficiently. In contrast, the research found that having access to complete trial transcripts on a computer was not as effective for jurors as was individual note taking. This suggests that active rather than passive involvement allows for more effective processing of large quantities of information.

Unfortunately, market researchers’ desire to make life easier for their clients, resulting in the now nearly ubiquitous photocopied handouts of research presentations, means that clients take far fewer notes than they used to do. Not only does this provide the opportunity for attention to wander – encouraging CPA – but it can also mean that the quality of client decisions based on the research findings can actually be lowered. It is probably unrealistic to think we can go back to a time when there were no handouts. However, we may wish to consider handouts that are far less detailed than they currently are.

Expertise. Hansen and Haas (2001) found that, at least for internally produced documents, teams that produced fewer documents covering fewer topic areas (thus establishing their expertise in those topics) but with higher quality had their document database accessed more frequently than if they had produced more documents on more topics. In other words, it did not pay to try to be a generalist.

While this has obvious implications for the reputation of a market research firm, it can also have a (negative) revenue impact. This could mean turning work down that does not fall within the firm’s very specific area of expertise. It is not unreasonable to believe that most market research firms would not pursue this path.

Repetition. Another possibility is through repetition of key information, either in a presentation or a written report. How many of us have been taught that the way to do a presentation is “Tell them what you’re going to tell them, then tell them, then tell them what you’ve told them”? Repeating key points, in particular key research findings or implications, is a standard means of communicating in the market research world as in most others.

However, Kronlund and Whittlesea (2005) found that repetition of certain items, while it increased accuracy of those items, also created the impression or illusion that other items had been given twice when in fact they had not. Could this possibly lead to some irritation on the part of the information recipient? Certainly it suggests we be a bit judicious in the use of straight repetition.

Instead we could make information retrieval easier through avoidance of cue overload. The cue overload principle states that the probability that an item can be recalled decreases as the total number of items associated with its functional
retrieval cue increases (e.g., Watkins and Watkins 1975). In other words, information is more readily recalled when some retrieval stimulus is uniquely or nearly uniquely associated with it.

Images and sounds. Meyer (1998) reports that he finds in a number of publications that information shown in picture form is processed both more rapidly and more completely (though he does not cite any specific sources). This is the concept that a picture is worth a thousand words.

However, the same does not necessarily apply to music or sounds. Moreno and Mayer (2000) tested, in experimental design, the addition of background music and/or sound effects to animated instructional presentations on lightning and hydraulic braking systems. In general, groups exposed to material enhanced with appropriate music and/or sound effects performed worse on retention of the information. This suggests that auditory effects can themselves create information overload because they take up some of the processing capacity of the individual. The moral of the story: use sound sparingly. The cute PowerPoint icons with sound effects may need to be retired.

Speed control. People tend to remember exceptions and inconsistencies (e.g., Hastie and Kumar 1979). And when people are allowed to control the speed of intake of information, they are more likely to be able to recall those details that are at variance with the overall picture (Bargh and Tsal, 1985). For market researchers, this can occur in market research reports where not only major findings but also minor exceptions are important for the client.

Unfortunately, in market research presentations it is the presenter rather than the audience who determines the speed of the material. One could argue that providing the client with a copy (hard or soft) for later follow-up at a more leisurely pace would overcome this limitation, but in fact this is trying to overcome a problem after it has been created rather than trying to avoid it in the first place. Impressions have already been formed, and we are all aware how difficult it can be to change an impression once it has come into existence. A potentially stronger solution involves interaction (see below).

Interaction. What is probably less utilised currently in market research presentations is high interactivity between the information and the recipient. Too often a market research audience is essentially passive, with dialogue limited to questions of clarification or to immediate application of selected results. The really good MBA programs around the world have applied adult education techniques to fully and actively involve students in the classroom learning environment, a very different scenario from the undergraduate lecture model. Bringing the voice of the customer into the research presentation - through audio or video clips of interviews or focus groups - has been powerful but still is passive. Developing ways to more totally engage your research clients in your presentation will require another level of creativity. At the same time, this will go a long way toward reducing the negative impact of the large body of information you have to present.

No one said it would be easy. Just as market research seems to be capable of making a significant contribution to the quantity of information available today, so do market researchers have the ability to reduce or ease their clients' information overload - as well as their own. Application of some of the techniques listed above - and being aware of some of the caveats associated with those techniques - may assist. We cannot stop the growth in information, nor would we want to. All we can do is put ourselves and our clients in the best possible position to make sound business decisions from the information we give them.
REFERENCES


Miller, G. A., 1956. The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. Psychological Review. 63 (2), 81-97.


Interested in becoming an AMSRS member?

The Australian Market & Social Research Society Ltd (AMSRS) is a not-for-profit professional membership organisation established to serve those practising or interested in market, economic and social research.

We have a national membership of over 2,000 researchers, marketing executives, market research managers and buyers, field managers, recruiters, interviewers, academics and students.

Being a member benefits those people:
- involved in or interested in market, economic, advertising or social research
- employed in the marketing or research sections of an organisation
- conducting qualitative and quantitative research
- who supervise field operations, interview or recruit participants for research projects
- who teach or study in marketing or social and market research

Membership entitlements include:
- discounted member rates to AMSRS and affiliate events
- a copy of the Annual Directory & Yearbook, and Research News
- regular updates on state and national industry information, developments and events
- a membership card and membership certificate
- Full members are entitled to use the postnominals MMSRS (Member of the Market & Social Research Society), vote in divisional and national elections of the AMSRS and receive the Society’s Journal AJMSR.

Note: Student members receive limited membership benefits so please see our website for full details.

All members are bound to observe the AMSRS Code of Professional Behaviour, which covers both ethics and standard conditions of conducting and reporting scientific research.

Contact us at:
Membership,
Australian Market & Social Research Society
Level 1, 3 Queen Street
Glebe NSW 2037

Phone: +61 (0) 2 9566 3100
Or check out Membership on our website at www.amsrs.com.au