

# DESIGNING BONSAI SURVEYS

THE SMALL BUT PERFECTLY FORMED SURVEY EXPERIENCE TO MEET THE NEEDS OF  
THE MODERN DAY MOBILE-BASED SURVEY CONSUMER

*Jon Puleston*

ESOMAR Office:  
Atlas Arena, Azië Gebouw  
Hoogoorddreef 5  
1101 BA Amsterdam  
The Netherlands  
Tel.: +31-20-664 21 41

Email: [customerservice@esomar.org](mailto:customerservice@esomar.org)  
Website: [www.esomar.org](http://www.esomar.org)

Publication Date: May 2014  
ESOMAR Publication Series Volume S364 APAC  
ISBN 92-831-0273-8

## COPYRIGHT

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system of any nature, or transmitted or made available in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior written permission of ESOMAR. ESOMAR will pursue copyright infringements.

In spite of careful preparation and editing, this publication may contain errors and imperfections. Authors, editors and ESOMAR do not accept any responsibility for the consequences that may arise as a result thereof. The views expressed by the authors in this publication do not necessarily represent the views of ESOMAR.

By the mere offering of any material to ESOMAR in order to be published, the author thereby guarantees:

- that the author - in line with the ICC/ESOMAR International Code of Marketing and Social Research- has obtained permission from clients and/ or third parties to present and publish the information contained in the material offered to ESOMAR;
- that the material offered to ESOMAR does not infringe on any right of any third party; and
- that the author shall defend ESOMAR and hold ESOMAR harmless from any claim of any third party based upon the publication by ESOMAR of the offered material.

Published by ESOMAR, Amsterdam,  
The Netherlands  
Edited by: Deborah S. Fellow

## ABOUT ESOMAR

ESOMAR is the essential organisation for encouraging, advancing and elevating market research worldwide.

With more than 4,900 members from over 130 countries, ESOMAR's aim is to promote the value of market and opinion research in illuminating real issues and bringing about effective decision-making.

To facilitate this ongoing dialogue, ESOMAR creates and manages a comprehensive programme of industry specific and thematic events, publications and communications, as well as actively advocating self-regulation and the worldwide code of practice.

ESOMAR was founded in 1948.

### ABOUT ESOMAR MEMBERSHIP

ESOMAR is open to everyone, all over the world, who believes that high quality research improves the way businesses make decisions. Our members are active in a wide range of industries and come from a variety of professional backgrounds, including research, marketing, advertising and media.

Membership benefits include the right to be listed in the ESOMAR Directories of Research Organisations and to use the ESOMAR Membership mark, plus access to a range of publications (either free of charge or with discount) and registration to all standard events, including the Annual Congress, at preferential Members' rates.

Members have the opportunity to attend and speak at conferences or take part in workshops. At all events the emphasis is on exchanging ideas, learning about latest developments and best practice and networking with other professionals in marketing, advertising and research. CONGRESS is our flagship event, attracting over 1,000 people, with a full programme of original papers and keynote speakers, plus a highly successful trade exhibition. Full details on latest membership are available online at [www.esomar.org](http://www.esomar.org).

### CONTACT US

#### **ESOMAR**

ESOMAR Office:  
Atlas Arena, Azië Gebouw  
Hoogoorddreef 5  
1101 BA Amsterdam  
The Netherlands  
Tel.: +31 20 589 7800

Email: [customerservice@esomar.org](mailto:customerservice@esomar.org)

Website: [www.esomar.org](http://www.esomar.org)

# DESIGNING BONSAI SURVEYS

## THE SMALL BUT PERFECTLY FORMED SURVEY EXPERIENCE TO MEET THE NEEDS OF THE MODERN DAY MOBILE-BASED SURVEY CONSUMER

Jon Puleston

### INTRODUCTION: BONSAI SURVEY DESIGN

This paper will explore the tools and technique needed to design “Bonsai surveys”: short, perfectly formed pieces of market research.

We have reached a crucial point in survey design. We all recognise that surveys must change to accommodate the ever-growing number of people wanting to use mobile phones and tablet devices, and that to compete with many other entertaining online activities, surveys must become more rewarding experiences for people with short attention spans. A new design ethos and a new survey design language have to emerge to deal with these issues.

Table 1 illustrates the challenge we face. The data comes from research undertaken by colleagues at Lightspeed Research in 2013, and shows that only three-quarters of smartphone users are willing to spend more than five minutes completing surveys, with less than a third willing to spend over fifteen minutes. How can we adapt studies that typically run 20 minutes (or longer) to the mobile-device environment?

**TABLE 1.**

MAXIMUM TIME DOING SURVEYS:	COMPUTER	TABLET	SMARTPHONE
5 minutes or less	2%	9%	27%
10 minutes or less	9%	24%	45%
15 minutes or less	19%	42%	65%
20 minutes or less	34%	65%	73%
25 minutes or less	42%	71%	77%
30 minutes or less	65%	81%	85%

Source: US data from LSR proprietary study, Nov 2013; n=1185.

This paper will look at:

1. How to identify the parts of your surveys that tire out respondents
2. How to hunt out and remove redundant questions in your surveys
3. A range of techniques you can use to shorten surveys using statistics and modelling techniques
4. How you can ask questions in more efficient ways
5. Using smarter questioning techniques
6. How to design surveys to work effectively across mobiles, tablets & PC devices

It will also espouse an iterative survey design philosophy, which means adding in more piloting protocols into the survey design process as a means of working out how to slim back your survey.

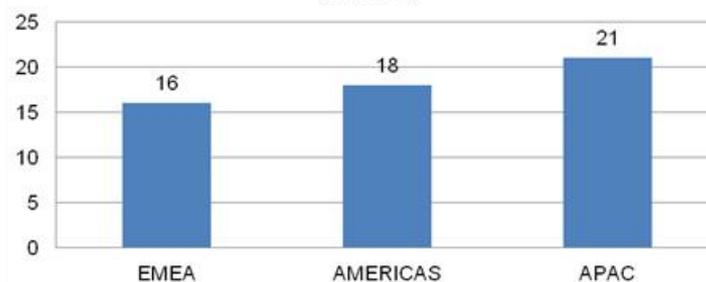
This paper is very much a “how to guide”, based on several years’ experience working with clients around the world, attempting to improve the design of their surveys. I have tried to identify the common problems I’ve seen that affect the length and engagement-levels of surveys, and to advance some ideas on how to deal with them.

I appreciate that much of the content is not going to seem ground-breaking. Many of the techniques I outline are relatively straightforward and might already be well-known to some of you. My aim of this paper is simple: to encourage more of these techniques to be commonly adopted as we have an industry wide survey obesity problem. From my personal experience I would estimate the average survey is made up of about 30% question “fat”, asking redundant unnecessary questions.

I want to challenge you to think about the efficiency of every question you are asking in your surveys.

I am delivering this paper at the ESOMAR Asia conference for two reasons. Firstly, Asia is without doubt leading the way in terms of mobile phone survey uptake. A third more respondents in Asia complete a survey via a mobile or tablet device compared to Europe and North America – around 20% as we speak, with some markets in Asia reaching 25%.<sup>11</sup> Secondly, Asian surveys tend to be significantly longer than the average European online survey by an average of 4 minutes, and suffer from around 10% higher dropout rates as a result, meaning the quality of data is falling behind other markets. (See figure 1.)

**FIGURE 1. AVERAGE SURVEY LENGTHS BY REGION (MINUTES)**



Source: US data from LSR proprietary study, Nov 2013; n=1185.

### THE FIVE KEY STEPS TOWARDS SHORTER SURVEYS

This first part of the paper outlines key steps to optimise the length of your surveys to maximise their effectiveness.

- Step 1: Buy some respondent goggles
- Step 2: Adopt an iterative survey development process
- Step 3: Go hunting for redundant questions
- Step 4: Set some question caps and option level quotas
- Sep 5: Apply some more advanced ascription and chunking technique

#### Step 1: Buy a pair of respondent goggles

If you are looking to re-engineer and shorten an existing survey, I recommend the first thing you do is purchase a pair of respondent goggles.

**FIGURE 2. RESPONDENT GOGGLES**



Putting on these goggles will immediately transform you into one of your own respondents. You should then read through your survey and take note of any questions you think you might get a little bored answering. Then pass the goggles to a few colleagues, friends, or family members, and ask them to do the same thing.

Cross-comparing everyone’s feedback will then give you a clear idea of which questions need work.

This may seem a facetious piece of advice, but putting yourself in a respondent's shoes is the best way to begin to understand how to improve your survey.

#### *Spotting the statistical signs of problem questions*

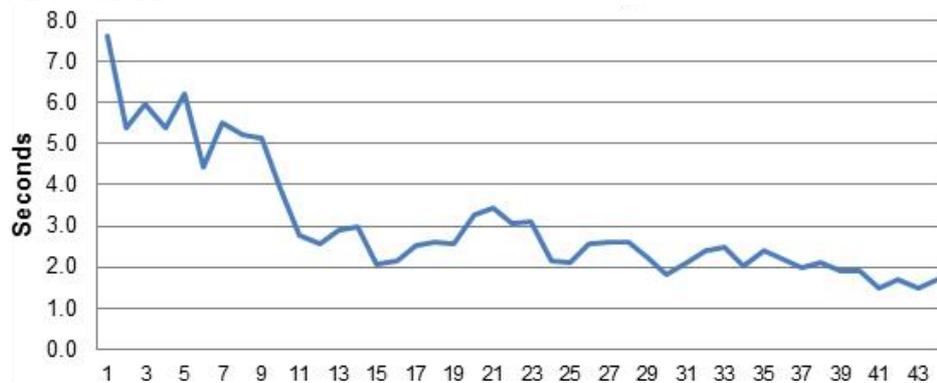
If you want to do this more seriously, and have data from previous surveys to examine, you can look at the common lead indicators of survey boredom. These are:

1. Dropout
2. Straight-lining
3. Reduced answer times

Dropout, particularly mid-survey, anything higher than about 1% per question is something to watch out for. Straight-lining, probably the most famous measure of bored respondent behaviour, which is a lead indicator of a more general problem of pattern answering, is something you need to monitor closely. A straight-lining measure of 8%, while not sounding much on its own, might indicate that up to 40% of respondents are not answering that question properly.

We find analysing average answer time to be the most useful way of understanding respondent boredom. Figure 3 illustrates how rapidly answer time can drop off when respondents are answering a repetitive series of questions. When it's falling below 2 seconds per option you have a problem that you need to deal with.

**FIGURE 3. ANSWER TIME DECAY**



*Tip:* Check if you are recording answer times in your surveys, this is something that most survey software can do but it is often not set up as a default practice, I would advise you to make it standard practice to measure answer time and encourage you to look at the data.

Things to watch out for:

- Higher than 1% dropout on any one question
- Any grid question answered in under 2 seconds per row
- Any multi-choice tick-selection question being answered in under 0.5 seconds per option

#### *The stapler test*

If you want a simple short cut measure to tell if your survey is too long or not print out your survey and try to staple it together and your staple cannot penetrate all the pages of the survey – *it's too long!!!*

**FIGURE 4. STAPLER TEST**



### Step 2: Adopting an iterative survey design philosophy

So, having identified some problem questions and realising the survey is too long, how do you best go about shortening the survey?

Nearly all the techniques to reduce survey length rely on you having some data that you can statistically analyse to make decisions on what questions to remove. So this presents the first problem for many researchers.

If you are trying to shorten an existing tracking survey, or a survey you have run previously, then fine, you're probably sitting on all the data you need. But when creating a survey from scratch you have no data. So how do you get over this problem?

To tackle this you need to adopt an iterative survey design approach, and integrate a pilot phase into the survey design process. This is probably the most critical piece of advice I can offer you in this paper.

Instead of designing the whole survey in one go, you firstly sketch out the main parts of the survey you're unclear about, and run a pilot. This will provide the data you need from which you can identify question redundancies and understand the sample requirements you need for the full study.

I recognise this is a big leap in thinking for many companies, and it does introduce some extra costs and time. We strongly recommend a staged piloting process in almost any survey when you are investing more than \$10k in panel provisions as this can almost be guaranteed to save you some money.

#### Tips for running iterative pilots

- If you are a client, the best thing is to leave this part in the hands of the research company, if you are a research company I suggest you don't involve the client in this part as it will hold things up.
- Don't attempt to set up the whole survey in the pilot, just the key sections you want to test and are unclear about.
- Don't get too involved in detail (such as trawling the script for every spelling mistake).
- Realise you are looking for clues, not rock-solid ±95% statistical certainty. You can normally make do with micro samples of 50-100 for each test cell. (If you don't get enough clarity from this, you can always gather more sample.)

### Step 3: Hunting out the redundant questions

One of the first things to look for in a survey that needs trimming is redundant questions: ones that give roughly the same or similar answers to other questions in the survey. These redundancies are commonly found in banks of attitudinal statements and word-association lists.

There are some sophisticated statistical ways to analyse this, but the simplest test and as useful as any it to run is a basic Pearson correlation which is a standard excel function, comparing the answers of every question with every other in the survey, and look for high levels of correlation. Where any two Likert questions have an answer correlation above R=±0.65, we would normally recommend removing one of them (or rotating them so half the sample sees each one).

If you have any long grid questions in your survey, this type of analysis will almost certainly discover economies. For this reason I would recommend conducting question correlation tests as standard practice on any survey you run.

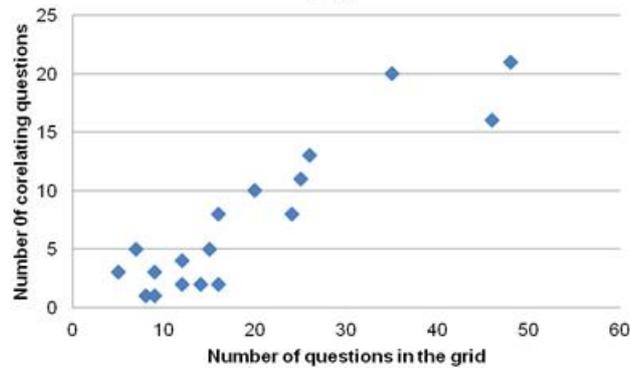
Figure 5 illustrates one visual example to give you an idea of what correlation looks like. It's a question correlation map produced by our DP team. The redder the colour, the stronger the correlation between one question and another. You will notice, for example, the block of red questions in the bottom right corner which are a set of questions all giving very similar answers that could be targeted for editing back.

FIGURE 5



To give some anecdotal evidence of the extent of this type of question redundancy, our data processing team analysed data from 10 random surveys conducted across the Asia Pacific region. In the 18 separate grid questions examined, they found an almost linear relationship between the number of options asked in a grid question and the level of question redundancy. On average, these 10 studies showed that 38% of the questions showed correlation above 0.65 as so could be considered candidates for removal. (See figure 6.)

**FIGURE 6. CORRELATING QUESTIONS**



These simple tests can be effective, but more sophisticated techniques will produce even more worthwhile results. Due to the sometimes non-linear relationship between the answers to two questions, a simple Pearson test will not effectively pick out a relationship – and cross-correlating answers to three or four questions, rather than two, will often make apparent even stronger correlations.

This type of analysis needs a more experienced marketing scientist or statistician and there are a number of companies that specialise in this type of work. This kind of expertise can be very worthwhile in the case of editing back a major tracker, for example.

*Focus on measuring just the things that are statistically worth measuring*

We often see clients demanding that data be gathered on extremely long brand lists, many of which are used by only a small fraction of the population, or “reason for purchase” lists that try to identify every minor factor, many of which, again, only a tiny fraction of the population may select.

This might be seen as being thorough by some clients, but such “thoroughness” is one of the main causes of bloated surveys.

Here is a rule I suggest you invoke – look at how many people in total will be answering your survey, and perform a basic statistical significance test to work out the 95% confidence range. Any activity with an incidence rate that is smaller than the error boundary for 95% confidence should be dropped from your survey.

So these are roughly the error boundaries:

- Sampling 100 = ±6%
- Sampling 200 = ±5%
- Sampling 400 = ±4%
- Sampling 1000+ = ±2%

What this is saying is that for example if you are sampling 200 people, then any brand with less than 5% usage is not statistically worth asking questions about.

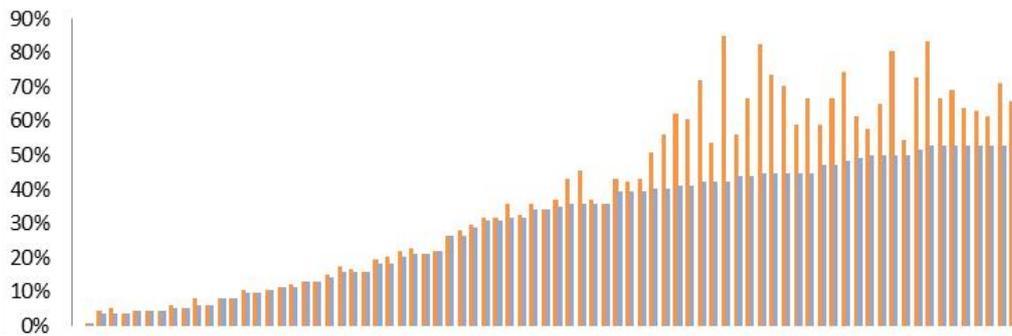
If you are measuring anything below 2% incidence rate be wary – the respondent accidentally tick option is at an error rate of ±1% so half your data points below this point could well be from accidental clicks.

*Tip:* My general advice is be as pragmatic as you can: whittle down the brand list to your key competitive set, and include an “other” option to cover the rest (and please don’t force respondents to write in their other choices into a box – that is a waste of time for everyone as I have never seen anyone use this data profitably).

**Step 4: Setting option caps & question level quotas**

At the opposite end of the scale, with some question options you might have more data than you actually need. Say, for example, you are asking 1,000 respondents about their shopping habits, you are interviewing that many people to reach the user of products in the smaller categories but you only need a maximum of 500 people to answer questions about any one product. A simple technique offered by most survey software solutions is to place a cap on the number of respondents for a particular topic, and then use a “least filled” selection process to ensure future respondents are routed to questions with least data. Figure 7 visualises this.

**FIGURE 7. CORRELATING QUESTIONS**



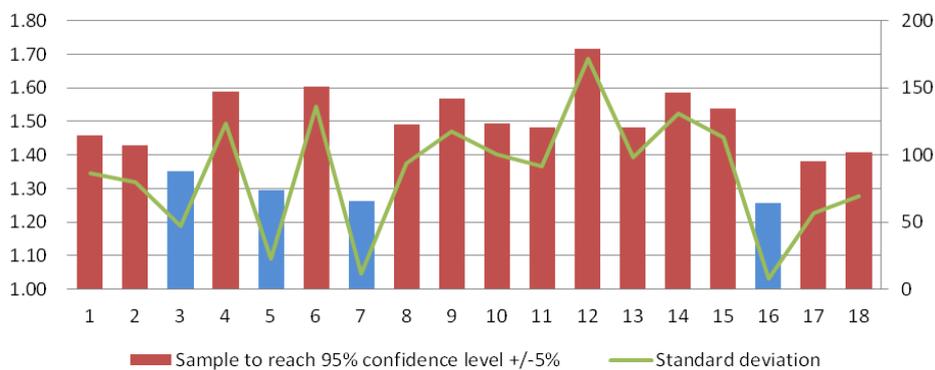
*Setting question-level quotas, not just survey-level quotas*

Another consideration is that some questions need fewer respondents to reach statistical certainty than others.

For example, a binary choice question such as “Which of these 2 ads do you like best?” could achieve 95% statistical certainty with a sample as low as 5 people technically, and very rarely would you need to interview more than 40 to gain a reliable answer. On the other hand, if you were asking about people’s media-consumption habits, and looking to demographically analyse this data, you might need to interview thousands.

Almost every survey contains a mixture of these two types of question, with some needing many respondents and some needing very few. Furthermore if you are asking a set of grid questions you will find that the sample requirements for each option will often vary dramatically too (see figure 8). It all boils down to the level of variation you are seeing in the data. The sample sizes you need for any one question are determined by the standard deviation in the answers and so by knowing this at a question by question level you can move away from setting sample sizes for a whole survey to setting sample sizes for individual questions and by using this approach major savings can be achieved.

**FIGURE 8. ASSESSING SAMPLE SIZES IS BASED ON STANDARD DEVIATION**



*Tip:* Setting a question-level quota cap is best done by a process of randomisation (so 1 in n respondents will answer each question) as this will give an even distribution of answers.

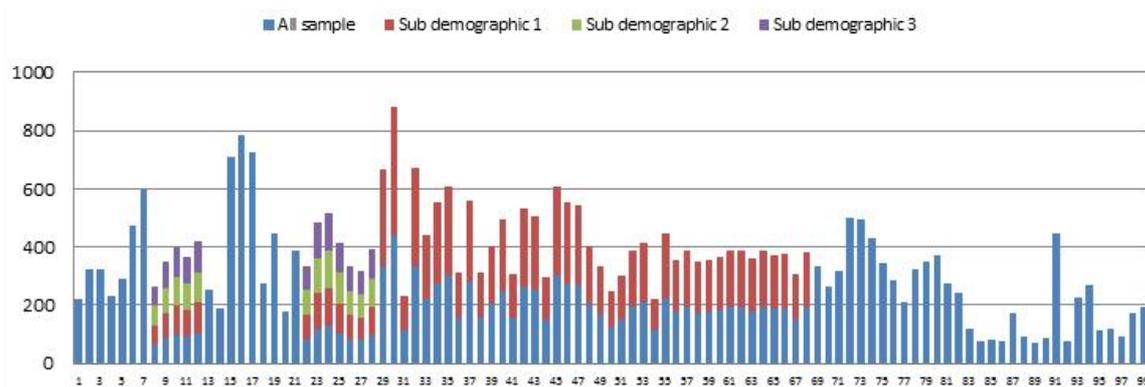
This is the one technique where the greatest level of savings can often be achieved in a survey. Using this technique alone we have often been able to reduce the length of a survey each respondent takes by 30% or more.

### Using a pilot to help set your sample requirements

*Tip:* Don't set in stone your sample requirements until you have run your pilot, it's not until then that you will really understand how many people you need to answer your survey. This is one of the real advantages of running a pilot.

For example, originally the research company had planned to poll 800 respondents but looking at the minimum sample required to get a statistically significant set of answers from each question from the demographic splits they were interested in analysing it was clear that they really only needed to interview a maximum of 600 respondents to achieve this and 80% of the questions could be answered with samples of 400 or less. In fact after the pilot 15% of the questions already had enough data so they did not need to be asked in the main wave. As a result they were able to reduce the length of this survey by 30% and reduce their sample requirements by 25%. (See figure 9.)

**FIGURE 9. SAMPLE REQUIREMENTS BY QUESTION**



*Tip:* Looking out for stable vs. unstable measures. Another thing to look for when running a tracker is how the answers to each question vary from wave to wave. You will find some are a lot more stable than others, and these you might consider removing from some waves, or, again, asking to a smaller proportion of the sample each time.

### Step 5: Using ascription, modelling & chunking techniques

To shorten your survey further, there are some more advanced techniques that you could consider using.

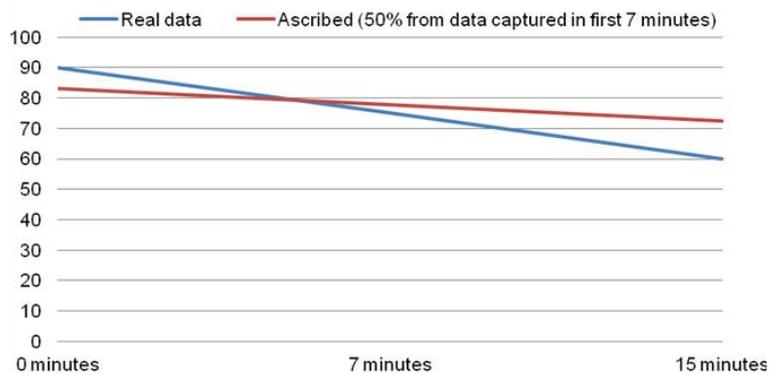
#### Using more full-scale ascription/modelling techniques

The modelling of answers is somehow seen as “cheating” in some quarters, and most clients are extremely nervous about using ascription, often viewing it as some sort of black magic or making up answers. But it is neither black magic nor cheating. I would cite Nate Silver in the case for the defence of this technique, as he used Monte Carlo style modelling techniques to predict all the US state elections results which is the process of modelling survey answers that I would most espouse.

My advice is to not be afraid of this technique: it can be an incredibly effective and valuable way of optimising some types of surveys but to do it properly it's a job for real marketing science experts. I am not going to go into the details of how to do it in this paper, just try and explain why you should consider doing it.

To illustrate the value of ascription techniques let's take an imaginary 15-minute survey, and say that at the beginning, respondents give answers to a Likert scale question with an accuracy of around 90%, and at this point can be cross-ascribed with an accuracy of 85%.<sup>2)</sup> Then the accuracy of their answers begins to fall. Seven minutes into the survey, the accuracy with which respondents are answering has fallen to 80% in ascription accuracy, and by the end it has dropped to as low as 60%. (These figures are not I am afraid typical of the type of pattern-answering we witness in some surveys.) There is a point at which round about 6 minutes or so in that the ascription data become potentially more accurate than the real data.

FIGURE 10. DATA QUALITY



The math implied here is admittedly simplistic, as the accuracy of answers varies dramatically by question format and topic. But while you are not necessarily going to want to cross-ascribe a whole survey, yet in the case of say a bank of 30 attitudinal questions at the end of a long survey (not uncommon), it would definitely be my advice to split this into two random groups, with each respondent only answering 15. I would almost offer a money-back guarantee that if you did this, and cross-ascribed the answers from each group to the other, the resulting data would be just as good, if not better, than from having all respondents answer all 30 questions.

*Survey chunking*

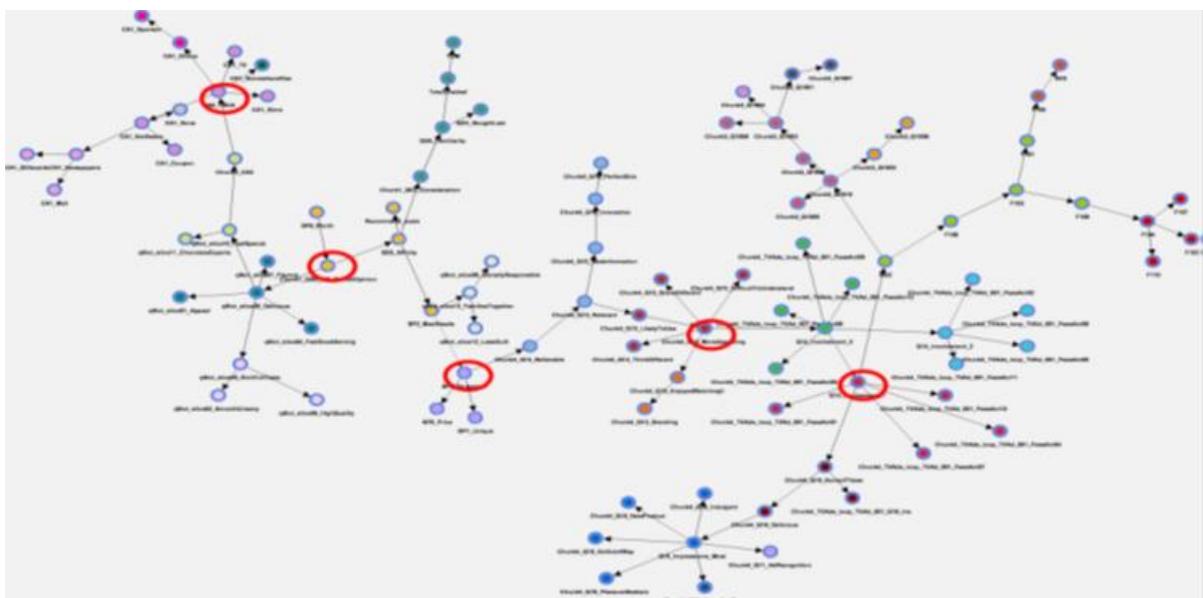
One of the techniques currently being actively explored by researchers trying to shorten surveys is the idea of breaking up surveys into chunks, and have these chunks completed by potentially different groups of respondents. This would allow the respondents to pick up and put down the survey, and choose how much of it they want to complete, without losing out on all their incentive.

The skill to doing this effectively is to understand what parts of the survey need to sit together, and what parts you can effectively separate. For example, you cannot ask spontaneous brand awareness and prompted brand awareness in separate parts of the survey – these questions are only useful when kept together.

To do this properly requires some quite advanced statistical techniques. Unlike most of the ideas in the paper, therefore, this is along with ascription is definitely not a DIY solution, and will need the help of a skilled marketing scientist.

How they go about this is to look at the strength of the underlying relationships between questions, or groups of questions, which can then be mapped out into a “nodal network” like the example illustrated in figure 11 that was produced by an expert from the TNS Marketing Science team. You can then make a decision on the cut points. As with butchery, the cuts are made at the weakest points, the nodes that clearly separate the different parts of the survey “animal”.

FIGURE 11.



If you are interested in reading more about this technique, I recommend “Modular Survey Design: Bite Sized Chunks 2”, an excellent paper published at the 2013 Casro conference by a colleague of mine, Frank Kelly, in association with Sherri Stevens at Millward Brown and Alex Johnson at Kantar.

Their early pioneering experiments exploring this methodology are very encouraging. They found that giving respondents the freedom to do a survey in parts motivated slightly more of them to go on to complete the whole survey, and resulted in greater respondent satisfaction with the experience. They found they were able to exploit more of the partial complete which is usually thrown away, effectively extending the value of the sample by upwards of 20%.

Frank Kelly and his team are currently exploring a new survey chunking approach called interlocking modularization which is a module randomization weighting process for determining combinations of modules in a multi-modulized chunked survey and including the survey dropouts that completed a module, effectively meaning you can use the data from the dropouts which would normally be thrown away. His paper on this topic has recently been published by the Casro Digital Research Conference (March 11/12, 2014).

## PART 2: WRITING & DESIGNING BONSAI SURVEYS

Having pared back your survey statistically, the next stage is the writing design process, where there are a lot more savings and reductions to be gained.

Again, let's break this into a 5-step process:

1. Consolidating and grouping questions: Using nested sub-question techniques and “dashboards”
2. Changing the question
3. Asking smarter more rewarding questions
4. Deciding on your design approach
5. Survey design ergonomics

### Step 1. Consolidating and grouping of questions

The first thing to consider is the possibility of consolidating and grouping questions together. One of the most frustrating things for a respondent is a survey with a badly organised structure, where they are asked to think about one topic or issue, then one or more others, before coming back to the first again. This makes them effectively pick up and put down information, like opening and closing drawers in their mind, which takes time and mental effort. The trick is to design a survey in which each drawer need only be opened once.

As an example, say we have a survey where we want to know what brands respondents are aware of, like, and use. A standard survey would ask about awareness, brand by brand, then liking, brand by brand, then usage, brand by brand.

It is theoretically more efficient to combine these into one simple thinking process. Take this example from this book survey. The original questionnaire involved three separate questions:

*Q1 Which of these books are you aware of?*

*Not heard of*

*Heard of*

*Read*

*Q2 (ask if heard of) Which of these books are you interested in reading?*

*Interested in reading*

*Not interested in reading*

*Q3 (ask if read) How much did you enjoy reading these books?*

*Did not like*

*Liked*

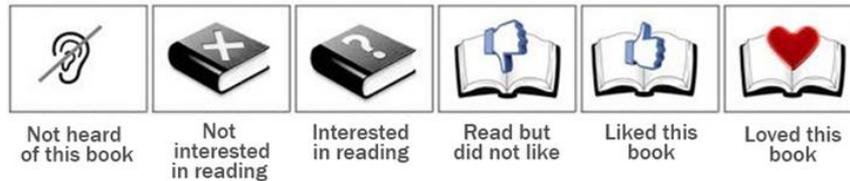
*Loved*

Using a combined scale you could ask these three questions in one go:

Not heard of this book	Heard of, not read but not interested in reading	Heard of not read but interested in reading	Heard of and read but did not like	Heard of read and liked	Heard of read and loved
---------------------------	---	--	--	----------------------------	----------------------------

There is a problem, though, with this approach and it's a big one. Respondents don't like processing long lists of choices like this, because they either have to hold a lot of information in working memory or read the list several times.

The trick to solving this issue is to communicate the choices through effective iconography. We often see icons in surveys as having a purely decorative role, but in this type of question, they can be highly functional.



A good icon, once read, acts as a quick visual prompt that can be processed in a fraction of a second, meaning respondents don't have to hold any of the choices in their working memory. The icons reduce the memory load, making the question much more straightforward to answer. This results in a question that combines the data of the earlier three parts, but with little more respondent effort than one of them. (Note also how the icons enable the language beneath to be tidied up so you don't need to repeat heard of and read!)

#### *Intelligent nesting of option choices*

One of issues I commonly encounter when trying to optimise surveys is long brand lists. Any list with more than 10 options sets off alarm bells in my mind and my immediate instincts is to want to shorten it. Long brand lists literally can paralyse a respondent's brain. They trigger dropout and often deliver back very poor data.

When you do have a long list I would consider using intelligent nesting of choices, which basically means breaking out long tick box lists into primary choices and follow up with sub-choice questions where you can. I would advise this for practical bonsai design reasons but also because it will potentially improve your data.

Here is an example question from a drinks tracker to show what I mean....

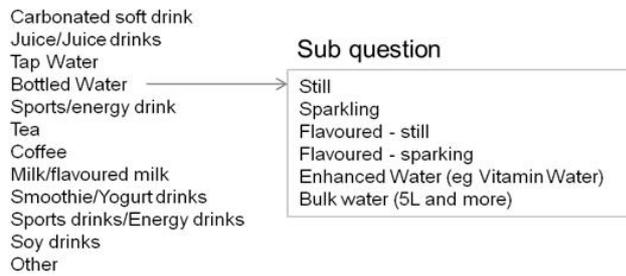
#### *What type of drink have you consumed in the last 24 hours?*

- Carbonated Soft Drinks Sparkling-Regular
- Carbonated Soft Drinks-Light
- Juices/Nectars
- Juice Drinks
- Vegetable Juices/Juice Drinks
- Tap water
- Bottle water - Still Unflavored
- Bottle water - Sparkling Unflavored
- Bottle water - Still Flavored
- Bottle water - Sparkling Flavored
- Bulk water (5L and more)
- Enhanced Water (e.g. Vitamin Water)
- Sports Drinks
- Energy Drinks
- Tea - Black
- Tea -Green
- Coffee
- Plain White Milk
- Flavored Milk
- Yogurt Drinks
- Smoothies
- Soy Based Drinks
- Other Non-Alcoholic RTD Beverages

Presented with this long list it's immediately hard work for a respondent to read though and process. It is also too long a list to present on a mobile screen without scrolling – which is something of a problem which I will discuss later in this paper. These types of long lists get poorly read and as a result can result in major data errors.

The revised question approach using sub question for bottled water:

*What type of drinks have you consumed in the last 24 hours?*



For the 90% of people who didn't consume any bottled water they are spared from having to read though this detailed list and this list now fits into a mobile phone screen.

## Step 2. Switching questions. Some other shorter ways of asking long grid questions

Aside from statistical techniques there are plenty of ways of reducing the amount of work involved in answering questions by simply changing the way you ask the question.

This is probably best illustrated by taking a real survey question and exploring the different ways it could be asked and looking at the time savings and efficiencies in doing this.

Here is an example of a not untypical type of question you might want to shorten somehow.

*How important on a scale of 1 to 10 are these 15 factors when deciding on which mobile phone to buy?*

- The design
- Functionality
- Quality of the speakers
- Quality of the camera
- Screen resolution
- Screen size
- Weight
- Battery life
- Operating system
- Mobile phone supplier
- Memory
- The number of free text
- The number of free minutes
- The amount of data

This question takes around 40 seconds to answer and to differentiate the opinions about all 15 issues requires a sample of around 1,000 respondents as the scores for many of the options are very similar.

*Ask a pre-shortening question:* A good way to make this less work for the respondent and to more efficiently hone in on the key issues is to ask a preliminary filter question to pick the most important factors, e.g. "Of these 15 factors which 5 are most important to you when choosing a mobile phone". Time saving doing this = around 10 seconds and the sample requirements to ask this question is lower too, only around 400 responses are needed to differentiate opinions of the top 5 issues.

*Use selective ranking:* If you want to simplify it further you could ask them not do any rating at all and ask them to "rank the top 5 factors". Time saving = 15 seconds and this reduces the minimum sample requirements to roughly 300 to gain statistical clarity.

*Choice based max diff:* Another way of differentiating opinions further is to present in small random sub sets of say five choices and ask them to pick the most and least important factor. This technique delivers back wonderful data free from any of the cross cultural issues you face interpreting rating scales data from different markets. Time saving = 15 seconds and minimum sample requirements is again around 300.

All three of these alternative techniques on one hand sacrifice some information, the detailed rating scores for each factor, but I would argue that much of this information is of little real use because the differences in rating scores can be so minimal they require huge samples to pick it apart. By being slightly more pragmatic about the information you gather you save time and require much smaller samples.

*Note:* You may be thinking, well we are only saving a few seconds here by doing this so what is the point? Across a whole survey saving 20 seconds here and 30 seconds there you can may be able to save three or four minutes by the more intelligent selection of questions and when you consider that the rate of dropout from respondents completing surveys on a mobile phone is anywhere up to 2%-3% a minute, a four minute saving could increase completion by 10% which is “not to be sniffed at” as we say in the UK.

### Step 3: Asking smarter more rewarding questions

Designing Bonsai surveys is not just about making it shorter. I want you to think about ways of asking smarter more rewarding questions too.

There are all manners of fresh and exciting new ways of asking questions that make surveys more interesting and rewarding experiences for respondents, born out of the learning of behavioural science, gamification & prediction markets.

As I have previously written several papers on this specific topic I will not dwell on this point in this paper too much but highlight a few of what I feel are the most important areas of new thinking.

#### *A shift from rating to binary questioning techniques*

The work by Daniel Kahneman, the Nobel prize winning psychologist exploring how we make decisions and his ideas around system 1 instinctive fast thinking and system 2 analytical slow thinking, places a big question mark over the value of asking rating style Likert questions generally in surveys.

Rating is a slow and complex “System 2” thinking process and easily gets corrupted by the idiosyncrasies of how each respondent chooses to rate things. On the other hand, binary questions tend to be answered more instinctively, and can be answered a lot faster and provide more reliable answers.

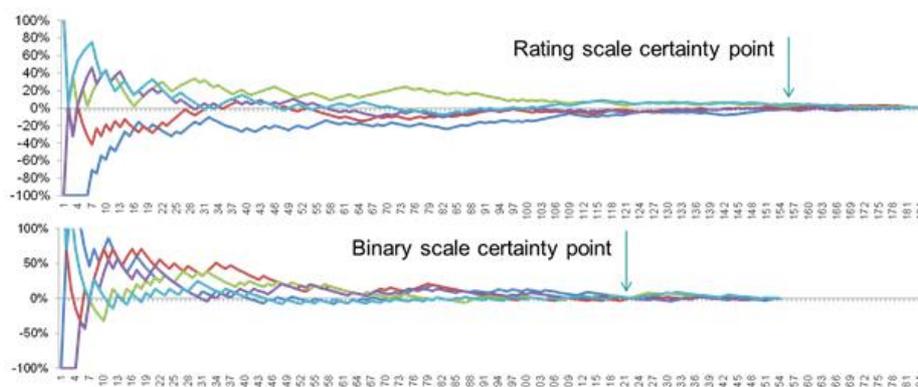
Say for example you go to the cinema to watch a film and I ask you did you like it or not, you can almost instantly tell me yes or no. On the other hand if I asked you to rate the film on a scale of 1-10, this takes quite a few seconds processing time and my answer is difficult to interpret as my score of 7 might be the same as your score of 8.

Rating = 2-5 seconds thinking

Binary = + <0.5 second thinking

There is growing evidence to suggest that binary questions deliver more reliable, better differentiated answers than rating scales. Recently we have been doing a lot of exploratory work that seems to confirm this (see figure 12, an example taken from one of our recent experiments), but why I like binary questions is that they make surveys shorter!

FIGURE 12.



*A shift from rating to choice based techniques*

I often see the misuse of rating protocols in surveys, where they are being used ultimately to make a decision. A classic example of this is in monadic ad testing where two groups of respondents rate an ad and then you compare the rating scores. It can be far more efficient to simply ask people to make a choice as to which they like best.

Take this example in figure 13: these two ads gain nearly identical rating scores but when asked to pick the best one there is a 70% preference for ad 1.

**FIGURE 13.**

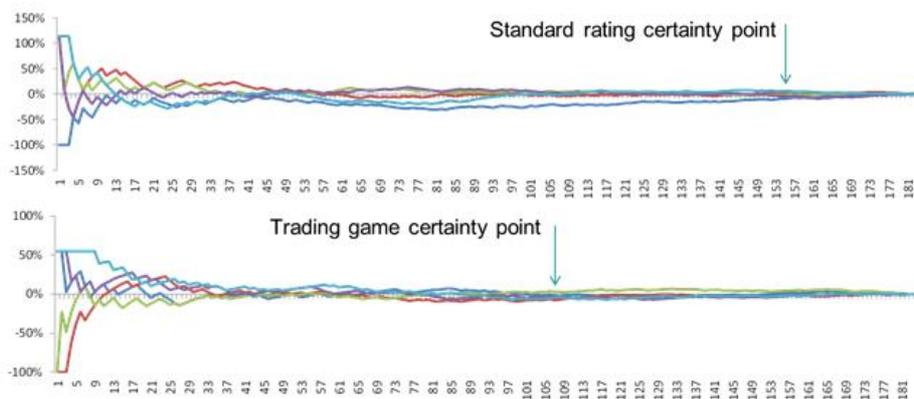


Rating these two ads would require sample of 200+ to pull them apart but a choice base approach reached a statistically significant answer with a samples of less than 20.

*Asking a more rewarding question*

The other factor that impact significantly on sample size requirements for a question is simply how engaged your respondents are in the task. The example in figure 14 illustrates this point very well. Two groups of people were asked to evaluate a series of ads. One group were simply asked to rate them, the other were playing a game where they had to trade ads based on whether they liked them or not. As a result they concentrated more, put more thought into their answers and as a result we found we needed much less sample to reach the statistical certainty point.

**FIGURE 14.**



If 20% of respondents are effectively making up answers in tasks like this because they don't care what they say,<sup>3</sup> we have found you need up to 40% more sample to smooth out the systematic errors this introduces.

In the SingTel case study at the end of this paper I will show how some of this thinking can be applied in a real survey.

*A shift from thinking of a respondent as bit of data to a quantum thinking machine!*

Probably the most exciting area of new thinking over survey question design is a shift from asking personal behavioural reporting and attitudinal questions towards more predictive “we thinking” techniques, a shift from treating respondents with bits of data to becoming quantum thinking machines.

The best example of this is the movement in political research away from asking thousands of people which way they will vote as a means of forecasting an election result to asking smaller groups of people to predict who will win using prediction markets.

In the case of the famous Iowa Election Market where small groups of people often as low as 20 have gambled on the outcome of elections, they have famously managed to out-predict the best performing polls on nearly 460 out of 580 occasions.

What they are doing here is using people's brains to solve a problem rather than treating people as bits of data, which is highly reliant on having representative samples that often can have systematic errors.

To reinforce this point is another example: a group of students at a meeting were asked in a piece of research if they thought they would clear up their coffee cups afterwards, and 50% said they would. Another group of students were asked a different question: "predict how many students will clear up after the meeting". Their average prediction was 15%, much closer to the 13% of students who actually cleared up.

The first way of asking this question could be seen as a typical market research approach and because of a cognitive bias in all students' viewpoints of their own behaviour it was a hopelessly inaccurate measurement technique. Changing the question to ask the student to solve a problem, they used their brains to answer it and so collectively they came up with a far more accurate answer.

This type of thinking is permeating into more mainstream research. For example instead of asking lots of people lots of question about what people think about your brand of shampoo, why not ask a smaller group of people to predict the main reasons why people buy this brand of shampoo and offer a prize (real or intangible) to the people that are able to identify the most common factors.

I believe this type of thinking has the potential to change the way whole vast tracts of market research is conducted, leading to shorter more effective pieces of research. Watch for a paper we are currently preparing on this whole topic!

#### **Step 4: The design process**

Now we move on to pure design considerations.

The first decision you have to make when designing surveys that you want to work across different platforms and devices is what is going to be your minimum design standard. The key thing really to think about is what to do about feature phones?

In many markets feature phones are still the dominant mobile format. The problem is that 95%+ of surveys that our panellists are asked to complete are simply too long and complicated to run on most feature phones. Of those that attempt to complete them, very few succeed in getting to the end, and abandonment rates can be upwards of 80%.

My advice is that if you're surveying in a market where you need to reach audiences via a feature phone survey, you should treat this as a separate design project and you need to have a completely cut down approach. You could probably see about 15 questions as the effective limit of the length of a survey on a feature phone. It certainly not impossible to do very effective surveys using just 15 questions, but it takes a real Bonsai survey design master to create them.

If you want to learn from a real Bonsai survey design master, I would advise you to study some of the work that Jannie Hofmeyr and his team at TNS in South Africa has been doing, forensically paring back surveys to the most necessary and important questions, condensing whole brand trackers into under eight questions.

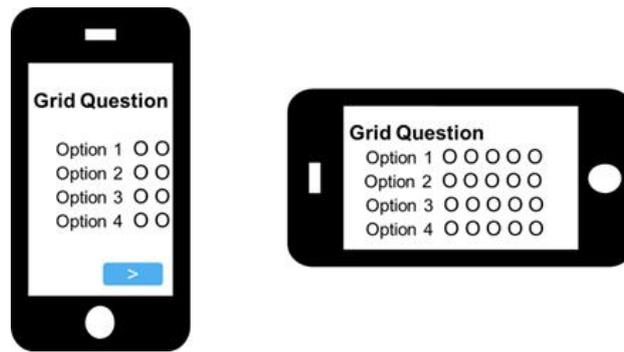
My advice for the rest of us is to forget about feature phones for the types of standard surveys most of us are producing and focus on making it compatible for Smart phones upwards as in most markets these are the primary mobile platforms used to complete surveys these days.

#### *The portrait vs. landscape issue*

The next thing to contend with when designing a platform neutral survey is the fact that the surveys we design on PCs are naturally landscape in format and the mobile phone is inherently a portrait format.

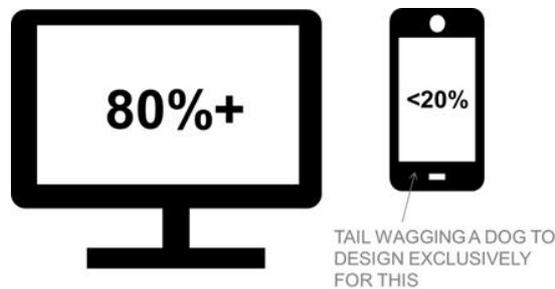
Respondents prefer to use mobile phones in portrait mode but the problem, however, is that most surveys we produce right now are very width-heavy, and many questions simply don't fit efficiently onto a mobile screen in portrait mode – not at least the grid question that 90% of surveys currently include.

FIGURE 15.



Also, pragmatically, as only a minority of surveys are currently being completed on mobile and tablet devices, and so adapting everything to work portrait-style could be seen as a bit like a tail wagging a dog.

FIGURE 16.



*Two solutions: Adaptive design or fixed landscape*

So the options we have to deal with this are either an adaptive design approach, adjusting the design specifically for mobile devices and using a different design on PCs and tablets; or using a “one size all fits all” approach, fixed in landscape mode.

The chosen option very much depends on the survey technology used, and the type and complexity of the surveys you are designing. Each approach has pros and cons.

*Adaptive design: Pros & cons*

The advantages of adaptive design are that respondents completing on a mobile phone don’t have to turn their phone around, surveys are easier to complete using one hand and by customising the survey for the mobile the integrity of the PC version of the survey is not compromised in any way.

Some survey software providers have developed sophisticated, fully customised design solutions to implement adaptive design which takes most of the work out of having to do this.

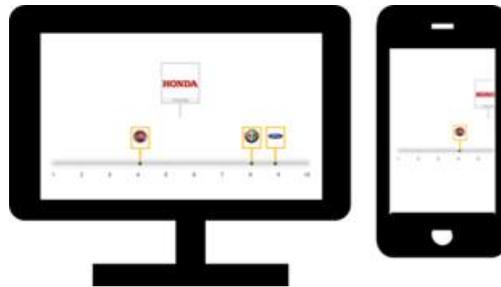
There are, however, some drawbacks. Firstly, some questions simply won’t naturally adapt to portrait mode, since a lot of content, such as advertising, is wider than it is deep. Adaptive design approaches can also fall down when it comes to formatting which means some questions looking very messy on a mobile device and not properly fitting the screen. This basically means having to effectively design and QA two surveys in many cases.

FIGURE 17.



The other consideration to bear in mind is that many of the most successful comparative techniques used in market research rely on side-by-side comparison, as we more naturally scan sideways than up and down and as many mobile app game developers have discovered, landscape is an all-round more efficient design format.

**FIGURE 18.**



*One size fits all landscape design: pros & cons*

So the second choice is to fix design in landscape and ask people to rotate their phone.

**FIGURE 19.**



The advantages of the fixed design in landscape approach is that it makes it a lot easier to design: the PC and mobile version can essentially be the same. It also means you have access to a greater range of more creative question formats.

The cons, apart from the inconvenience of rotating the phone, are that you have to adapt the overall design of your survey to meet the minimum design standards of a mobile phone. So your whole approach to the design of your survey has to change.

*So which do you go for?*

For us right now, as a service provider with little control over survey content, we find it difficult to use adaptive design approaches. In the typical survey we are asked to adapt, there are always one or two questions that just don't work in portrait mode, and so, for practical reasons, we tend to opt for fixed landscape design.

However, if you are in greater control of your survey content, and able to write and design surveys to ensure all questions are naturally adaptive and using the right survey software, I would recommend an adaptive route as this can provide a better experience for respondents.

### **DESIGNING “ONE SIZE FITS ALL”, PLATFORM-NEUTRAL SURVEYS**

If you are opting for using a one size approach here are some tips:

1. Be conscious of the “sacred viewing zone”
2. Make every pixel count!
3. Avoid scrolling at all costs
4. Make the most of interactive question formats
5. Scale down and think about perspective ratios not pixels

#### **The sacred viewing zone**

The most important thing to consider when designing platform-neutral one size fits all surveys is the “sacred viewing zone” at the top of the page. Try to fit as much content into this space as you can, as this is what will be visible on a mobile device.

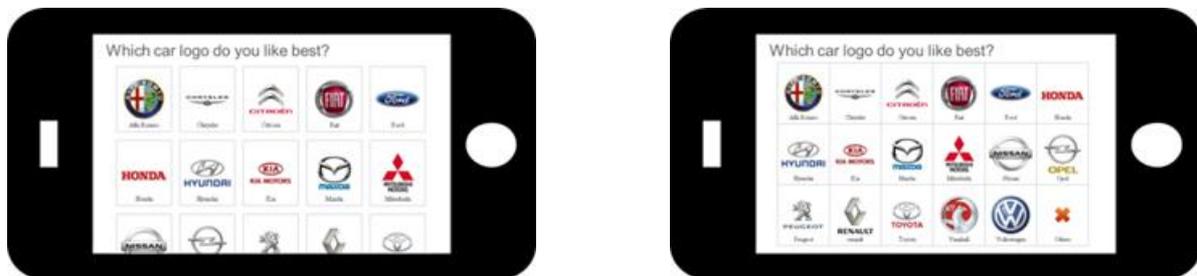
FIGURE 20.



### Make every pixel count

The next piece of advice is to remove gaps between option choices (see figure 21) to ensure as much content as possible appears in the sacred viewing zone.

FIGURE 21.



If your survey technology allows, we also recommend switching from what might be described as “old school” buttons, since the buttons themselves take up valuable space. Use point-and-click box choices instead.

FIGURE 22.



### Use auto-nexting where possible

It's best to avoid respondents having to scroll down to find the “next” button, because with some survey software, scrolling back results in the navigation bar further reduces the available sacred viewing space.

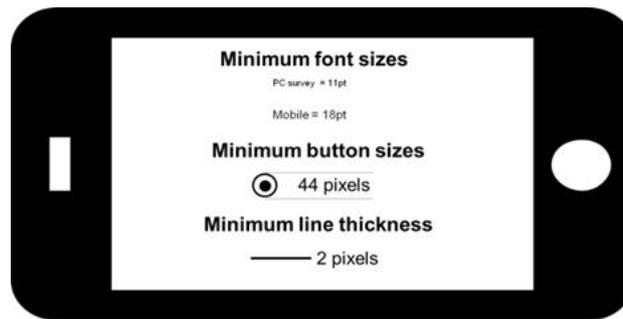
### Increased font and minimum button sizes

We recommend increasing the minimum font size used in your survey to 18pt, and if you have space, even bigger for question titles.

All your radio and tick-choice buttons also need to increase in size to a minimum height of 44 pixels (this is the Apple standard).

Lines also need to become a bit thicker if you use them. 2 pixels is the minimum, as some newer smart phones use extremely high pixel density, sometimes making single-pixel things disappear.

FIGURE 23.

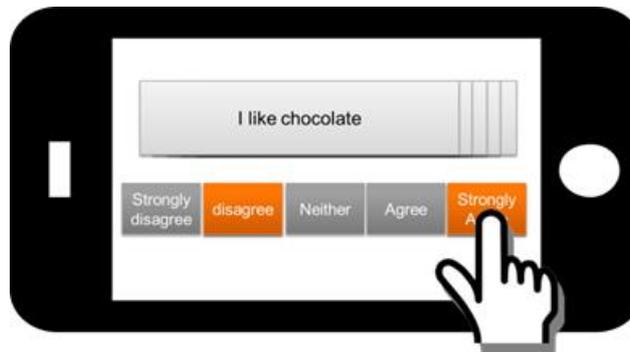


### Exploiting interactive question formats

Whilst interactive question techniques might have been considered a nice thing to have in traditional surveys to improve engagement they become a whole lot more valuable weapons when designing a platform neutral survey.

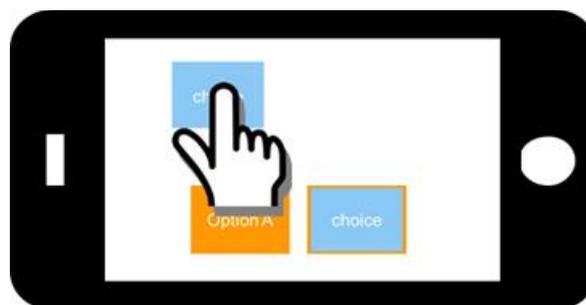
Dynamic grids, where as you select an option the next scrolls in dynamically, for example are a brilliant question format for mobile devices, as traditional grids often don't fit on a mobile screen.

FIGURE 24.



Dragging and dropping is a far more intuitive process on a touch screen so these types of question become far more valuable questions in your repertoire.

FIGURE 25.

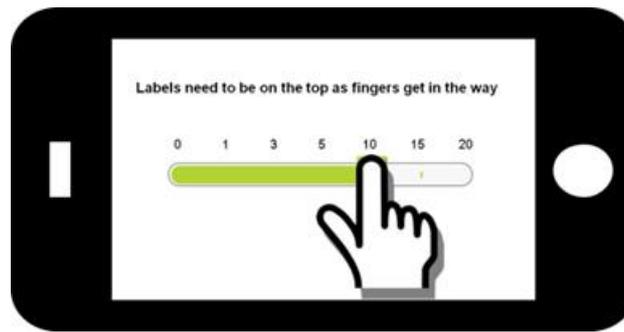


As more and more of these question formats become available in HTML I would push for all your surveys to include them.

### Rethinking question layout

There are lots of smaller things to consider. For example, make sure you push open-ended text boxes as high up the screen as possible, to give space for the keyboard when it pops up. Labels need to go on top of sliders rather than below, so they're not obscured by the respondent's finger.

FIGURE 26.



Longer instructions and question text is sometimes best placed on a separate screen as it can take up valuable sacred viewing space.

Furthermore try to make every question as short as you can: think like a tweeter. There are often huge space savings to be had by cutting back on the question text.

Don't be shy of using icons and visuals where you have space. They can really bring a survey to life, and, as explained earlier, have an important role in communicating choices more effectively.

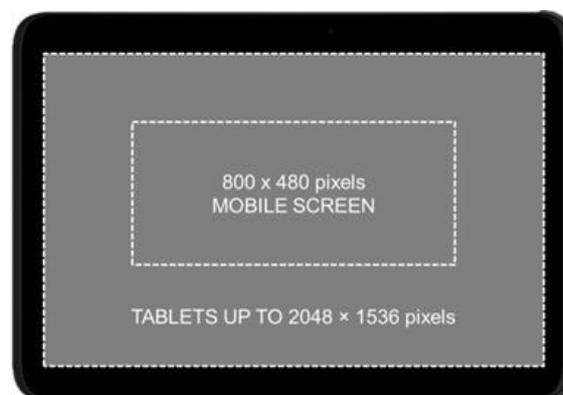
FIGURE 27.



### Scale down not up

Remember that your surveys will also be completed on tablet devices, which can have supersized screens. When designing your surveys, it's better to design them so they scale down rather than up.

FIGURE 28.



### Get hold of a survey design testing kit

The one thing you will also need is a survey design testing kit like the one shown in figure 29. This is a bit more expensive I am afraid than the respondent goggles I recommend you also purchase, but it's crucial, when designing platform-neutral surveys, to be able to see what they look like on a variety of devices.

FIGURE 29.



### CASE STUDY: CREATING A BONSAI U&A STUDY FOR SINGTEL

As part of this paper Singtel, the Singapore telcom company, kindly agreed for us to explore how we could apply some of this Bonsai survey design thinking to one of their surveys.

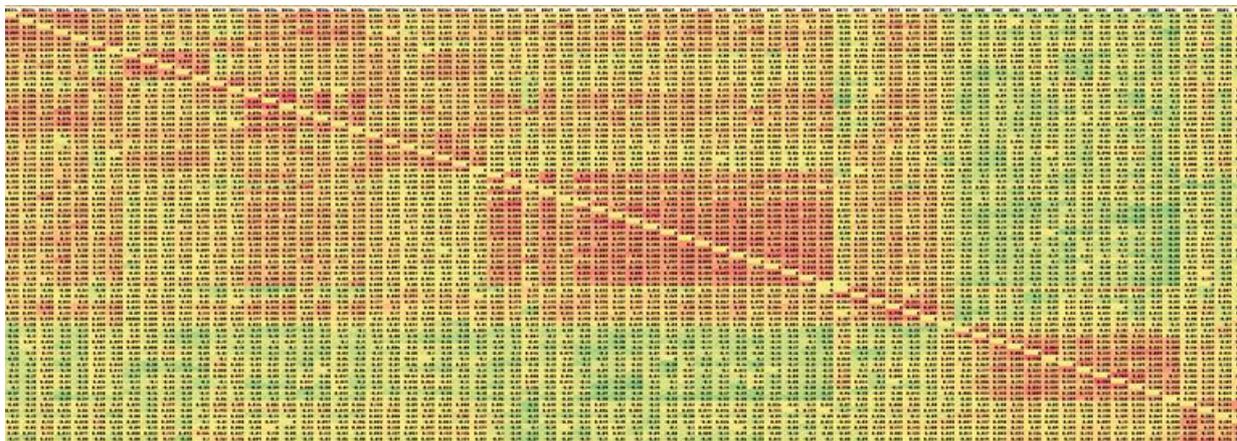
The survey they asked us to look at was a fairly typical 25 minute U&A study exploring the use and attitudes towards mobile phones. The survey included a wide range of questions about competitive mobile phone services and payment tariffs. A large section of the survey was dedicated to mapping out what features of the phone were being used including what types of apps and about how the phone was being used competitively compared to PC and tablet devices. It also included a range of attitudinal questions.

The first step was to sketch out the core of the survey and run a quick pilot on 100 respondents. We didn't bother asking any of the demographic questions and some of the core questions that we knew were not going to change.

From this pilot we were able to immediately identify a range of ways in which the survey could be shortened.

Conducting Pearson correlation test results, this enabled us to identify 24 cross correlating questions (with correlation score about 0.65) which we were able to consolidate, saving one minute. Figure 30 is a visual to show you what that test looks like visually - patches of red are the closely correlating questions.

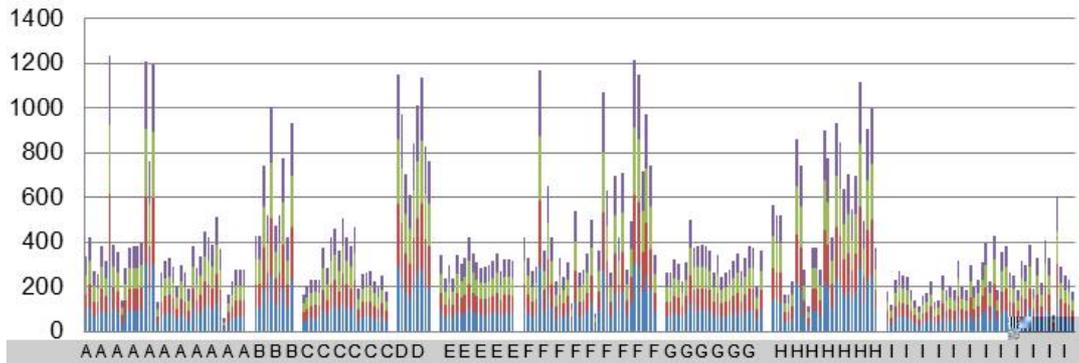
FIGURE 30.



We conducted some question level sample size analysis and determined that the maximum size of the sample needed to conduct this survey was around 1,000 respondents if they were planning to cut the data up to four ways for some sub demographic analysis. However as only 10% of the questions needed samples this big pragmatically we felt we could reasonably reduce this to 800. As you can see most of the rest of the questions only needed to be answered by less than

half this total sample or round about 400. So by rotating sections A, C, E, G & I and using question level quotas we estimated this survey could potentially be reduced in length by around 8 minutes.

**FIGURE 31. MINIMUM SAMPLE REQUIREMENTS**

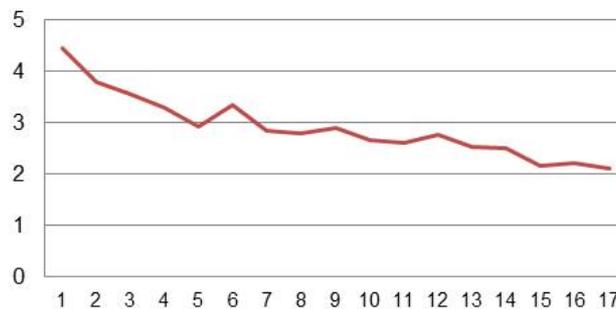


Overall this estimated 25 minute survey could be asked to 25% less people than originally planned and trimmed back efficiently to around 15 minutes without any significant loss of data.

**Survey design considerations**

As in many U&A survey one of the main problems from a respondent point of view was the very repetitive nature of some of the questions being asked. This was apparent when studying answer time data where we observed rapid decays of attention on many of the main sets of question. Figure 32 illustrates a set of questions asking what activities were undertaken on a mobile phone, PC and tablet.

**FIGURE 32. DECAY IN ANSWER TIMES**



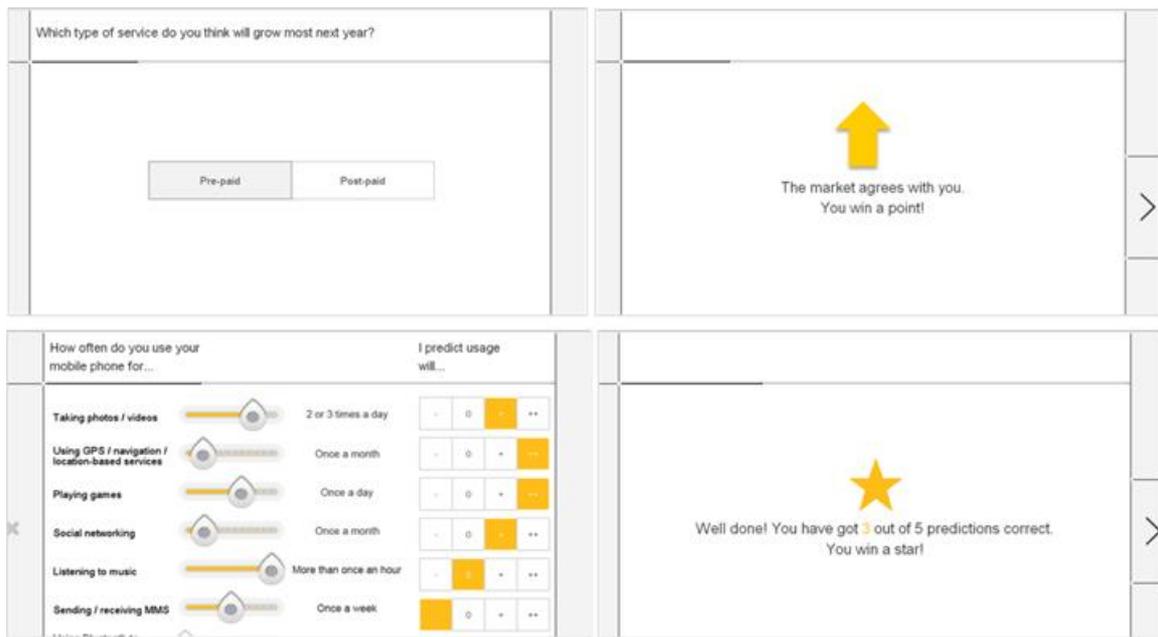
So what could we do to improve this? Our approach when encountering these types of issues it to think about ways of making the survey a more satisfying and rewarding experience for the respondents.

**Applying a “Questification” mechanic**

In previous papers I have talked about the concept of “Questification”, the idea that respondents are prepared to put more effort and thought into a survey if they feel there is a real clear purpose to what they are doing.

Our approach to questify this survey was to ask them to take part in a special survey to help predict how mobile phone usage will change in the future. We felt this was a more exciting task for respondents to get involved with and we adapted the questions around that theme.

FIGURE 33.



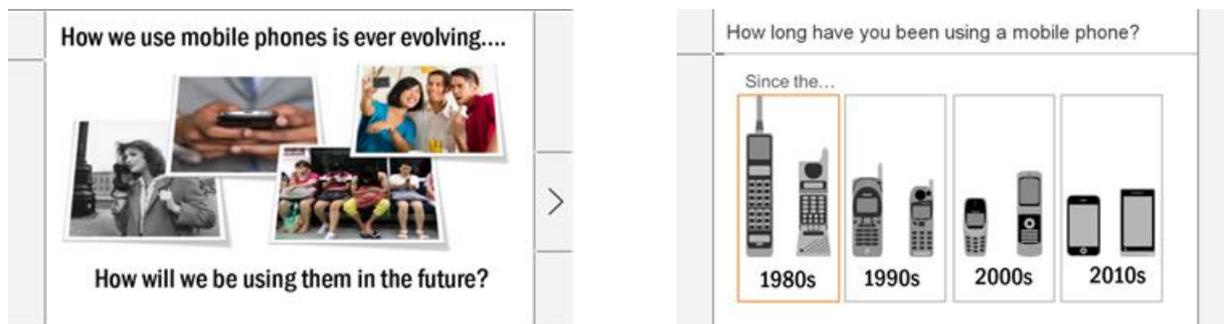
Throughout the survey we asked respondents to make predictions about how they thought mobile phone usage would change in the future and compared their answers to what other respondents thought. We pitched this as a game: if they were able to predict what the market of other respondents thought consistently they won star points at the end of each section.

**The pure design aspects**

As for the design aspects we opted for a fixed horizontal design format and made maximum use of all the creative question formats we had in our armoury. We edited back a lot of the question text to ensure the question fit efficiently into the “sacred viewing zone”. All bar three questions in the whole survey we were able to design so they could be answered without the need for any form of scrolling if taken on a mobile device.

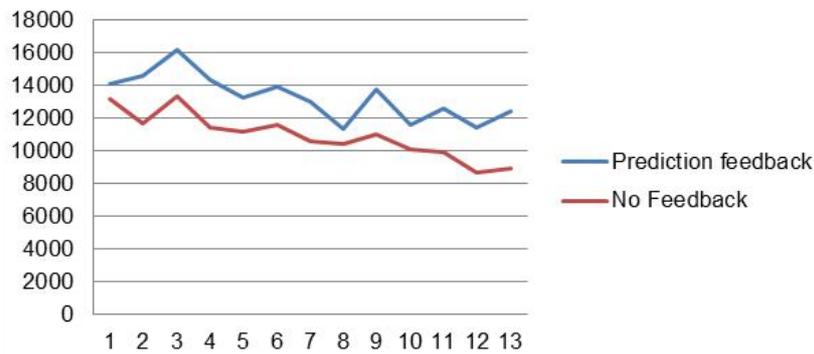
We noted in the initial pilot a significant proportion of respondents dropping out near the start and so we made particular effort to redesign the intro to the survey to make it more welcoming. In the example illustrated in figure 34, the first question was really designed as a catch to grab their attention and gain their interest.

FIGURE 34.



To test the effectiveness of this concept we conducted a second iterative pilot using split samples, one group being asked the question without any prediction feedback and the other group we gave feedback on their predictions and awarded them stars if they were performing well.

This feedback process really did succeed at emotionally drawing them into the topic, answer times for like for like questions increased by an average of 18% (which may not sound much but considering two-thirds of the time spent answering a question is taken up just reading it this 18% increase amounted to a near doubling of the average consideration time for each question).

**FIGURE 35. COMPARATIVE ANSWER TIME**

What we were also able to observe were statically significant improvements in answer quality, an estimated 50% reduction in pattern answering measured by the differences in standard deviation of answers at a respondent level.

Now the overall survey completion time went up by three minutes, which may in some camps be viewed as an issue, but not in this case. More respondents actually completed the survey (about 5%) and their survey satisfaction scores rose by 20%. Linguistic analysis of the open ended feedback revealed extremely enthusiastic reactions towards the reward mechanics in particular, and they were investing that extra time because they wanted to and were motivated to.

## CONCLUSIONS

I hope this paper has helped to teach you some of the basic technique to enable you to become a Bonsai survey designer.

It is clear from exploring this topic in detail that in nearly every survey there are significant economies to be made in survey length and sampling requirement. We have found that it's quite possible to reduce the average 20 minute survey down to 12 minutes with a thorough and more systematic statistics driven approach.

The key economies to be made come from hunting out and removing redundant questions that nearly every survey we tested had, and shifting to setting quotas at a question level rather than a survey level. We estimated this technique alone could reduce a typical survey length by 30%.

I hope I have been able to demonstrate that by asking more thoughtful and well organised questions you are not only going to improve the experience for respondents by making the survey shorter, but also improve the quality of your data.

I would also challenge you to be more pragmatic, and think about what you really want to know. Chop back your brand lists and option choices. Never be satisfied that you are asking a grid question, there so many better ways this information can be gathered. Think about what you are going to do with the data more and use this to determine how you should ask the question.

To achieve all this I would advocate a shift towards a more iterative survey design philosophy and more closely integrate marketing science skills into the everyday survey design process.

When designing Bonsai surveys your main focus should be on making every pixel count. Try and condense each question so it fits into the "sacred viewing zone" that means respondents, whether they are completing the survey on a mobile, tablet or PC, will all experience the same quality of question.

A summary of advice:

1. Think like a respondent
2. Adopt an iterative design approach
3. Hunt out the redundant questions
4. Apply question level caps & quotas
5. Don't be afraid to model, ascribe & chunk
6. Combine and nest questions making best use of iconography
7. Switch to shorter more effective question techniques
8. Treat respondents as quantum thinking machines & ask smarter more rewarding question
9. Exploit open ended questions
10. When designing your survey think of the sacred viewing zone & make every pixel count

I recognise this process is not always easy and one of the biggest hurdles you face will be in convincing some of your clients and possibly your marketing department of the need to implement changes like these. Good luck with that!

## ENDNOTES

1. Source: Lightspeed Research 2014 panel data
2. Questions can often be ascribed with accuracy of 85% or higher
3. 20% is a not uncommon number, in fact you might consider it to be the average.

## REFERENCES

- Kelly F, Johnson A, Stevens S (2013) Modular Survey Design: Bite Sized Chunks CASRO Online Research Conference
- Kelly F, Johnson A, Stevens S (2012) Modular Survey Design for Mobile Devices CASRO Online Research Conference
- Smith R, Kotzev I, Miller C, D Kachhi (2012) Modularizing Surveys To Meet Today's Respondent Challenges
- Cape, P, (2010) Questionnaire Length, Fatigue Effects and Response Quality Revisited ARF Re:think
- Puleston, J Eggers, M (2012) Dimensions of online survey data quality ESOMAR congress
- Puleston, J & Rintoul D (2012) Can Gaming Techniques Cross Cultures ESOMAR Asia
- Malinoff, B & Puleston, J. (2011) How far is too far? Traditional, flash and gamification interfaces, and implications for the future of market research online survey design ESOMAR 3D Miami.
- Puleston, J. & Sleep, D. (2011) The game experiments: Researching how game techniques can be used to improve the quality of feedback from online research. ESOMAR Congress Amsterdam.
- Downes-Le Guin, T., Baker, R., Mechling, J. & Ruyle, E. (2011). All fun and games? Myths and realities of respondent engagement In online surveys. ESOMAR Congress, Amsterdam.
- Sleep, D. & Puleston, J. (2009). Panel quality: Leveraging interactive techniques to engage online respondents. ARF Congress, <http://www.gmi-mr.com/solutions/resources.php>.
- Strube, S. & Zdanowicz, Y. (2008) Maximizing respondent engagement through survey design. CASRO Panel Conference, Miami.
- Puleston, J and Sleep, D (2008) Measuring the value of respondent engagement: Innovative techniques to improve panel quality. ESOMAR Online Dublin
- Reid, J., Morden, M. & Reid, A. (2007) Maximizing respondent engagement: The use of rich media. ESOMAR Congress.

### *Other*

- Krosnick, J. (1991) "Response strategies for coping with the cognitive demands of attitude measures in surveys." Applied Cognitive Psychology. Vol 5, 213-36.
- Sudman, S and Bradburn, N M (1982) Asking questions: A practical guide to questionnaire design. Jossey-Bass: San Francisco.
- Miller, G A (1956) 'The magical number seven, plus or minus two: Some limits on our capacity for processing information'. Psychological Review 63: 81-97.
- Alreck, P L and Settle, R B (2004) The survey research handbook, 3rd ed. McGraw Hill Irwin.

## THE AUTHOR

Jon Puleston is VP Innovation, GMI, United Kingdom.