

# Watch What I Do! - Using graphical input controls in web surveys

Nicola Stanley & Stephen Jenkins

## Abstract

This paper considers the pragmatic aspects surrounding the deployment and use of graphical image-based controls for collecting web survey response data. It compares the relative merits of graphical control surveys in terms of the respondent experience of survey completion including usability, engagement and enjoyment of taking part in the survey. The design effort required of the survey development team is also considered. In particular the paper explores whether interactive web graphical scales can be a direct replacement for traditional web surveys or have more specialised application such as brand perceptual mapping techniques and use with specific respondent groups including established internet panels and young people.

## Keywords

Graphical scales, usability, engagement, web surveys, click and drag, brand / perceptual mapping, gap analysis, numerical scales, descriptive scales, visual control, Internet panels.

## 1. Background

Following an initial rush of enthusiasm to exploit the new technology and explore new ways of reaching respondents, most web questionnaires published are simply electronic equivalents of their paper counterparts. Whilst they often incorporate sophisticated control of the interview by using routing, text-substitution, answer masking and other techniques, respondents' response controls have largely remained unchanged. Multi-choice questions are represented either as lists of check boxes or radio buttons (or sometimes a mixture of the two). Open questions typically are characterised by a text box into which the respondent enters their reply.

At the other end of the survey process, survey results presented to lay audiences often make use of charts and pictures and other graphical elements. For example, ratings may be shown on a graphical scale, proportions of two parts may be shown as a single-element stacked bar chart and aggregate respondent perceptions may be shown on a multi-dimensional map.

If the presentation of results is aided by the construction of a suitable graphical image, then why should not the interview process be similarly improved by using suitable graphical images as a means of inputting response data? The purpose of this paper is to challenge the current status of web survey design and to explore the added value that graphical control scales can have on the survey process.

## 2. About graphical control scales

The main motivation for researchers to use graphical control scales in surveys is to increase respondent involvement in repetitive and low interest surveys through the incorporation of interactive game-type elements (Luetters et al, 2007). Some researchers (for example, Couper et al, 2006 and Luetters et al, 2007) have observed that the use of interactive visual analogue and graphical control scales in web surveys can result in higher rates of non-completion, higher rates of missing data and longer completion times compared to standard radio button scales.

Our approach was to develop experimental surveys to explore the use of a variety of graphical input controls on web surveys conducted with a UK based internet consumer panel. We specifically wanted to try a number of different scale types in the same survey to see if we could detect a consistent non-response bias. We were also interested in the possibility that respondents may find our questionnaire more engaging simply because we had used graphical controls throughout.

Controls we explored included:

Abstract sliders with and without reflected numerical values – stepped sliders corresponding to single-response choice questions; continuous sliders corresponding to open quantity questions.

Slider controls such as a grab-able thermometer meniscus – corresponding to open quantity questions.

Multi-dimensional mapping graphics onto which the respondent drags icons representing items (brand logos), dropping them at appropriate locations – corresponding to batteries of open or closed-choice grids.

All were implemented using standard JavaScript technology and thus did not require the downloading of any special applets or controls onto respondents' computers.

Two important considerations were taken into account in the survey development (as seen in Figure 1). Firstly the need for 'no reply' options on the graphical scales. This is a critical design element that is often overlooked in graphical control scales. Secondly the need for some guidance directions for respondents on the use of the graphical control scales. The team had no prior understanding of how intuitive the use of graphical scales would be to an established internet consumer panel.

Figure 1: Graphical scale showing 'no reply' start position and guidance on scale use

**Your beliefs**

*Move the pointer on the scale upwards for agreement and smiley faces appear or downwards for disagreement and sad faces.*

**Please state how much you agree or disagree with each of the following statements:**

I care about food miles when I buy products



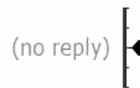
I think that global warming is NOT a real issue



I believe that supermarkets should stop giving customers free plastic bags



I feel that food manufacturers put too much packaging on their products



### 3. Methodology of the trial

Two web surveys were designed in order to compare the relative merits of graphical image-based web survey ratings against those of traditional standard web survey scales. Environmental responsibility was chosen as the topic of the survey because of its current high profile within the media and the possibility to use a variety of scales to explore many aspects of consumers' claimed habits and attitudes and their perceptions of brands.

A script for a standard ten minute survey was produced. Following that a half day project team meeting was held in order to identify ways in which the survey questions could be designed as interactive graphical control scales.

Dimensions that were explored as part of the study were:

Engagement - how motivated respondents are to interact with the graphical components.

Usability - how easy and intuitive graphical control interfaces are for respondents to use.

Accuracy - comparison of results with those produced using traditional web survey scales.

A number of demographic questions covering dimensions which the team felt might make a difference to how respondents interacted with the graphical control survey were included in both surveys; these were age, gender, hours spent on personal internet use per week and level of qualifications achieved.

A web developer and an image designer worked up the initial drafts of the questionnaires which were then subject to a number of revisions as the team's understanding of graphical scales developed.

The links to each survey were sent simultaneously by e-mail to equivalent samples of an internet consumer panel. The samples were composed of 300 respondents per survey and matched for age and gender. The rate of returned responses was the same for the two surveys over the two days/nights that they were in field. The fieldwork was stopped when 112

responses to each survey were received. All surveys were checked and those that were very incomplete or had suspect answers were omitted, resulting in a total of 105 responses from each survey going forward for analysis.

#### 4. Discussion of results

The results are discussed under the main headings of engagement, usability and accuracy.

##### 4.1. Engagement

Although the sample of the respondents was matched for age and gender, the actual profiles of the completing respondents was skewed for age (Table 1). The graphical survey appears to be very appealing to the 25 to 34 age group and of lesser interest to 17 to 24 year olds. Over three quarters (20/24) of the 25 to 34s who received the graphical survey responded compared to only just over a third (9/24) who were sent the standard survey. There was a lower response to both surveys for the younger age groups than for the older age groups. This lower response rate is more marked in the case of the graphical control survey where only a third (7/21) of the 17 to 24 year olds responded to the graphical survey compared to close to half (10/21) of those who received the standard survey. This goes against one hypothesis of ours that graphical surveys are more appealing to younger age groups. There were also no clear indications in the data that the decision to take part in the survey or not was affected by respondents' gender, level of internet usage or educational attainment.

Table 1: Respondent profiles

<b>Age Bands</b>	<b>Graphical Control</b>	<b>Standard</b>	<b>Panel Sample</b>
17 to 19	1%	4%	4%
20 to 24	6%	6%	10%
25 to 34	22%	8%	16%
35 to 44	17%	18%	20%
45 to 54	20%	27%	24%
55 to 64	7%	10%	6%
65 or more	28%	28%	20%
<b>Gender</b>			
Male	46%	52%	48%
Female	53%	47%	52%
no reply	1%	1%	

In terms of time taken for respondents to complete the questionnaire, the graphical survey showed an average response time of just less than 11 minutes and the standard version as close to 10 minutes (Table 2). Each survey also had one respondent taking just over 70 minutes (72 and 71 minutes) to complete - these responses were removed from the final analyses.

The use of graphical controls in surveys has often been criticised for resulting in longer response times and higher drop out rates (Couper et al, 2006). This view was not however supported by these findings. The minimum time for completion of the graphical survey was 4 minutes compared to 2 minutes for the standard survey and we could hypothesize that respondents are spending extra time reading the brief guidance notes on how to use the scales. They could also be spending this small amount of additional time considering their responses because they are more engaged with the subject of the survey when questions are presented to them in a colourful and interactive way.

Table 2: Statistics of response times (minutes)

Survey	Mean	Mode	Median	Minimum	Maximum
Graphical	10.8	6.0	9.8	4.3	34.2
Standard	9.7	4.3	7.5	1.7	27.4

Some basic indices of engagement were asked directly of respondents and overall the graphical control survey was scored significantly higher than the standard survey (95% confidence level) for:

The subject being of interest (87% scored 'yes' for the graphical survey versus 76% for the standard survey)

The question style being enjoyable (86% scored 'yes' for the graphical survey versus 72% for the standard survey).

Those aged 65 or older were less likely to find the question style of the graphical survey to be enjoyable (i.e. 77% compared to 89% for those aged below 65). Of the 5 respondents who answered 'no' when asked if they found the question style enjoyable, 4 were aged 65 or above and all 5 were male. These negative scores for graphical control scales accounted for the views of 4 out of the total of 14 male respondents aged over 65 who had responded to the graphical control survey.

A significantly higher number of women (93%) rated 'yes' when asked whether they found the graphical control question style to be enjoyable compared to men (77%). This difference was not observed for the standard survey.

Scores for these engagement indices across respondents' personal internet usage levels (hours/per week) showed no significant differences within either survey. There were also no significant differences for age, gender or qualification classifications for the standard survey across all questions.

The final question on the surveys asked respondents to include any open text comments they had on the survey itself. Some 35 (33%) respondents from the graphical survey added a comment and only 21 (20%) from the standard survey, these figures again demonstrating a greater level of respondent engagement with the graphical survey. These comments were evenly provided by males and females across the age groups. An analysis of the comments is shown in Table 3 below.

Table 3: Analysis of open comments covering engagement topics

Comment Classification	Graphical	Standard
Interesting / entertaining / enjoyable style	15%	5%
Educating	6%	1%
Relevant topic / interesting subject	4%	3%
Comments about recycling / green issues in general	3%	2%
Importance of using the results	1%	-

Overall the comments conclude that there was a high level of interest and engagement with the graphical control survey questions and the subject area. The top four comments made by respondents from the graphical survey were firstly that it was an interesting/enjoyable/entertaining style (15%). Six of these 16 comments were from those aged 65 and over showing some level of appreciation of the graphical design by this age group. The second most commented topic was that the survey was educating (6%) and made them think about the topic. All the comments about educational value were made by women. Thirdly respondents described the survey as a relevant/topical/interesting subject (4%); again, these comments were made only by women. Some of the actual responses that demonstrate the main views are shown in Table 4.

Table 4: Engagement open text comments made about the graphical survey

**Some of the engagement advantages of the graphical control survey were seen to be:**

**Interesting/enjoyable/entertaining style**

"I thoroughly enjoyed answering the questions. It was an original and entertaining way to do it. Hope that more surveys come along this way."

"I have never done a survey which used these types of question styles before, even though I find the subject area interesting the styling also helped maintain interest."

"Fantastic interactivity one of the best I've taken part in."

"I love this type of survey; it makes it so much more enjoyable."

**Educating**

"This is a very important matter and we need to educate people!!! So this survey is a great thing to get out to people!"

"I think it's a good thing and really gets you thinking more about our everyday actions!"

"This survey has certainly made me think about what our family do and don't do to help the environment. It has actually made me feel quite guilty! Time for some changes I think!"

#### 4.2. Usability

Although the standard survey was scored slightly higher for being easy to respond to (91% versus 89% for the graphical survey) and for being quick to load (97% standard versus 92% graphical) these differences were not significant.

When asked about whether they needed to refer to the scale use guidance notes on the graphical survey, over two thirds (73%) of the respondents answered either 'yes' or

'somewhat'. Just over a quarter (27%) found the scales to be intuitive with no need for them to refer to the explanations.

Trends in the data show that respondents with qualifications from university/college were more likely to refer to the explanations on how to use the scales (43% said 'yes' of those with university/college qualifications (59) compared with only 24% of those without (46)).

Scores for these usability indices across respondents' personal internet usage levels (hours/week) showed no significant differences within either survey.

An analysis of the last question on the surveys which asked respondents to write any comments that they had about the survey (Table 5) demonstrated the importance of guidance notes for the graphical survey (mentioned by 4% of respondents). These were all respondents who had received college/university qualifications.

Table 5: Analysis of open comments covering usability topics

Comment Classification	Graphical	Standard
Importance of the guidance notes	4%	-
Good survey / good design	2%	6%
Difficulty / dislikes	2%	1%
Specific about scales (negative)	2%	5%
It's easy	1%	-

It is of interest to note that the standard survey had more negative comments made about the scales used than the graphical survey. These negative comments all referred to concerns that would also have been of relevance to the graphical survey such as the fixed format postcode question.

Some of the actual responses made that demonstrate the main views on the usability for the graphical survey are shown in Table 6.

Table 6: Usability open text comments made about the graphical survey

**Some of the usability disadvantages of the graphical control scales were described as:**

Need for guidance notes	Difficulties/dislikes	Specifics about the scales (negative)
"It did confuse me at first so the explanations helped."	"Slightly more complicated at times than other surveys."	"The 'temperature bulb' approach offered no improvement over older established methods."
"It took a moment or two to get used to the style. The guidance was very helpful."	"I feel that the style of the questionnaire was infantile - almost imbecile! As an experienced statistician I cannot see how any reliable results can be expected from it."	"On the two axis grid question there was no option for brands I was unaware of so therefore had to try and squeeze them all into a neutral space in the middle which is not such a good judge of true responses."
"Although some questions weren't immediately obvious about how to answer them, the accompanying description		[Note: respondents did not require to

helped."  
 "You do have to be careful  
 though to ensure you  
 understand what the scaling  
 means."

'click and drag' all the images. This  
 clearly demonstrates the importance  
 of highlighting 'no reply/don't know'  
 areas on graphic scales]

We gained a general view of the popularity (or 'easiness') of each question by undertaking an analysis of 'no replies' to each comparable question. This analysis demonstrated that overall there is not a great deal of difference in reply rates across the majority of questions within each survey. The majority of scales were completed by 98% of respondents and above. However there was a lower reply rate at 93% and 95% for two of the graphical scales and in each case 100% of respondents completed the comparative questions in the standard survey. Consideration should be given to make these questions more intuitive and simple to use for respondents in the future. The more difficult scales were:

Dual variable slider bar (100/105 respondents completed)

When asked how much volume of their household waste they recycled the graphical control survey provided respondents with a sliding bar that automatically calculated the relative percentages of recycled and non-recycled waste. The percent figures chosen were shown on the image (Figure 2).

Figure 2: The dual variable slider bar

**What proportion (volume) of your overall waste would you say is recycled and what is put in the refuse bin?**  
*Click & drag the scale to show the relative amounts.*



In contrast the standard survey required respondents to type the quantity of '% recycled' in a numeric box. The results from the standard survey show that all but two of the respondents typed in figures for '% recycled' that ended in either 0 or 5. As the graphical control slider bar enabled respondents to allocate % scores in increments of 1, the amounts inputted using this scale were more varied, however it was also more difficult to slide the bar accurately to a specific amount.

Opacity faded slider (98/105 respondents completed)

The opacity faded slider on the graphical control survey provided a gauge that enabled respondents to choose 'yes', 'not sure' or 'no' and any position of certainty/uncertainty across this scale (Figure 3). They were asked to use this scale to show whether they knew the difference between an environmentally friendly product and a fair-trade product. The levels of certainty/uncertainty chosen coincided with overlaid images of a tick, question mark or cross becoming either bolder or fainter. Analysing the scores showed that 23% chose 'yes' and no one chose 'not sure' or 'no'. The rest of the results were allocated in varying degrees of certainty between 'not sure' and 'yes' (53%) and 'not sure' and 'no' (16%). This contrasts with the standard survey where respondents had to choose one of the three fixed answers, here 52% ticked 'yes', 36% 'not sure' and 11% scored 'no'. The level of certainty of those feeling strongly that they do know the difference is clearer in the fixed choice option where 52%

ticked 'yes' however the truer picture is likely to be reflected in the graphical scale where only 23% were completely certain that they meant 'yes'.

Figure 3: The opacity faded slider

Would you say that you personally knew the difference between an environmentally friendly product and a fair-trade product?

Move the point on the scale to the far left for "yes" (a tick) and to the furthest right (a cross) for "no". The middle is for "not sure" (a question mark). You can move the pointer to show different levels of certainty for "yes", "no" and "not sure".



#### 4.3. Accuracy

In comparing the use of the scales across the two surveys there is an indication that respondents are more likely to use the upper end of the scale in some graphical control questions compared with the comparative standard scales. This results in a trend for higher mean scores in the graphical compared with the standard scales. Similar observations were made by Thomas and Couper, 2007 who found that for interactive visual analogue scales, end-anchored scales yielded higher mean values than fully anchored scales.

Examples of higher mean scores for the graphical scales are seen in the purchasing choice 'thermometer gauge' and the two dimensional brand perceptual mapping exercise.

The importance scale used for the 'thermometer gauge' question was an anchored sliding scale. Anchor point descriptions showing only when respondents moved the pointer along the scale (Figure 4). The comparative standard scale used was a fully anchored drop down box with scale descriptions.

The graphical control scores for the sliding 'importance' gauge resulted in respondents recording significantly higher mean scores for their importance to purchasing decisions than for those who used the drop down box in the standard survey (Table 6).

These results suggest that respondents using the graphical scales are more motivated to using the higher end of this graphical scale than those using the standard drop-down box scale. In effect they are forced to move the meniscus across all points of the scale in order to view the scale description at each level.

Figure 4: Grab-able thermometer meniscus

Click & drag the thermometer liquid to show the level of importance.

How important is the environmentally friendliness of a product, brand or service to your purchasing choice?

Important



How important to your purchasing choice of a product, brand or service is the support it provides for the communities and people that produce it?

An Absolute Must



How important is the level of trust that you have in a product, brand or service to your purchasing choice?

(no reply)



Table 6: Comparison of mean scores for importance in purchasing choices

Rating	Graphical	Standard
Environmental friendliness	3.3	2.9
Support for producer communities	3.2	2.9
Level of trust	3.8	3.5

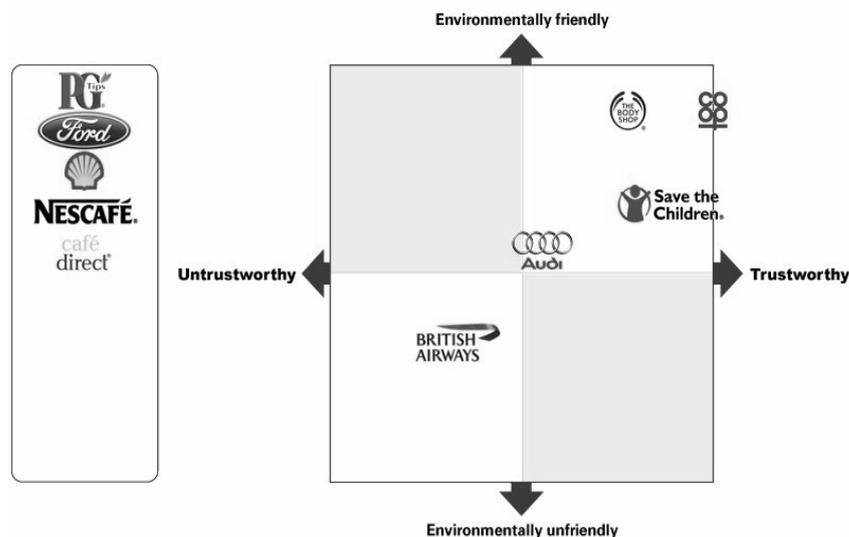
Note: A mean score was calculated by assigning a score of 5 for 'an absolute must', 4 for 'very important', 3 for 'important', 2 for 'of little importance' and 1 for 'unimportant'.

An analysis of the perceptual brand mapping for the respondents using a 'click and drag' graphical control two dimensional map (Figure 5) versus a standard radio button scale also showed that there was a tendency for those using the graphical control scale to use higher levels in the scale than those using the standard scale (see Figure 6).

Figure 5: Interactive brand mapping

Click & drag each of the images to a position on the square to show the level of trustworthiness and environmental friendliness. The centre of the image marks the actual position chosen. You can overlap images and also use the edges of the square as points for maximum or minimum scores.

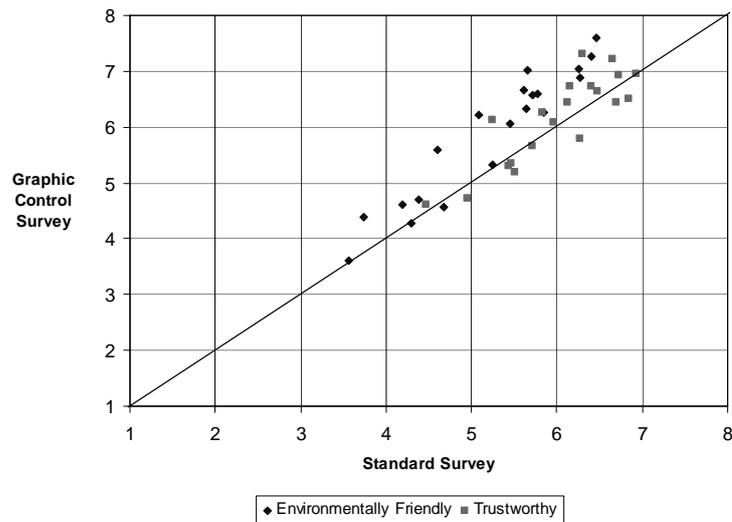
Please indicate how environmentally friendly you think each of the following brands/businesses are and how much you personally trust them.



The two dimensional brand mapping space used end-anchored bipolar scales, with the axes intersecting at the neutral midpoint. The standard 10 point radio button scale was end-anchored (e.g. completely untrustworthy/completely trustworthy) with an anchored neutral mid-point (e.g. neither untrustworthy nor trustworthy).

Whether the observed increased use of the higher end of the scale in these graphical scales is a more accurate reflection of the true value of a score compared to standard scales, or whether it is introducing a bias is open to debate and requires further work.

Figure 6: Score comparisons for graphical brand mapping and standard brand rating



## 5. Further work

We would see that our project has only scratched the surface of the possibilities that graphical input methods for questionnaires promises. There is, consequently, a vast amount of experimental work that could be carried out. In order to keep the work focussed we consider that there are two overall goals:

- Reduce potential set-up time, probably by introducing more standardisation into controls.

- Provide guidelines for questionnaire designers on the application of these controls.

As part of this we would specifically want to:

- Compare the use of different types of graphical control in the same situation with the aim of coming up with some heuristics as to which form of graphical control is appropriate when.

- Compare use of abstract images with the use of subject-related images. For example, if one is asking a question relating to proportion of time, which is a better way to represent time passing: an abstract scale, a container indicating percentage full, or a clock face?

## 6. Conclusions

We developed experimental surveys to explore the use of a variety of graphical input controls on web surveys. Controls we investigated included continuous and stepped sliders, and controls that involved the respondent dragging and dropping brand logos onto one- and two-dimensional image maps.

One fairly significant point to be made about graphical controls is how to design the control with a default response of 'no reply' and making it clear to the respondent that a response has not yet been given. This is especially true of slider controls where the slider needs to be somewhere in order for the respondent to recognise it. A similar problem arises with the use of image maps and the solution here seems to be to allocate a special place to where respondents drag icons to indicate 'don't know'.

Another potential problem is the novelty of graphical input methods for survey respondents. In our opinion, care does need to be taken to ensure that respondents know how to interact, either by making the control intuitively 'obvious' or by including appropriate instructions.

Currently the production of a graphical control survey requires more investment in development time and closer working with the developer, designer and researcher, however as graphical control scales become more standardised the additional time required on survey development will reduce.

Overall, we were encouraged that many respondents across all ages from the internet consumer panel found the graphical inputs acceptable, enjoyed completing the questionnaire and were looking forward to more surveys of this type in the future. The increased level of respondent engagement with the subject area through the use of interactive and colourful graphical scales is an important advantage over standard scales. We found no significant disadvantage of graphical scales for response rates or completion times. We therefore consider that the use of graphical controls such as those used in our survey could improve response rates and levels of interest from established consumer internet panels.

## References

- Couper, M.P., Tourangeau, R., Conrad, F.G., Singer, E. (2006), "Evaluating the effectiveness of visual analog scales: A web experiment", *Social Science Computer Review* 2006, 24
- Luetters, H., Westphal, D., Heublein, F. (2007), "*SniperScale: Graphical scaling in data collection and its effect on the response behaviour of participants in online studies*", German Online Research Conference (GOR)
- Thomas, R.K., Couper, M.P. (2007), "*A comparison of visual analog and graphic rating scales*", German Online Research Conference (GOR)

## Acknowledgements

The rest of the project team, all from Snap Surveys Ltd: Ralph Sutcliffe, Senior Researcher, Andrew Checkley, Web Software Developer and David Mitchell, Web/Image Designer.

## About the Authors

Dr Nicola Stanley is Managing Director of Silver Dialogue Ltd., a stakeholder engagement consultancy. She is a market research and communications specialist, with specific interest in novel approaches to motivate and empower individuals and groups through dialogue techniques. She has over 18 years' of experience gained through consumer/business research and corporate communication roles. She can be contacted at Silver Dialogue Ltd., PO Box 6470, Rushden, Northamptonshire, NN10 0ZW, UK, or by email [nicola\(at\)silverdialogue.com](mailto:nicola(at)silverdialogue.com).

Dr Stephen Jenkins is Technical Director of Snap Surveys Ltd. and is responsible for the direction and development of Snap survey software. He has over 30 years' experience in the design, development and use of survey software. He is a founding member of The Triple-S Group and an author of the Triple-S survey interchange standard. He can be contacted at Snap Surveys Ltd., Mead Court, Thornbury, Bristol BS35 3UW, UK, or by email [sjenkins\(at\)snapsurveys.com](mailto:sjenkins(at)snapsurveys.com).

Copies of the full surveys can be requested from the authors.